Rewriting the Politics of Publication: Tracking 25 Years of Debate in Academic Reform with Bibliometric Analysis

Tarid Wongvorachan * and Suchada Naknarin

Posted Date: 31 July 2025

doi: 10.20944/preprints202507.2644.v1

Keywords: academic publishing; scholarly communication; publishing reform; editorial gatekeeping; research assessment; knowledge equity; bibliometrics; bibliometric analysis; sci-hub; San Francisco declaration on research assessment

 doi:10.20944/preprints202507.2644.v1

Article

Rewriting the Politics of Publication: Tracking 25 Years of Debate in Academic Reform with Bibliometric Analysis

Tarid Wongvorachan [1,*] and Suchada Naknarin [2]

[1] Faculty of Education, Chulalongkorn University, Bangkok, 10330, Thailand
[2] Faculty of Medicine, Thammasat University, Pathum Thani 12120, Thailand
* Correspondence: wongvora@ualberta.ca

Abstract

Academic publishing is central to scholarly knowledge validation and dissemination, yet persistent issues such as editorial gatekeeping, prestige-driven evaluation metrics, and global inequities in knowledge access continue to shape the field. These challenges have significant implications for research integrity, inclusivity, and the equitable dissemination of educational scholarship. The conversation around these topics is on-going, but little is known about how these discussions have evolved within the scholarly literature. This study presents a three-fold bibliometric analysis of scholarly literature on academic publishing reform published between 2000 and 2025. The analysis examines the literature landscape on the topics of (1) Gatekeeping and Editorial Bias, (2) Prestige-driven Metrics and Research Assessment, and (3) Barrier and Equity Issues in Research Accessibility. Using structured searches in Scopus and bibliometric analysis techniques (i.e., performance and conceptual analysis), this study identifies key publication trends and thematic patterns within each area. Time-sliced analyses further explore how discourse has shifted in response to major milestones such as the San Francisco Declaration on research assessment and the rise of the sci-hub shadow library. By providing a comprehensive overview of how these critical conversations have evolved across global scholarship, this study contributes to a deeper understanding of the dynamics, thereby shaping the future of scholarly communication and offering insights to inform ongoing reform efforts for educators, researchers, and policymakers to foster a more inclusive, transparent, and equitable academic publishing landscape.

Keywords: academic publishing; scholarly communication; publishing reform; editorial gatekeeping; research assessment; knowledge equity; bibliometrics; bibliometric analysis; sci-hub; San Francisco declaration on research assessment

1. Introduction

Academic publishing lies at the heart of scholarly knowledge production, validation, and dissemination; It is the primary mechanism through which researchers communicate findings, establish credibility, and advance their careers [1]. Within education and across other disciplines, publication in peer-reviewed journals remains a critical marker of scholarly legitimacy and influence [1,2]. As such, the structures and norms of academic publishing significantly shape not only the flow of information but also whose voices are amplified, what kinds of knowledge are recognized, and how scholarly impact is assessed [1,3,4].

Despite its centrality, academic publishing continues to face longstanding and systemic challenges [5]. Among the most prominent are issues of editorial gatekeeping and bias, the dominance of prestige-driven research metrics, and enduring inequities in access to both publishing opportunities and scholarly knowledge [6–9]. These concerns have drawn increasing attention from researchers, institutions, and policymakers alike, sparking debates about the future of scholarly communication and the values that underpin it [7,10].

 doi:10.20944/preprints202507.2644.v1

One core challenge involves editorial gatekeeping and bias, which refer to the ways in which journal editors and peer reviewers serve as powerful arbiters of knowledge production [8]. Although peer review is designed to maintain academic rigor, it is also susceptible to implicit bias, ideological conformity, and institutional or geographic favoritism [11,12]. For example, journal editors may demonstrate preferential acceptance toward submissions from authors with certain characteristics (i.e., affiliated with prestigious institutions or personal familiarity), which is reflected in invite-only submission practices [13–15]. These dynamics can lead to the exclusion of marginalized perspectives, disciplines, or regions, especially those from the Global South or underrepresented groups, thereby reinforcing existing power hierarchies in academia [15,16].

A second area of concern centers on the prestige economy of academic publishing, which relies heavily on journal-based metrics such as the Journal Impact Factor (JIF) and h-index (i.e., the number of time cited of an author) to evaluate research quality and researcher performance [17]. These metrics, often criticized for their reductionism and lack of contextual nuance, have come to dominate funding decisions, hiring practices, and promotion criteria [18,19]. As a result, researchers may tailor their work to align with metric-based incentives, potentially privileging novelty and visibility over rigor, relevance, or social impact [20].

Third, the issue of access and equity in scholarly communication remains prominent in the scholarly publishing ecosystem. High article processing charges, subscription paywalls, and infrastructure disparities continue to limit who can publish, read, and build upon scholarly work [21,22]. While the open access movement has made significant inroads, access to academic knowledge remains uneven, particularly for under-resourced institutions and scholars in the Global South [23,24]. Initiatives like Sci-Hub and institutional repositories have challenged the commercial publishing model, but they also raise complex legal and ethical questions [25].

In response to these challenges, a range of reform efforts has emerged over the past two decades. Declarations such as the San Francisco Declaration on Research Assessment (DORA) [26,27] and the Leiden Manifesto [19] have called for more responsible and transparent approaches to research evaluation. Meanwhile, platforms promoting open science, preprints, and alternative metrics (altmetrics) have gained traction [28]. A number of open access initiatives have been introduced in recent years to address the persistent issue of academic paywalls to make scholarly knowledge more widely accessible to readers [29]. Yet despite growing momentum, the landscape of reform remains fragmented, with competing visions and uneven adoption across disciplines and regions.

While individual aspects of academic publishing reform have been widely discussed, the scholarly discourse remains dispersed across journals, disciplines, and national contexts. Few studies have systematically examined how this body of literature has evolved over time, particularly in response to major interventions like San Francisco DORA or the rise of Sci-Hub [30]. This limits our understanding of how the conversation around publishing reform has developed, who is driving it, and what themes are gaining prominence or fading from view.

To address this gap, the present study conducts a bibliometric analysis of literature on academic publishing reform from 2000 to 2025. Specifically, it maps scholarly discourse across three interrelated themes: (1) editorial gatekeeping and bias, (2) prestige-driven metrics and research assessment, and (3) barriers to and equity issues in research accessibility. Bibliometric methods provide a systematic way to trace publication trends, identify influential contributors, and analyze the thematic structure of scholarly conversations [31]. By offering a longitudinal view of how publishing reform has been discussed, this study contributes to a more comprehensive understanding of the dynamics shaping scholarly communication.

The central research question of this work is: *How has scholarly literature on academic publishing reform evolved across the themes of editorial gatekeeping and bias, prestige-driven research metrics and evaluation, and barrier and equity issues in research accessibility over the past 25 years?* By answering this question, the study aims to inform ongoing reform efforts and provide educators, researchers, and policymakers with insights to foster a more inclusive, transparent, and equitable academic publishing ecosystem.

 doi:10.20944/preprints202507.2644.v1

2. Literature Review

2.1. Overview of Academic Publishing's Role in Scholarly Communication

Academic publishing plays a vital role in scholarly communication as both a process and a system for validating, curating, and disseminating knowledge [32]. In the academic community, the term "publish" refers not only to the formal disclosure of a manuscript's content but also to the broader act of making information publicly accessible [32,33]. The publishing process officially begins when authors finalize their manuscripts and submit them to academic platforms, typically peer-reviewed journals or scholarly publishers [34]. To scholars, publishing is central to their academic identity and professional advancement through external recognition from peers [2,35].

The publication process itself is highly collaborative and structured, as described by Hyland [2] as follows: From the author's perspective, selecting an appropriate journal involves evaluating the journal's aims, scope, and previously published work, as well as adhering to submission guidelines. After the submission, authors must then monitor the editorial process, respond to reviewer comments, revise the manuscript as necessary, and approve the final proofs. Journal editors and publishers support this process by managing submission platforms, coordinating peer review, validating references, handling formatting, and preserving the final record [36]. This process displays the interdependence between authors, editors, and publishers, as well as reflects academic publishing's function as both a technical infrastructure and a system of scholarly communication.

Aside from knowledge dissemination, academic publishing serves as a quality assurance system. Peer review, the primary method of quality control through validation by other scholars, is designed to evaluate the rigor, reliability, and academic integrity of submitted work [2,37]. As discussed by De Vries et al. [37], peer review can be in single-blinded or double-blinded format and involve at least two reviewers and an editor who assess the manuscript's intellectual contribution, including originality, relevance to existing literature, clarity, argumentative quality, theoretical and practical implications, and alignment with a journal's editorial scope. These factors collectively determine whether a manuscript is accepted and ultimately help define what is considered quality academic work [38]. Lastly in the process, journal editors play a critical role in synthesizing reviewers' recommendations to make publication decisions, which shapes the direction of certain perspectives, topics, and research voices in academic dialogue [2].

The implication of academic publishing extends beyond knowledge record and dissemination. In today's data-driven environment, publishing track record plays a critical role in evaluations for hiring, promotion, tenure, and research funding [39]. One of the most influential tools in these evaluations is the JIF, widely used as a metric of journal prestige and scholarly productivity [2,40]. Scholars are incentivized to tailor their research to align with high-impact journals, as career advancement, recognition, and funding are often tied to metrics [39]. Academic publishing thus operates as both a regulatory system and a reward structure, commodifying scholarly output and equating impact with citation [2]. Simultaneously, university rankings increasingly rely on publication output, further entrenching the importance of publishing metrics [41]. According to Bradshaw and Brook [42], citation and impact scores not only influence journal rankings but also shape which voices are amplified in global scholarly discussions. In other words, based on the current publication system, the value of research is often perceived to lie not only in its dissemination, but also in the professional capital it generates.

Collectively, the mentioned dynamics show that publication is a crucial component for academic actors of all levels, which is why it is very crucial to make sure that the landscape of the academic publishing ecosystem is as equitable, inclusive, and representative to scholars of all backgrounds (e.g., race, ethnicity, institutional reputation, and career stage). This commitment to equity must be reflected across every facet of the ecosystem: from access to knowledge as the foundation for research production, to the editorial decisions that shape what contributions are recognized, and the evaluation processes that determine research value.

 doi:10.20944/preprints202507.2644.v1

2.2. Editorial Gatekeeping and Bias

Despite knowing how the publication ecosystem should be, that is ideal. The reality is far from ideal. Editorial gatekeeping refers to the decision-making processes through which journal editors, editorial boards, and peer reviewers determine which submissions are accepted for publication [2]; While these mechanisms are designed to uphold academic standards and ensure quality control, they also act as filters that shape the directions of academic discourse. In practice, editorial gatekeeping can introduce a variety of biases that reflect systemic asymmetries within academia. Scholars from underrepresented regions, non-native English-speaking backgrounds, or those affiliated with lesser-known institutions often face significant challenges in navigating this system, particularly when their work challenges dominant theoretical, methodological, or political paradigms [3]. This exclusionary practice is most commonly associated with high-ranking or 'top-tier' journals, where selective editorial practice creates a bottleneck that researchers seek to overcome in order to advance their careers [43]. Despite their training in scholarly evaluation, editors and reviewers are not immune to cognitive biases that can compromise the fairness of their assessments, which is an issue that lies at the heart of editorial gatekeeping dynamics.

Research has identified forms of bias such as favoritism toward prestigious institutions or well-known authors, confirmatory bias (where reviewers prefer work aligning with their own views) [44], positive publication bias (favoring studies with significant findings) [45,46], and even geopolitical or ideological interference (assuming that research from certain institutions are better or worse in general) [3]. Moreover, studies show that a large proportion of articles rejected by elite journals are eventually published elsewhere, suggesting that rejection often reflects misalignment with editorial or disciplinary preferences rather than intrinsic quality [47,48]. The prevalence of desk rejection, where manuscripts are dismissed without external peer review, can centralize opportunity to be heard within a limited editorial cohort, thereby amplifying existing biases and narrowing the trajectory of scholarly discourse [49]. Ultimately, peer review is a human process shaped by subjective judgments, disciplinary norms, and social dynamics. While efforts to standardize and anonymize the process exist, they cannot fully eliminate the influence of existing power structures and beliefs. As such, editorial gatekeeping, though central to scholarly publishing as a quality control measure, may perpetuate inequalities in academic visibility and recognition if left unchecked.

In response to these concerns, a range of reform initiatives has emerged, with open peer review gaining particular attention as a model that fosters transparency, accountability, and inclusivity. Key features of open peer review include open identities, publication of peer review reports, wider community participation, and preprint dissemination [50,51]. These developments represent a shift away from opaque, hierarchical review structures toward more democratic models of scholarly communication. Another similar development is the post-publication peer review model, where the peer evaluation process occurs after the manuscript is publicly disseminated. In this system, authors may identify potential reviewers and submit their names alongside the manuscript to the publisher [52]. The publisher then verifies the credentials of the reviewers and facilitates their evaluation, which becomes accessible as part of the manuscript's public record [52]. This model opens up opportunities for scholars to be heard while at the same time engages the manuscript in a quality assurance process similar to the traditional peer review practice. Complementing these efforts is a growing emphasis on research integrity and publication ethics, which is outlined in the UK Concordat to Support Research Integrity [53,54]. The concordat especially emphasizes honesty, fairness, transparency, and accountability throughout the life cycle of research, including the peer review and editorial aspect [53].

Together, these reforms seek to counteract gatekeeping bias, wherein editorial decisions may be shaped by implicit norms or institutional preferences, leading to the exclusion of novel, interdisciplinary, or regionally diverse perspectives. Addressing such bias is not merely a procedural concern but an ethical imperative to uphold the values of equity and openness in scholarly publishing. As gatekeeping is deeply associated with questions of prestige and recognition, the following section discusses the role of evaluation metrics in reinforcing these structural dynamics.

 doi:10.20944/preprints202507.2644.v1

2.3. Prestige-Driven Metrics and Research Assessment

Over the past decades, bibliometric indicators such as the JIF, h-index, and citation counts have become dominant tools for evaluating scholarly productivity and influence [2,40]. Initially developed to track the dissemination and usage of research, these metrics have evolved into high-stakes proxies for academic quality, frequently used by institutions in hiring, tenure, promotion, and funding decisions [55,56]. A study by McKiernan et al. [55] revealed that 40% of universities with a strong research orientation and 18% of master's-level institutions utilize JIF in their promotion decision. However, the emphasis on bibliometric measures has also incentivized scholars to prioritize publishing in high-impact journals and focus on increasing such metrics, which may not always align with research quality or innovation [57]. Reflecting these pressures, a large-scale survey by D'Souza et al. [58] was administered with over 5,000 authors to highlight key challenges faced by authors in the academic publishing ecosystem; The survey revealed that JIF was used as the most important element in journal selection. These developments of metric-oriented culture illustrate how bibliometrics have shaped academic publishing. While these metrics help track certain aspects of scholarly output, they can also cause problems such as encouraging researchers to prioritize boosting scores instead of producing work that's meaningful or useful to real audiences, as these metrics are use as proxies for prestige of the work itself [58,59]. In some countries, researchers receive financial rewards for publishing in international journals. As a result, early-career academics often feel pressured to focus on getting their work published at specific venues, sometimes at the cost of other important roles like teaching and mentoring [2].

Ironically, the current use of metrics like JIF is distorted from the original intention of its creator [40]. The growing reliance on these prestige-driven metrics has drawn considerable critiques. These metrics, while intended to measure scholarly impact, have created perverse incentives that encourage unethical practices such as excessive self-citation, citation cartels, coercive citation during peer review, and the manipulation of authorship through gift authorship or credit slicing across large teams [60]. Studies show that such behaviors distort the academic record and foster new forms of misconduct, including citation rings, manipulated peer review, and the rise of predatory publishing [20]. Quantitative analyses reveal significant disparities in self-citation practices across countries, suggesting cultural or policy-driven motivations behind citation inflation [61]. Institutions, especially those seeking rapid advancement in global university rankings, have been found to adopt metric-driven strategies that may signal manipulation of metrics such as lead or corresponding author count, STEM output, and publication number [62]. These practices are further reinforced by funding models like the UK's Research excellence framework, where journal rankings and quasi-exponential scoring systems disproportionately reward publications in elite, monodisciplinary journals, thereby systematically disadvantaging interdisciplinary and non-English or locally focused research [63]. As academic evaluation shifts toward a corporate model that emphasizes performance metrics, scholars face growing pressure to align their work with citation-rich fields, often at the cost of innovation, diversity, and public trust [64]. To address these systemic issues, improved oversight, transparent bibliometric tools, and a redefinition of academic excellence beyond numerical indicators are urgently needed.

In response to growing concerns about the misuse of traditional bibliometric indicators (e.g., JIF or h-index) numerous reform movements have emerged advocating for more responsible, inclusive, and context-sensitive approaches to research evaluation. One major initiative is the San Francisco DORA, first declared in 2012 and formalized in 2013, which calls for de-emphasizing journal-based metrics in favor of assessing research on its own merits [27]. DORA encourages transparency, recognition of diverse research outputs (such as datasets, software, and societal impact), and discourages the use of JIF to evaluate individual researchers or as promotional tools for publishers [27]. The DORA also urges researchers to assess research based on its content, rather than its metrics [27]. Complementing this, the Leiden Manifesto, published in Nature, outlines ten principles for the responsible use of metrics, including the combination of qualitative expert review with quantitative indicators, recognition of disciplinary differences, and greater support for locally relevant or non-English research [19,65,66].

 doi:10.20944/preprints202507.2644.v1

Alongside these frameworks, altmetrics have emerged as a complementary tool to capture the broader impact of research through online mentions, media coverage, and policy uptake, though concerns remain about platform dependency and data consistency [28]. Institutional initiatives, such as the coalition for advancing research assessment, are also gaining momentum by embedding these principles into policy and practice. For example, a Canadian Delphi study identified 50 indicators aligned with DORA such as methodological rigor, equity, policy relevance, and community engagement [67]. Collectively, these movements seek to reduce systemic bias, foster diversity, and restore integrity in scholarly communication by redefining excellence through transparent, context-aware, and inclusive assessment systems. These critiques not only challenge the dominance of numerical indicators but also expose how prestige-based evaluation can reinforce barriers to inclusion and knowledge accessibility through commodification.

2.4. Barrier and Equity Issues in Research Accessibility

While academic publishing is central to scholarly communication, it has also become a site of deep inequality, shaped by commercialization and global disparities. In recent decades, the consolidation of publishing power among a few large commercial entities has led to an increasingly profit-driven model, where access to both publishing and reading scholarly work often hinges on institutional wealth [68]. In fact, affordability in publication is also a main concern among research authors as revealed by D'Souza et al. [58]'s study. The academic publishing landscape has evolved into an oligopoly dominated by high-profit commercial publishers such as Elsevier and Springer Nature [69]. These publishers leverage market power, bundled contracts, and article processing charges (APCs) to generate substantial revenue, often at the expense of universal access and equity. van Bellen et al. [70] analyzed large-scale bibliometric data and reveal that despite some growth among small and regional publishers, particularly in the humanities and social sciences, major publishers still dominate the global output across multiple databases including, Web of Science (WoS) and OpenAlex. Björk [71] explains how low market competition and "big deal" bundling strategies allow these publishers to maintain high profits, which in turn inflate subscription costs and severely constrain library budgets. As a result, a small number of commercial publishers hold enormous influence over the scholarly communication ecosystem, entrenching systemic inequality. A testament to this oligopoly concern became more prominent when an antitrust lawsuit was filed to such publishers, alleging that they took advantages against scholars through actions such as exclusivity enforcement (i.e., single submission policy), zero peer-review remuneration, and forced transfer or copyright after acceptance [72]. Market concentration, combined with opaque pricing, continues to restrict access and visibility, especially for scholars in under-resourced institutions and regions. However, meaningful transformation will require challenging entrenched commercial incentives and redefining the values that underpin academic publishing [9,70,71].

Although open access publishing removes paywalls for readers, it often shifts the financial burden to authors through APC instead, thereby creating a different form of accessibility barrier [29]. These challenges are especially acute for scholars from low- and middle-income countries and less-funded disciplines, who often must pay out-of-pocket or forgo publication opportunities altogether [73,74]. High APCs, particularly in prestigious journals, contribute to a stratification effect in academic publishing, enabling well-funded institutions to dominate high-impact open access venues while marginalizing under-resourced scholars [75]. This barrier therefore contributes to the system that privileges wealthier researchers and institutions [76]. Meanwhile, the high cost of journal subscriptions continues to restrict access for many academic libraries, independent researchers, and the general population, reinforcing global knowledge silos. Additionally, the funds allocated to cover APCs may divert resources from other essential research needs such as staffing, travel, or equipment, ultimately limiting the broader development of academic work [77,78].

As barriers to accessing scholarly knowledge persist, particularly for those in under-resourced institutions and regions, alternative access models have proliferated to challenge the dominance of commercial academic publishing. Several open access models were developed as follows: Green

 doi:10.20944/preprints202507.2644.v1

open access (i.e., self-archiving by authors in institutional or subject repositories, typically after an embargo period) allows for free dissemination of preprints or postprints without requiring payment to publishers [79]. Hybrid open access (i.e., subscription journals offering individual articles as open access for a fee) provides partial openness but has been criticized for "double dipping", charging both subscription fees and APC [80]. Gold open access (i.e., full and immediate access provided by journals, often funded by APCs paid by authors or institutions) enhances visibility but can be financially exclusionary for researchers in low-resource settings [79]. Diamond open access (i.e., journals offering immediate and free access without APCs, supported by academic institutions or public funding) represents the most equitable model, though its scalability and sustainability remain under scrutiny [81].

Additionally, illicit shadow libraries like Sci-Hub have also emerged in response to paywalls and restrictive publishing structures. Founded in 2011 by Alexandra Elbakyan, dubbed as Science's Pirate Queen, Sci-Hub bypasses publisher restrictions to provide free access to millions of paywalled academic articles, surpassing even what the legitimate route can provide in some universities [82]. For many researchers, especially in the Global South, Sci-Hub has become a vital resource, reflecting both desperation in the face of unaffordable subscriptions and resistance to the commodification of publicly funded research [83,84]. Despite repeated legal challenges and domain shutdowns, Sci-Hub has continued to operate through alternative domains and mirror sites, demonstrating its resilience and the persistent demand for unrestricted access to scholarly literature [85]. However, access inequalities still also persist, not only in terms of financial barriers but also in structural imbalances such as English-language dominance, geographic centralization of editorial authority, and the privileging of institutional prestige, all of which shape whose knowledge is legitimized within global academia [86]. In response, reform initiatives such as Plan S [87], Project DEAL [88], and Digital Research Infrastructure for the Arts and Humanities [80] have called for and took effort in initiating structural shifts by emphasizing principles of openness, transparency, and inclusivity to dismantle paywall-based models, promote fairer evaluation practices, and encourage the sustainable dissemination of knowledge.

3. Present Study

The preceding literature reveals three persistent and interconnected challenges in academic publishing: editorial gatekeeping and bias, prestige-driven metrics and their influence on research behavior, and enduring inequities in access to both publishing and scholarly knowledge. While these issues have been widely discussed across scholarly communities, much of the discourse remains fragmented, often focusing on isolated aspects or individual reforms. What is less understood is the landscape of academic literature pertaining to these matters, as well as how they evolve over time, particularly in response to major interventions such as the San Francisco DORA or the emergence of Sci-Hub.

To address this gap, the present study adopts a bibliometric approach to systematically map the development of scholarly conversations on publishing reform from 2000 to 2025. Bibliometric analysis offers a robust method for uncovering patterns in publication activity, thematic evolution, and the structural relationships within the literature. Therefore, this study aims to answer the research question of *How has scholarly literature on academic publishing reform evolved across the themes of editorial gatekeeping and bias, prestige-driven research metrics and evaluation, and barrier and equity issues in research accessibility over the past 25 years?* The following section outlines the methodological design employed to answer this question.

4. Method

4.1. Document Search and Collection

Bibliometric metadata of scholarly work regarding the three topics (i.e., Gatekeeping and Editorial Bias, (2) Prestige-driven Metrics and Research Assessment, and (3) Barrier and Equity Issues in Research

 doi:10.20944/preprints202507.2644.v1

Accessibility) from 2000-2025 was collected from the the Scopus database. Scopus was selected as the sole data source over other prominent databases such as WoS due to its more comprehensive and detailed metadata offerings. Specifically, Scopus allows for the extraction of rich bibliometric information in CSV format, including citation details (e.g., authors, title, year, citation count, document type, and open access status), bibliographic data (e.g., affiliation, publisher, and serial identifiers like ISSN), and content-related elements such as abstracts and both author-provided and indexed keywords. In addition, Scopus includes other valuable fields such as funding information, conference details, and references when available.

In contrast, WoS provides significantly less detailed metadata, typically limited to authors, titles, sources, and abstracts. Although it is technically possible to merge datasets from Scopus and WoS, doing so introduces complications due to the incompatible structure and differing metadata standards between the two platforms. For example, discrepancies in total citation counts and variations in keyword indexing methods, such as Scopus using "Author Keywords" versus WoS using "Keyword Plus", can undermine the consistency and reliability of merged data. Moreover, the absence of many metadata fields in WoS affects the completeness of the dataset and can potentially distort analysis results and their interpretations. Given these considerations, Scopus was chosen for its superior metadata richness and higher export limit, allowing the extraction of up to 5,000 records at once, compared to WoS's capacity of 1,000 records export. These advantages make Scopus the more suitable choice for conducting a robust bibliometric analysis in this study.

4.2. Search Strategy and Dataset Construction

The bibliometric search was conducted in Scopus for the period 2000 to 2025 using a structured and systematic strategy. Searches were limited to records in English and were conducted within the fields of article title, abstract, and keywords. Only peer-reviewed literature types of articles, editorials, reviews, book chapters, conference papers, and books were included. Preprint citation indexes were excluded. The data collection took place between June 9th and June 19th, 2025.

The search was conducted separately for each of the three core research topics by breaking them down into more specific subtopics to ensure thorough coverage of the database. For example, the topic of Gatekeeping and Editorial Bias was divided into three distinct search blocks: (1) Gatekeeping and Editorial Control, (2) Bias in Peer Review and Inclusion, and (3) Hierarchy and Prestige Barriers in Publishing. Each subtopic was searched independently, and the resulting datasets were later merged. Duplicate entries were identified and removed using DOI matching to ensure dataset integrity.

For the second and third research topics, the datasets were further segmented into two time periods to allow for temporal analysis based on significant historical developments. The topic of Prestige-Driven Metrics and Research Evaluation was split into two periods: 2000–2013 (pre-DORA) and 2014–2025 (post-DORA), referencing the launch of the San Francisco DORA as a pivotal moment in the change of evaluation practices. In contrast, the Barrier and Equity Issues in Research Accessibility dataset was divided into 2000–2011 (pre-Sci-Hub) and 2012–2025 (post-Sci-Hub), recognizing Sci-Hub's disruptive role in research accessibility. Note that despite other shadow libraries such as Library Genesis and Z-Library exist, they were excluded from this study due to their broader content scope, which includes general-interest books, images, comics, audiobooks, and magazines. Table 1 presents the complete search strings used for each topic and subtopic, including the specific time slices for an analysis of the evolution of scholarly discourse before and after key events associated with each thematic area.

 doi:10.20944/preprints202507.2644.v1

Table 1. Search strings and time slice period of the study.

#	Topic	Key Concept	Search String	Time Slicing
1	Gatekeeping and editorial bias	Editorial gatekeeping, Editorial subjectivity / bias, Biased desk rejection, Biased, peer review process, Confirmation bias in peer review, Positive publication bias, Inclusion/exclusion of replication or negative results, Hierarchical publishing / prestige bottlenecks	Block A: Gatekeeping & Editorial Control Scopus: ("editorial gatekeeping" OR "editorial bias" OR "editorial subjectivity" OR "editorial decision*" OR "insider bias" OR ("desk reject*" AND (bias OR subjectivity OR favoritism OR transparency))) 731 documents found on Scopus Block B: Bias in peer review / inclusion ("peer review bias" OR "publication bias" OR "positive results bias" OR "positive publication bias" OR "null results exclusion" OR "negative results exclusion" OR "negative findings exclusion" OR "null findings exclusion" OR "file drawer problem") AND ("peer review" OR "academic publishing" OR "scientific communication" OR "research dissemination") 516 documents found on Scopus Block C: Journal Prestige barriers "prestige journals" OR "hierarchical publishing" OR "publishing hierarchy" OR "publication hierarchy" OR "journal prestige" OR "publication bottleneck" 191 documents found on Scopus	No time slicing
2	Prestige-driven metrics and research evaluation	Overemphasis of publication metrics (e.g., impact factor, citation counts, journal quartiles. Distorted research evaluation. Hypercompetition in publishing, "Natural selection of bad science", distorted publication practice (e.g., p-hacking, selective reporting).	Block A: Metrics overreliance ("impact factor" OR "journal impact factor" OR "journal quartile" OR "journal rankings" OR "citation-based metric*" OR "research metric*") AND ("overreliance" OR "misuse" OR "abuse" OR "limitations" OR "criticism" OR "flaws" OR "problem*" OR "bias*" OR "gaming" OR "inappropriate") 3571 documents found on Scopus Block B: Research assessment & evaluation "research assessment fairness" OR "research evaluation" OR "faculty evaluation fairness" OR "academic evaluation fairness" OR "promotion and tenure" OR "RPT" OR "review promotion tenure" 1704 documents found on Scopus Block C: Problems / distortions caused by metrics "natural selection of bad science" OR "p-hacking" OR "selective reporting" OR "salami slicing" OR "research malpractice" OR "reductionist evaluation" OR "metrics distortion" OR "prestige-driven publishing" 1812 documents found on Scopus Block D: Broader discourse documents "DORA declaration" OR "San Francisco Declaration on Research Assessment" OR "Leiden Manifesto" OR "responsible metrics" OR "responsible research assessment" OR "responsible evaluation" 130 documents found on Scopus	2000–2013 = pre-DORA/Leiden 2014-2025 = post-DORA/Leiden adoption
3	Barrier and equity issues in research accessibility	Commercialization of publishing and academic publication oligopoly. Paywalls from subscription-based models. Accessibility problem from article processing charges (APCs). Open access as an inequitable partial solution. Regional and institutional inequities in access. Sci-Hub as a symptom of access issues. Global disparities in scholarly publishing participation	Block A: Publishing Commercialization & oligopoly "commercial publishing" OR "academic publishing oligopoly" OR "publishing oligopoly" OR "commercial publishers" OR "academic publishing industry" OR "publishing market concentration" 536 documents found on Scopus Block B: Barriers to access "academic journals paywall*" OR "subscription-based journal*" OR "subscription cost*" OR "restricted research access" OR "barriers to research access" OR "access to research" OR "research access" AND ("inequit*" OR "disparity" OR "librar*") 318 documents found on Scopus Block C: Open access challenges & APCs "article processing charges" OR "cost of open access" OR "open access inequity" OR "APC waiver" OR "affordability of open access" OR "predatory open access" 553 documents found on Scopus Block D: Global equity & participation "equity in scholarly publishing" OR "global inequities in publishing" OR "research accessibility" OR "knowledge equity" OR "research participation barriers" OR "low-income country publishing" OR "underfunded institutions" OR "global South research access" 143 documents found on Scopus Block E: Sci-Hub and shadow libraries "Sci-Hub" OR "shadow library" OR "pirate library" OR "illegal access to research" OR "circumventing paywalls" OR "Sci-Hub citation impact" OR "Sci-Hub and research access" OR "black open access" OR "shadow libraries" 155 documents found on Scopus	2000-2011 = pre Sci-hub 2012-2025 = post Sci-hub

4.3. Bibliometric Analysis

Bibliometric analysis, which is the quantitative evaluation of publication patterns, research output, and scholarly impact within a specific field, was conducted using the Bibliometrix package in R [89]. Two primary analytical approaches were employed: performance analysis and conceptual analysis. Performance analysis was performed using the biblioAnalysis function to assess the productivity and influence of various research actors, including institutions, countries, and sources (e.g., journals or conferences). This analysis provided insights into publication output, scientific production trends, and the identification of dominant countries regarding discourses of the topic. Note that while author name is a standard component of performance analysis, it is not emphasized in this study, as it offers limited meaningful insight into the overarching research objective (i.e., understanding the structural and thematic landscape of the literature). Additionally, highlighting individual authors may invite unnecessary or unsolicited attention, which this study aims to avoid.

Conceptual analysis was conducted to explore the thematic structure and intellectual development within each research topic. Author keywords served as the primary data source for this analysis. Keyword co-occurrence networks were constructed using the biblioNetwork function and visualized with the Fruchterman-Reingold layout algorithm to reveal structural relationships and thematic clusters within the field. Further, conceptual cluster analysis was executed using the conceptualStructure function, applying the Multiple Correspondence Analysis (MCA) method. MCA is a statistical technique that reduces the dimensionality of categorical data to identify meaningful groupings of related terms [90]. This analysis was configured with a minimum term occurrence (minDegree) of 4, five clusters (clust = 5), and a minimum of 10 documents per cluster, ensuring the thematic integrity and interpretability of the clusters.

A thematic map analysis was then performed using the thematicMap function to examine the developmental trajectory of research themes [91]. The analysis included the top 250 terms (n = 250)

 doi:10.20944/preprints202507.2644.v1

and used a minimum frequency threshold (minFreq) of 4 per thousand documents. The resulting map classifies clusters into four quadrants, each representing different stages of thematic maturity:

- Emerging or Declining Themes (low centrality, low density): This quadrant represents topics that are either underdeveloped or losing relevance, often characterized by weak integration with other themes and limited research activity.
- Niche Themes (low centrality, high density): This quadrant represents Well-developed but isolated topics with strong internal coherence and limited external linkage.
- Basic Themes (high centrality, low density): This quadrant represents foundational topics with wide relevance but lower development, often serving as conceptual anchors for the field.
- Motor Themes (high centrality, high density): This quadrant represents both well-developed and highly connected topics, indicating the forefront of the research domain.

For topics subjected to time-sliced analysis, the thematicEvolution function was employed to visualize and quantify thematic changes over time [91]. The configuration is similar to that of the thematic map (N = 250, minFreq = 4). The evolution was interpreted by examining keyword continuity and directional arrows between time periods. When arrows connect themes across time, they signal sustained relevance and thematic development. Conversely, the absence of a connecting arrow suggests either the emergence of a new theme or the obsolescence of a prior one. Additionally, convergence of multiple earlier themes into a single later theme suggests interdisciplinary integration or the coalescence of research areas into a unified topic.

To enhance the thematic relevance and clarity of the conceptual analyses, a rigorous keyword cleaning process was undertaken. Terms that did not meaningfully contribute to the core subject matter, such as Gatekeeping and Editorial Bias, were systematically excluded, even if they appeared as prominent keywords in the dataset. This step was critical to eliminate noise and prevent the distortion of thematic structures by unrelated or overly broad topics. For instance, in the conceptual analysis of gatekeeping and editorial bias, the following terms were removed: “news production,” “online journalism,” “journalism,” “news,” “newspapers,” “online news,” “social media”. Although some of these terms were frequently used, they represented tangential or unrelated domains, and their inclusion would have diluted the thematic precision of the analysis.

Additionally, synonymous and closely related terms were consolidated to reduce redundancy and improve the interpretability of the conceptual structure. This consolidation was necessary to avoid fragmentation of semantically similar ideas across multiple clusters, which could obscure the true conceptual landscape of the field. For example, “publication,” “publishing,” “academic publishing,” and “scholarly publishing” were treated as a single unified concept. Similarly, “journal prestige” and “prestige” were merged, as were variations of “journal” such as “academic journals,” “academic journal,” “journal,” and “journals.” Other merged pairs included “editorial decisions” and “editorial decision.” These standardization efforts ensured a more coherent representation of dominant themes and helped preserve the analytical integrity of the conceptual mapping.

Despite deliberate efforts to remove irrelevant keywords and consolidate synonyms, some residual or synonymous terms may still appear in the results. In certain cases, seemingly unrelated terms may appear because they reflect the disciplinary context of the analyzed documents. For instance, a paper addressing publication bias may originate from the clinical sciences, thereby introducing domain-specific clinical terminology into the conceptual clusters. This trade-off was deemed acceptable, as it preserves the broader conceptual alignment of the dataset while minimizing disruptions to thematic coherence and interpretability. Additionally, some synonymous terms may persist in keyword-dense results due to the limitations of manual synonym consolidation. Given the complexity involved in this process, it is plausible that some redundancies were overlooked. However, to mitigate the impact of such inconsistencies, the interpretation of conceptual analysis results was based on the triangulation of multiple methods (e.g., keyword co-occurrence network, MCA, and thematic map). This approach ensured that interpretations were grounded in converging patterns across analyses, rather than being skewed by any single artifact or missed keyword.

 doi:10.20944/preprints202507.2644.v1

Appendix A presents the full list of removed terms and terms consolidated as synonyms. Finally, the results of the conceptual analysis are interpreted qualitatively by drawing on insights from relevant literature. These findings are then integrated with the performance analysis results to construct a comprehensive view of the scholarly landscape across the three core topics. For Topics 2 and 3, this integrated analysis also enables the identification of key changes in thematic focus following major field-changing events (e.g., the emergence of Sci-hub).

5. Results

5.1. Performance Analysis

Results from the performance analysis are presented first, as they offer a foundational overview of the research landscape, including scientific productivity, key sources, and frequently appeared keywords. This initial overview provides context for interpreting the findings from the subsequent conceptual and thematic analyses. In this analysis, we opted to use Author Keywords rather than Keyword Plus. While both indexing terms are commonly employed in bibliometric mapping, Author Keywords are typically more context-specific, as they are selected by the authors to reflect the core themes and subject matter of their research. In contrast, Keyword Plus terms are algorithmically generated based on the titles of cited references, which can result in less precise or contextually relevant descriptors [92]. As such, Author Keywords are better suited for capturing the intentional focus and domain-specific language used by researchers, thereby enhancing the accuracy and interpretability of conceptual analyses [92]. This study analyzed a total of 10,101 scholarly documents across three core themes related to academic publishing reform: Gatekeeping and Editorial Bias (N = 1,420), Prestige-driven Metrics and Research Evaluation (N = 7,044), and Barrier and Equity Issues in Research Accessibility (N = 1,637). These sample sizes represent de-duplicated results obtained from structured searches conducted in the Scopus database. Table 2 presents a comparative summary of the performance analysis for each topic, highlighting key metrics such as publication output, growth rate, source distribution, and top keywords.

Table 2. Performance analysis results of the three topics.

	Topic 1: Gatekeeping and editorial bias	Topic 2: Prestige-driven metrics and research evaluation	Topic 3: Barrier and equity issues in research accessibility
N of analyzed documents (Total document - duplicates)	1420 (1438/18)	7044 (7217/173)	1637 (1705/68)
Annual growth rate %	10.7	9.08	8.09
Average citations per year per document	2.955	3.588	1.583
Document Types (N)	article (880) book (16) book chapter (61) conference paper (47) editorial (157) review (259)	article (4558) book (24) book chapter (197) conference paper (848) editorial (165) review (1252)	article (1138) book (17) book chapter (102) conference paper (129) editorial (49) review (202)
Maximally and minimally productive year (N of publications)	2019 (158) 2000 (5)	2022 (600) 2000 (32)	2024 (171) 2003 (10)
Top corresponding author's country (N of documents)	USA (365)	USA (1096)	USA (399)
Top 5 sources	1. Review of financial studies 2. PLOS one 3. Scientometrics 4. BMJ Open 5. Review of corporate finance studies	1. Cochrane database of systematic review 2. PLOS one 3. Scientometric 4. Journal of clinical epidemiology 5. Research evaluation	1. Learned publishing 2. Insights: The UKSG journal 3. Scientometrics 4. Publications 5. Serial librarians
Top 5 keywords	1. Peer review 2. Publication bias 3. Bias 4. Prestige bias 5. publication	1. Impact factor 2. Research evaluation 4. Bibliometrics 4. Publication bias 5. P-hacking	1. Article processing charge 2. Scholarly publishing 3. Open access 4. Publishing 5. Sci-hub

Note. For the top 5 keyword components, keywords that share synonymous meaning as explained above were merged. Keywords that are not relevant to the subject matter were also excluded as mentioned above.

5.1.1. Gatekeeping and Editorial Bias

The dataset for Topic 1 comprised 1,420 unique documents, extracted from an initial pool of 1,438 records. Although it represents the smallest dataset among the three topics, it exhibited the highest annual growth rate at 10.7% over the past 25 years. The most productive year was 2019, with 158 publications, compared to only 5 publications in 2000. This trend highlights an increase in attention to the editorial gatekeeping issue in the last decade. The majority of documents were journal articles (n = 880), followed by review articles (n = 259) and editorials (n = 157). This distribution suggests that the discourse is both empirical and reflective, which is consistent with the normative concerns often raised in peer review reform and editorial transparency. On average, documents in this topic received 2.955

 doi:10.20944/preprints202507.2644.v1

citations per year. Geographically, the United States led in corresponding authorship (n = 365), which highlights its dominant position in editorial policy discourse and journal management.

The most prominent sources included Review of Financial Studies, PLOS ONE, Scientometrics, and BMJ Open; This distribution reflects a combination of disciplinary application and meta-level inquiry. The top recurring keywords were peer review, publication bias, bias, prestige bias, and publication, which collectively point to persistent concerns about fairness, credibility, and the influence of institutional prestige in scholarly publishing. These keywords highlight ongoing critiques of the legitimacy and inclusivity of current editorial systems.

5.1.2. Prestige-Driven Metrics and Research Evaluation

The second dataset, focusing on prestige-based evaluation metrics and research assessment practices, was the largest of the three, comprising 7,044 unique documents after removing duplicates from an initial pool of 7,217. This topic showed an annual growth rate of 9.08%. The most productive year was 2022, during which 600 documents were published, which reflects increased global scrutiny of performance indicators such as journal impact factors, citation counts, and university rankings. The dominant document type in this dataset was the journal article (n = 4,558), accompanied by a substantial number of reviews (n = 1,252) and conference papers (n = 848). The volume of conference proceedings suggests that discussions around research evaluation are both academically active and responsive to evolving policy frameworks. Editorials (n = 165), book chapters (n = 197), and books (n = 24) were also present, indicating some degree of conceptual and theoretical exploration beyond empirical analysis.

This topic recorded the highest average citation rate across the three datasets at 3.588 citations per document per year, reflecting substantial academic engagement. Similarly to the first topic, the United States was the top country for corresponding authorship (n = 1,096), reinforcing its strong institutional and policy presence in shaping research evaluation discourse. Key publishing venues included the Cochrane Database of Systematic Reviews, PLOS ONE, Scientometrics, Journal of Clinical Epidemiology, and Research Evaluation. These journals represent a mix of methodological work, bibliometric-level analysis, and policy engagement. The most frequent keywords were impact factor, research evaluation, bibliometrics, publication bias, and p-hacking. These terms reflect central themes such as the quantification of scholarly value, the limitations of traditional metrics and its entailing problem.

5.1.3. Barrier and Equity Issues in Research Accessibility

The third dataset, centered on barriers to research access and equity in knowledge dissemination, has the lowest growth rate of 8.09% among the three topics. The most productive year was 2024, during which 171 documents were published. As with the other topics, journal articles were the predominant document type (n = 1,138), followed by reviews (n = 202) and book chapters (n = 102). This topic recorded the lowest average citations per year per document at 1.583, possibly due to the cross-disciplinary nature of accessibility discussions and slower citation cycles in fields such as library science, education, and development studies. The United States also led in corresponding authorship (n = 399), though the topic itself likely features more global representation as it could be more relevant to researchers in the global south communities as they are directly affected by the equitable access issue [24].

Leading publication venues included Learned Publishing, Insights: The UKSG Journal, Scientometrics, Publications, and Serial Librarian. These sources reflect a mix of scholarly publishing, library science, and open access advocacy. The most frequent keywords of article processing charge, scholarly publishing, open access, publishing, and Sci-Hub highlight core concerns around the economic and legal barriers to accessing research, as well as grassroots responses to these barriers. The appearance of Sci-Hub in particular shows the tension between formal publishing infrastructures and informal access mechanisms.

 doi:10.20944/preprints202507.2644.v1

5.1.4. Time Sliced Analysis: Prestige-Driven Metrics and Research Evaluation Before and After DORA

Table 3 displays and juxtaposes performance analysis results of the second topic: Prestige-driven metrics and research evaluation before and after the emergence of San-Francisco DORA. The pre-DORA period (2000–2013) yielded 1,664 documents, with a relatively high annual growth rate of 18.35%, indicating a rapid early-stage expansion of discourse on research evaluation and metrics. Scholarly output reached a pre-event peak in 2013 (286 documents). The average citations per document per year was 3.572, already showing substantial engagement. Most publications were journal articles (n = 1,004), but conference papers (n = 300) and reviews (n = 277) also made significant contributions, suggesting active debate and methodological development. The USA led as the top country for corresponding authors (n = 296). Prominent journals included PLOS ONE, Scientometrics, and the Cochrane Database of Systematic Reviews. Top keywords emphasized impact factor, bibliometrics, research evaluation, and bias, highlighting early critical engagement with citation-based indicators.

Table 3. Performance analysis results of the second topic: Prestige-driven metrics and research evaluation before and after the emergence of San-Francisco DORA.

	2000-2013 pre-DORA and Leiden Manifesto	2014-2025 post-DORA and Leiden Manifesto
N of analyzed documents	1664	5380
Annual growth rate %	18.35	-1.26
Average citations per year per document	3.572	3.593
Document Types (N)	article (1004) book (1) book chapter (30) conference paper (300) editorial (52) review (277)	article (3554) book (23) book chapter (167) conference paper (548) editorial (113) review (975)
Maximally and minimally productive year (N of publications)	2013 (286) 2000 (32)	2022 (600) 2025 (281)
Top corresponding author's country (N of documents)	USA (296)	China (825)
Top 5 sources	1. PLOS one 2. Scientometric 3. Cochrane database of systematic review 4. Journal of clinical epidemiology 5. Lecture note in computer science	1. Cochrane database of systematic review 2. PLOS one 3. Scientometrics 4. Journal of clinical epidemiology 5. BMJ open
Top 5 keywords	1. Impact factor 2. Bibliometrics 3. Research evaluation 4. Bias 5. Citation analysis	1. Research evaluation 2. Impact factor 3. Bibliometrics 4. P-hacking 5. Publication bias

Note. Unable to determine the number of duplicated documents as the split is done after deduplication. For the top 5 keyword components, keywords that share synonymous meaning as explained above were merged. Keywords that are not relevant to the subject matter were also excluded as mentioned above.

In the post-DORA period, the dataset expanded considerably to 5,380 documents, yet the annual growth rate slowed to -1.26%, indicating a plateau or saturation effect in publication volume. The most productive year is 2022 (600 publications). Despite the slower growth, citation activity slightly increased, with an average of 3.593 citations per document per year. Publication types diversified slightly, with a greater share of book chapters (n = 167) and editorials (n = 113), pointing to broader conceptual discussions post-DORA. Notably, China overtook the USA as the top country by volume of corresponding authors (n = 825), suggesting geographic shifts in leadership. Publication sources remain relatively the same.

Thematically, the most frequent keywords evolved: while research evaluation, impact factor and bibliometrics remained prominent, newer terms like P-hacking and publication bias emerged as prominent keywords; this suggests a deepening concern with research integrity and reproducibility, reflecting a shift in scholarly focus from merely assessing impact to critically examining the processes and incentives that shape scientific outputs. Such developments may reflect the influence of DORA's advocacy for more responsible approaches to evaluating scientific contributions, marking a transition toward a more mature and reflexive discourse on research assessment.

 doi:10.20944/preprints202507.2644.v1

5.1.5. Time Sliced Analysis: Barrier and Equity Issues in Research Accessibility Before and After Sci-Hub

Table 4 displays and juxtaposes performance analysis results of the third topic: Barrier and equity issues in research accessibility before and after the emergence of Sci-hub shadow library. The pre-Sci-Hub period yielded 276 publications with a relatively higher annual growth rate of 10.19%, but this growth was on a relatively small base. The average citations per document per year was relatively low at 0.8838, suggesting that access equity was a niche or emerging issue within scholarly publishing. The most productive year was 2011 (32 documents). Publication types were dominated by journal articles (n = 160), with smaller shares of reviews (n = 43), conference papers (n = 38), and book chapters (n = 16). The USA was the top contributing country (n = 81). Core journals included Serials Librarian, Information Services and Use, and First Monday, pointing to a strong library and information science orientation. Top keywords included publishing, open access, electronic publishing, libraries, and internet, which reflect an infrastructural framing of the access issue, centered on digital transition and institutional roles.

Table 4. Performance analysis results of the third topic: Barrier and equity issues in research accessibility before and after the emergence of Sci-hub shadow library.

	2000-2011 Pre Sci-hub	2012-2025 Post Sci-hub
N of analyzed documents	276	1361
Annual growth rate %	10.19	5.8
Average citations per year per document	0.8838	1.725
Document types (N)	article (160) book (3) book chapter (16) conference paper (38) editorial (16) review (43)	article (978) book (14) book chapter (86) conference paper (91) editorial (33) review (159)
Maximally and minimally productive year (N of publications)	2011 (32) 2003 (10)	2024 (171) 2012 (37)
Top corresponding author's country (N of documents)	USA (81)	USA (258)
Top 5 sources	1. Serials librarian 2. Information services and use 3. D-Lib Magazine 4. College and research libraries 5. First monday	1. Learned publishing 2. Insights: the UKSG journal 3. Scientometrics 4. Publications 5. Publishing research quarterly
Top 5 keywords	1. Publishing 2. Open access 3. Electronic publishing 4. Libraries 5. Internet	1. Open access 2. Article processing charge 3. Publishing 4. Sci-hub 5. Predatory journals

Note. Unable to determine the number of duplicated documents as the split is done after deduplication. For the top 5 keyword components, keywords that share synonymous meaning as explained above were merged. Keywords that are not relevant to the subject matter were also excluded as mentioned above.

Following the rise of Sci-Hub, the dataset grew to 1,361 documents, although the growth rate slowed to 5.8% annually. Importantly, the average citation rate nearly doubled to 1.725, indicating increased academic attention and possibly controversy. The most productive year was 2024 (171 publications), showing peak interest more than a decade after Sci-Hub's appearance, suggesting a delayed but accelerating concern over equity and access. The USA remained the leading country for corresponding authorship (n = 258). There was a shift in publication venues to include Learned Publishing, Insights, Scientometrics, and Publishing Research Quarterly, suggesting a broader disciplinary base regarding science evaluation and communication. In the post-Scihub period, the top keywords of article processing charges, predatory journals, and Sci-Hub itself emerge in addition to the existing top keywords of publishing and open access; This could indicate a more politicized and economically framed discourse, focused on both formal and informal access structures.

 doi:10.20944/preprints202507.2644.v1

5.2. Conceptual Analysis

For conceptual analysis results, similarly to the pattern of reporting of performance analysis results, results of the full dataset will be reported first to provide an overview to conceptual landscape of the three topics, then the juxtaposition of the concept before and after the key event of topic 2 and 3 will be discussed later to examine how the conceptual landscape has evolved overtime.

5.2.1. Keyword Co-Occurrence Network

Gatekeeping and Editorial Bias

Figure 1 displays the keyword co-occurrence networks of this topic. Cluster 1 (Blue) is the largest, comprising 10 key terms. Peer review is the most prominent node, followed by publication bias. Other related terms in this cluster include impact factor, transparency, funding, open science, reproducibility, review, decision making, and editorial decisions. As the largest and most central cluster, this group emphasizes the procedural and normative dimensions of academic publishing. Peer review emerges as the dominant node, reflecting its foundational role in gatekeeping mechanisms. The co-occurrence of terms like publication bias, impact factor, transparency, and open science suggests ongoing concerns about how evaluation metrics and systemic opacity influence editorial decisions.

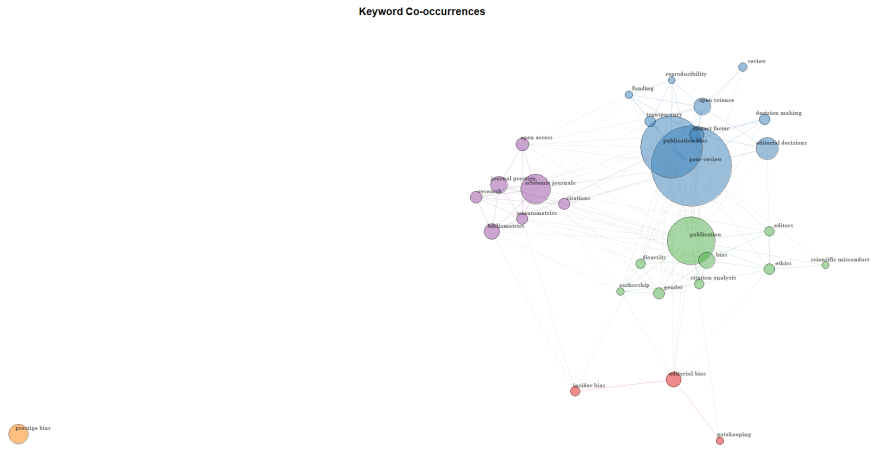

Figure 1. Keyword Co-occurrence Networks of the Gatekeeping and Editorial Bias Topic.

Cluster 2 (Green) includes 9 terms, with publications as the biggest node. Surrounding it are diversity, authorship, gender, bias, citation analysis, ethics, editors, and scientific misconduct. This cluster focuses on the socio-ethical implications of editorial practices, with publications at its core. Terms such as diversity, gender, authorship, and bias highlight concerns about representational equity and implicit bias in scholarly gatekeeping. The presence of ethics, scientific misconduct, and editors suggests that questions of fairness, inclusion, and editorial responsibility are central to this conceptual grouping. Citation analysis appears as a methodological anchor, possibly reflecting attempts to empirically measure disparities in visibility and recognition across demographic or geographic lines.

Cluster 3 (Purple) consists of 7 terms, with academic journals as the most prominent node. Other terms include open access, journal prestige, research, bibliometrics, scientometrics, and citations. The central term, academic journals, could indicate that this cluster is concerned with structural aspects of scholarly communication. Keywords like open access and journal prestige could represent ongoing tensions between traditional prestige-based publication models and newer, more inclusive or accessible alternatives. The inclusion of scientometrics and citations links this discussion to the evaluation of research performance and the stratification of journals, further connecting it to the mechanisms of editorial bias and prestige-driven gatekeeping.

 doi:10.20944/preprints202507.2644.v1

Cluster 4 (Red), the smallest, contains 3 related terms: editorial bias, insider bias, and gatekeeping. Although the smallest, this cluster is conceptually significant as it directly captures the core theme of gatekeeping. The grouping of editorial bias, insider bias, and gatekeeping suggests a focused discourse around the consolidation of editorial power and the potential for exclusionary practices. Additionally, the term prestige bias appears as a peripheral, standalone concept, forming its own isolated cluster. Its isolation may suggest it is either underexplored in empirical research or framed independently from the other dominant discourses, despite its relevance to multiple clusters

Overall, this landscape reflects a growing academic concern with the accountability, transparency, and equity of editorial processes. At its core, the landscape is shaped by the interaction between institutional mechanisms (e.g., peer review, editorial decisions), ethical and social considerations (e.g., diversity, bias, misconduct), and structural features of the scholarly communication system (e.g., journal prestige, bibliometrics), and systemic issue (e.g., insider bias, editorial bias). While each cluster emphasizes a distinct thematic area, together they form a comprehensive picture of how gatekeeping operates, both overtly through editorial decisions and subtly through systemic biases embedded in publishing norms, metrics, or favoritism.

Prestige-Driven Metrics and Research Evaluation

Figure 2 displays the keyword co-occurrence networks of this topic. Cluster 1 (Green) is the largest, consisting of 14 terms. The most prominent node is journal impact factor, followed by research evaluation. Other related terms include altmetrics, bibliometrics, scientometrics, h-index, social media, impact, journal ranking, citations, citation analysis, journals, research impact, and bibliometric analysis. As the largest and most conceptually dense cluster, this group focuses on the tools and indicators used to evaluate research. Journal impact factor is the most prominent term, reinforcing its role as a dominant evaluative standard in academia. The presence of h-index, citations, journal ranking, and bibliometric analysis reflects a system heavily reliant on quantitative indicators. However, the inclusion of altmetrics and social media signals growing interest in alternative forms of research visibility and impact, especially those that capture influence beyond academic citations. This cluster embodies the core of prestige-driven evaluation, where visibility, ranking, and impact are frequently treated as proxies for quality.

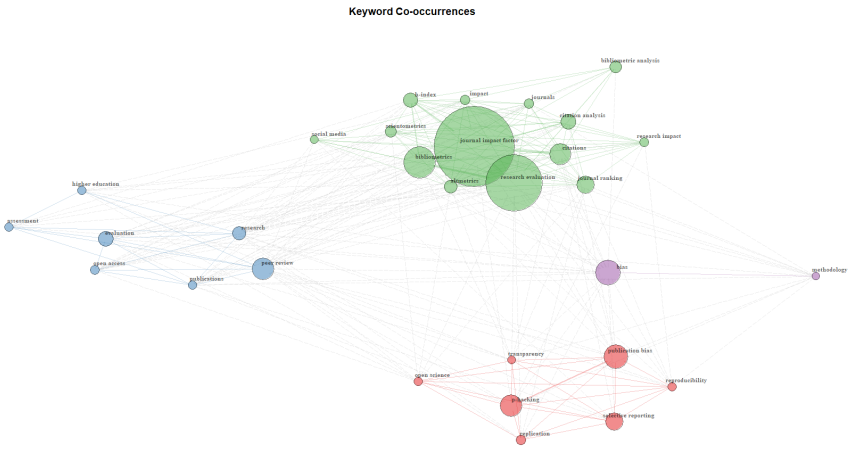

Figure 2. Keyword Co-occurrence Networks of the Prestige-driven Metrics and Research Evaluation Topic.

Cluster 2 (Blue) and Cluster 3 (Red) are equal in size, each containing 7 terms. The Blue cluster comprises terms such as assessment, higher education, evaluation, open access, publications, research, and peer review. This cluster focuses on the institutional and procedural dimensions of research evaluation. Terms like assessment, higher education, evaluation, and peer review suggest a concern with how metrics are used within academic settings and often influencing hiring, funding, and

 doi:10.20944/preprints202507.2644.v1

promotion decisions. The co-occurrence of publications, open access, and research points to ongoing debates about how to balance accessibility with perceived quality or impact. This cluster connects evaluative tools to the broader educational and institutional ecosystem in which they are embedded.

The Red cluster comprises terms such as open science, transparency, p-hacking, replication, selective reporting, publication bias, and reproducibility. This cluster highlights the ethical and epistemological risks of prestige-driven evaluation. Terms like p-hacking, selective reporting, and publication bias reflect how metric pressure can distort research practices, incentivizing statistically significant (but potentially unreliable) findings. The presence of open science, transparency, replication, and reproducibility suggests a countermovement aiming to restore scientific integrity and resist distortions driven by impact-focused publication strategies. This cluster serves as a critique of the unintended consequences of metric obsession and supports ongoing reforms in scientific practice.

Cluster 4 (Purple) is the smallest, comprising just 2 terms: bias and methodology. Though small, this cluster is also conceptually important. These two terms point to growing awareness of bias in research methodology. Overall, the largest thematic space is represented by terms related to quantitative indicators of research (e.g., impact factor), its involvement with the institutional and evaluative contexts of higher education and publication systems (e.g., peer review, higher education), questionable research practice that may stem from research evaluation (e.g., p-hacking, selective reporting), and the broader topic of research bias in research methodology (e.g., bias).

Barrier and Equity Issues in Research Accessibility

Figure 3 displays the keyword co-occurrence networks of this topic. Cluster 1 (Green) is the largest, consisting of 16 terms. Open access is the most prominent node, followed by scholarly publishing and article processing charge. Additional related terms include scientific journal, predatory journal, bibliometrics, Scopus, peer review, Directory of Open Access Journals (DOAJ), hybrid journals, gold open access, journal impact factor, open science, predatory publishing, commercial publishing, and institutional repositories. This dominant cluster focuses on the mechanics, platforms, and economic structures of scholarly publishing. Open access is the central term, with links to scholarly publishing and article processing charge. This co-occurrence reflects the costs of participation in open publishing models. The presence of terms like predatory journal, gold open access, and hybrid journals reflects both the diversification and exploitation of the open access movement. This cluster also includes operational elements such as Scopus, DOAJ, institutional repositories, and bibliometrics, showing how the formalization and indexing of open-access content is shaping scholarly visibility. Interestingly, the coexistence of commercial publishing and journal impact factor keywords in this cluster suggests that prestige-driven metrics and profit-based models continue to influence the spaces nominally committed to scientific dissemination.

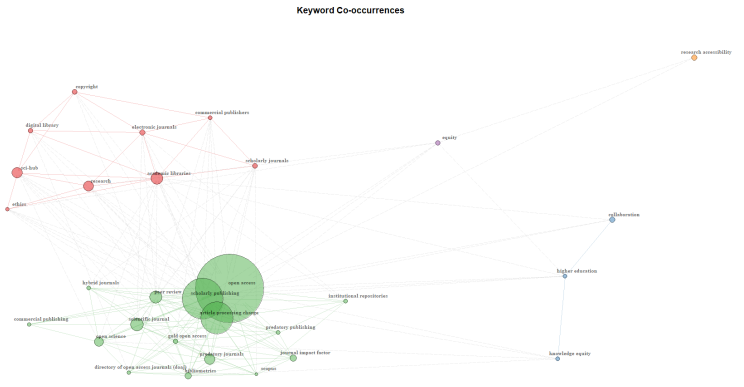

Figure 3. Keyword Co-occurrence Networks of the Barrier and Equity Issues in Research Accessibility Topic.

 doi:10.20944/preprints202507.2644.v1

Cluster 2 (Red) comprises 9 terms such as ethics, research, Sci-Hub, academic libraries, electronic journals, copyright, digital library, commercial publishers, and scholarly journal. This cluster centers on the moral and legal dimensions of research accessibility. Terms such as Sci-Hub, copyright, and commercial publishers frame a discourse around access as both an ethical and legal issue. The inclusion of academic libraries, digital libraries, and electronic journals shows that both traditional and digital infrastructures continue to play a crucial role in research access provision, by serving as the foundation of subscription-based models and open access ideals. This cluster represents an interplay between authorized access models and unauthorized or legally grey systems that emerge in response to persistent barriers.

Cluster 3 (Blue) is smaller, with 3 terms: higher education, collaboration, and knowledge equity. The co-occurrence of this cluster suggests efforts to address inequities in knowledge access and production. This cluster could represent a shift in the discourse from technical discussions of access to broader concerns about justice, participation, and inclusion. Additionally, two terms of equity and research accessibility are positioned at the periphery of the conceptual map and are considered standalone nodes forming their own micro-clusters. This spatial positioning implies that, while frequently appearing, these terms may not yet be fully operationalized or empirically connected within the broader literature.

Overall, the co-occurrences in this network represent the tensions between means of access, commercial interests, and scholarly inclusivity. The central focus of the network revolves around the dynamics of open access publishing, from its promises of democratizing knowledge to the economic and institutional barriers that persist (Green, 2019). Across clusters, several key themes emerge: the economic cost of access (e.g., article processing charges), ethical challenges (e.g., predatory publishing, copyright), institutional roles (e.g., libraries, repositories), and the global inequities tied to who can access and participate in knowledge production.

5.2.2. MCA-Based Conceptual Clustering

Gatekeeping and Editorial Bias

Figure 4 displays the keyword clusters from MCA of this topic. This, and the following figures, highlight the most prominent keywords within each cluster. Due to visual limitations and overlap, not all smaller keywords are displayed, only those with the highest prominence appear in the output. The biggest cluster (blue) has prominent keywords such as editorial decision making, manuscript, transparency, publication, ethics, interventions, academic journals, conflict of interest, bias, editor, peer-review, academia, media, authorship, citation, metrics, journal quality, and scopus. This cluster captures the process behind editorial decision-making. Central terms such as editorial decision making, manuscript, peer-review, and publication emphasize the procedural core of the gatekeeping process. These are complemented by ethical concerns like transparency, conflict of interest, ethics, and bias editor. The inclusion of authorship, citation, metrics, journal quality, and scopus indicates a broader evaluative framework in which editorial decisions are shaped not just by content but by prestige indicators and visibility metrics. For example, journals with a high impact factor may be more selective in accepting manuscripts than their lower peers. This cluster reflects a literature focused on reform, standardization, and the need for greater accountability in editorial practices.

 doi:10.20944/preprints202507.2644.v1

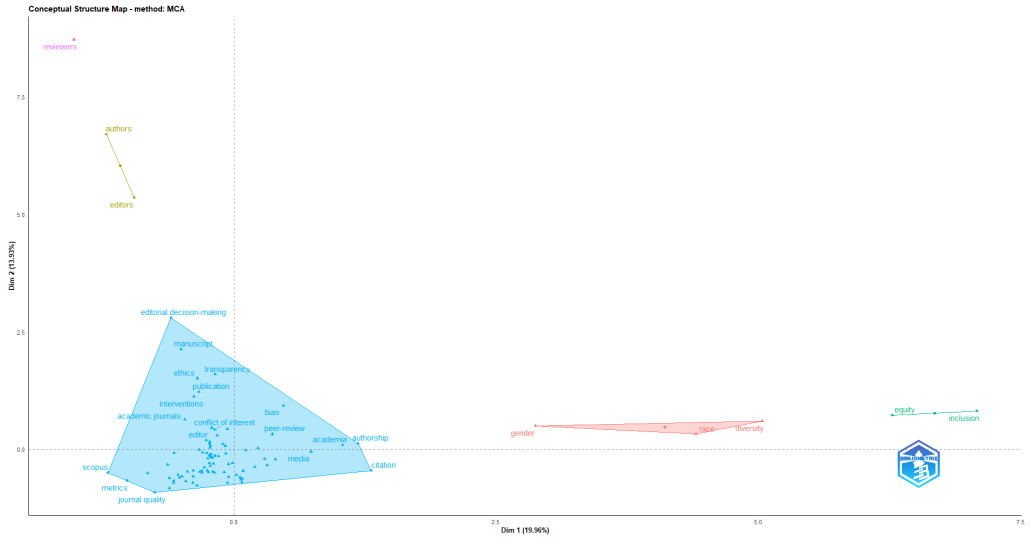

Figure 4. MCA Keyword Cluster of the Gatekeeping and Editorial Bias Topic.

The red cluster has prominent keywords of gender, race, and diversity. This cluster reflects the socio-demographic critiques of editorial practices. Gender, race, and diversity form a conceptual triad that points to systemic biases in academic publishing, particularly in who gets published, cited, or invited to contribute. The presence of this distinct cluster signals a dedicated strand of research exploring how identity categories intersect with publishing outcomes, highlighting equity gaps in participation, recognition, and authority.

The green cluster has keywords of equity and inclusion. While related to the red cluster, this group focuses more on institutional responses. Equity and inclusion represent values and policy goals often cited in editorial guidelines and reform discourse. This cluster appears slightly conceptually removed from the procedural focus of the blue cluster, which may suggest that equity and inclusion goals are often articulated separately from operational reforms, perhaps more as aspirational ideals than as directly integrated practices.

The yellow cluster has keywords of authors and editors. This small cluster reflects the relational dynamics at the heart of editorial decision-making. The keywords authors and editors encapsulate the asymmetrical power relationship that shapes publication outcomes. The isolated position of this cluster suggests that while these roles are central to the publishing process, discussions of their interaction are often separated from the broader systemic aspect. There is also one isolated keyword: reviewers. The solitary position of reviewers is notable. As key actors in the gatekeeping system, reviewers wield substantial influence, yet their role appears conceptually distant in the MCA space. This suggests a literature gap or conceptual underdevelopment regarding how reviewers contribute to or mitigate editorial bias.

Overall, the MCA output on gatekeeping and editorial bias reveals a conceptually stratified structure composed of interrelated but distinct thematic clusters. The largest cluster (blue) reflects a comprehensive operational view of editorial decision-making, where various publishing actors and institutional processes interact. The smaller clusters (red, green, and yellow) delineate more focused socio-structural issues, such as identity-related bias (gender, race), equity and inclusion goals, and the role-based dynamics between authors and editors.

Prestige-Driven Metrics and Research Evaluation

Figure 5 displays the keyword clusters from MCA of this topic. The blue cluster, which is the largest, has keywords such as article, impact, social media, and replicability. This cluster reflects the current shift toward article-level metrics and altmetrics, where visibility and influence are increasingly measured by social dissemination (e.g., Twitter mentions, downloads) in addition to traditional citations. The inclusion of replicability could mean a growing awareness that prestige and visibility

 doi:10.20944/preprints202507.2644.v1

may not necessarily equate to methodological quality. This cluster, therefore, represents the tension between quantitative performance indicators and the replication crisis, suggesting that prestige and rigor may contradict each other in some cases.

The yellow cluster has keywords such as effectiveness, critical appraisal, sample size, randomization, and p-value. These terms suggest a strong emphasis on methodological soundness in empirical studies, especially in fields that rely on experimental or quasi-experimental designs (e.g., biomedical sciences, psychology). This cluster focuses on evaluative validity from a scientific standpoint, questioning whether the studies that gain prestige (or drive metrics) are actually methodologically sound. It suggests a literature base that pushes for more rigorous statistical standards and better study design to ensure that research findings are both valid and generalizable.

The red cluster has keywords such as efficacy, methods, scientific misconduct, reporting guidelines, validity, and effect size. This cluster critiques the integrity of research practices under prestige pressure. The keywords reflect concerns about how prestige-based incentives may distort research behavior, leading to exaggerated claims, poor reporting, and even unethical conduct. This cluster aligns conceptually with reformist movements that advocate for transparency, registration, and adherence to rigorous reporting standards. It highlights how the pursuit of high-impact publication may compromise methodological and ethical standards if left unchecked.

The purple cluster has keywords such as novelty, preprints, and PubMed. This cluster indicates a literature focused on new forms of dissemination and the value placed on being first, fast, and visible. This cluster captures an innovation-centered view of prestige, where novelty is rewarded and preprint platforms serve as both accelerators and disruptors of traditional peer review. PubMed, as a biomedical indexing platform, reflects the centrality of database visibility in research evaluation, where inclusion alone can influence perceived legitimacy and reach [93]. The green cluster has keywords such as blinding and acceptance. This smaller cluster focuses on editorial and review processes, specifically on how manuscripts are evaluated prior to publication.

Overall. This MCA map shows interconnected discourse between article-level metrics with issues around research replicability and social visibility, while the surrounding clusters represent methodological integrity (yellow, red), innovation in dissemination (purple), and acceptance processes (green).

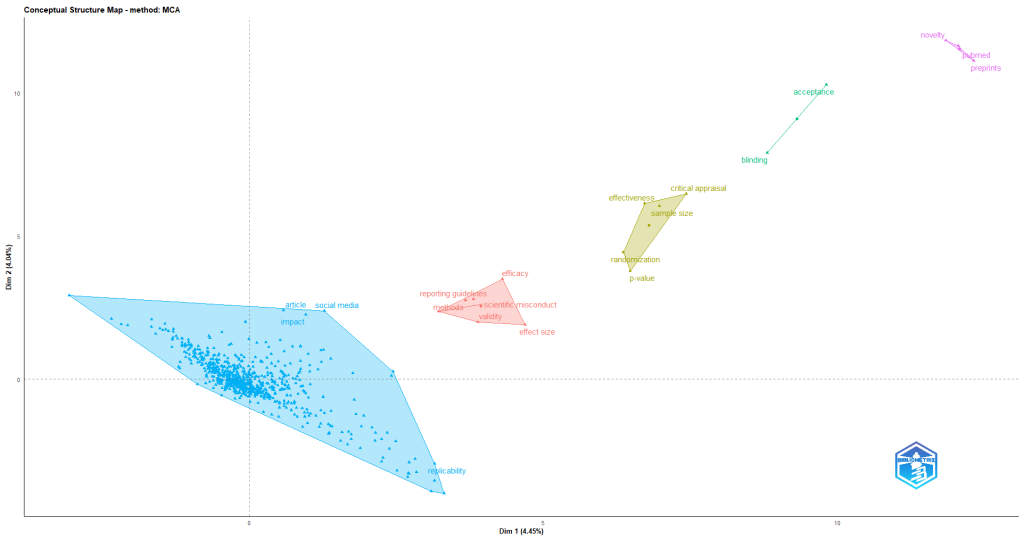

Figure 5. MCA Keyword Cluster of the Prestige-driven Metrics and Research Evaluation Topic.

Barrier and Equity Issues in Research Accessibility

Figure 6 displays the keyword clusters from MCA of this topic. The blue cluster, which is the largest, has keywords such as publishing model, creative commons, open access policy, green, self-archiving, unpaywall, gold OA, metrics, impact, citations, Beall's list, bibliometrics, altmetrics,

 doi:10.20944/preprints202507.2644.v1

publishers, journal, institutional repository, subscriptions, pricing, licensing, commercial publisher, open data, Elsevier, Springer, h-index, research funding, repository, research evaluation, science, databases, grey literature, academic library, shadow libraries, evidence-based practice, information literacy, knowledge, teaching, barriers, and information access. This extensive cluster reflects the broad architecture of access and its limitations. Terms such as publishing model, open access policy, green OA, and gold OA capture the variety of access routes researchers can pursue or encounter. Simultaneously, the presence of Beall's list, commercial publisher, Elsevier, Springer, and subscriptions reveal the tension between commercial control and free access and dissemination ideals, particularly between the cost of legitimate journals and the rise of their predatory counterparts. The inclusion of metrics, impact, citations, bibliometrics, and altmetrics indicates that access is increasingly evaluated in terms of measurable performance, linking it to research evaluation systems that may reinforce inequities. Academic library, institutional repository, shadow libraries, and grey literature reflect alternative or grassroots strategies to circumvent barriers, often invoked in response to the high costs of journal access or APCs. Terms like evidence-based practice, information literacy, and teaching show the downstream educational and societal consequences of restricted access, revealing that barriers extend beyond academia and into public knowledge ecosystems.

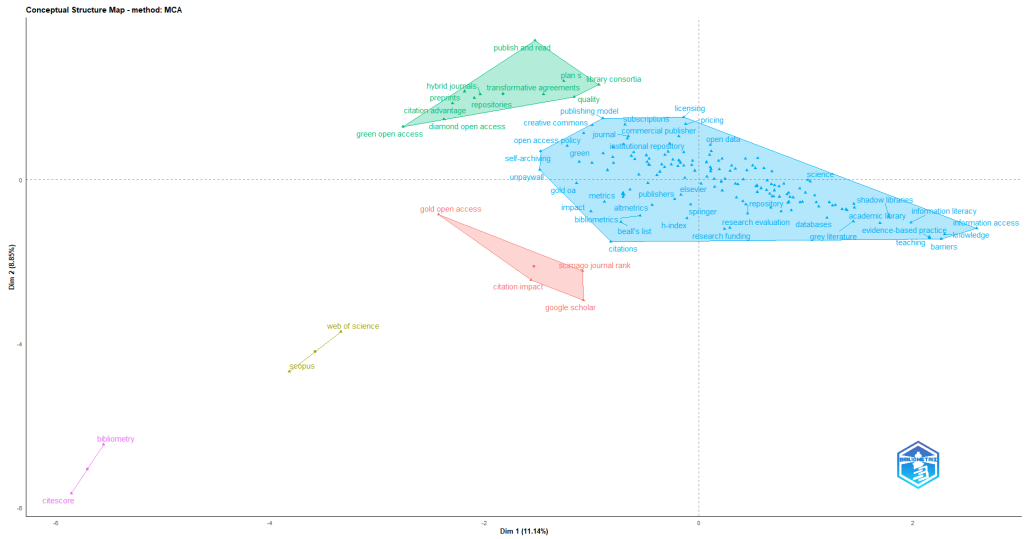

Figure 6. MCA Keyword Cluster of the Barrier and Equity Issues in Research Accessibility Topic.

The green cluster, which is the second-largest, has keywords such as green open access, diamond open access, citation advantage, preprints, repositories, hybrid journals, transformative agreements, quality, Plan S, publish and read, and library consortia. The terms in this cluster point to strategic attempts by funders, consortia, and institutions to restructure publishing economics (e.g., green open access, diamond open access, hybrid journals). Importantly, the term citation advantage and quality indicate that these newer models are not only framed as ethically superior, but also academically beneficial, thereby challenging assumptions that open models compromise scholarly prestige. This cluster reflects a shift toward institutional alignment, where universities, libraries, and funders collaborate to create sustainable, equitable alternatives to legacy publishing.

The red cluster has keywords such as gold open access, citation impact, Google Scholar, and Scimago Journal Rank. This cluster captures a hybrid discourse that connects accessibility with prestige, where open access is pursued not purely for equity but also for increased visibility and citation performance. For instance, their interaction could imply how even equity-driven models like gold open access could increase visibility and citation, thus connecting prestige metrics with open science practice.

Two smaller clusters also appear: The yellow cluster has keywords of Web of Science and Scopus. This cluster reflects the role of indexing platforms. Inclusion in these databases is often a prerequisite

 doi:10.20944/preprints202507.2644.v1

for legitimacy, yet they are dominated by commercial or Western-centric frameworks. This cluster reveals how access and recognition are mediated by selective visibility, reinforcing global disparities in publishing. The purple cluster has keywords of bibliometry and CiteScore. This cluster highlights the technical instruments used to rank and evaluate journals. This cluster encapsulates the metricization of access, where open publishing is increasingly judged not on openness alone, but on its metric footprint.

Overall, this MCA landscape suggests that barriers to access are not merely technical or economic, but deeply systemic by relating to entrenched prestige hierarchies, commercial interests, and uneven global infrastructure. The result presents an interaction between the ecosystem of research access issues, its remedy, and its relationship with performance metrics governed by commercial publishers. In other words, the literature landscape represents discourse on equity-versus-exclusivity tension that defines contemporary academic publishing.

5.2.3. Thematic Map

Gatekeeping and Editorial Bias

Figure 7 displays the thematic map of this topic. The motor theme quadrant includes three prominent keyword clusters. Cluster 1 has keywords such as peer-review, publication bias, publication, academic journals, and bias. This cluster represents its foundational critique of traditional publishing practices, focusing on how bias in peer review and publication decisions distorts the research practice. Cluster 2 includes open science, reproducibility, replication, editorial, and methodology. This cluster represents the reformation effort, advocating for greater methodological transparency and editorial accountability through open science practices. Cluster 3 has keywords such as diversity, gender, decision making, equity, and gender bias, suggesting a strong thematic link between editorial practices and broader issues of fairness and representation in academic publishing. It represents a justice-oriented dimension, reframing editorial practices as not just epistemic filters but also sites of structural inequity, with the representation and fairness aspect as the focus.

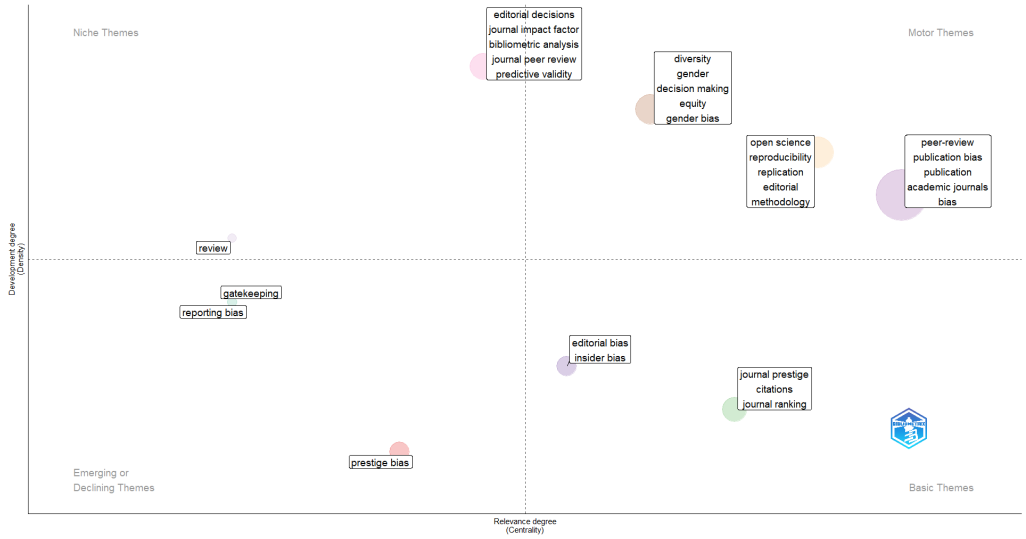

Figure 7. Thematic Map of the Gatekeeping and editorial bias Topic.

The niche theme quadrant has one keyword cluster and one isolated keyword. The cluster has editorial decisions, journal impact factor, bibliometric analysis, journal peer review, and predictive validity. This cluster explores the quantification of editorial practices, analyzing the reliability, outcomes, and influence of peer review using empirical tools. The isolated keyword is "review". The isolated keyword may reflect a broader element of the editorial practices.

The emerging or declining theme quadrant has three isolated keywords of gatekeeping, reporting bias, and prestige bias. These terms are highly relevant to systemic critiques but remain thematically

 doi:10.20944/preprints202507.2644.v1

fragmented in the current literature. Their placement suggests that although scholars recognize their importance, integrated conceptual frameworks or empirical applications are still lacking.

The basic theme quadrant has two clusters: Cluster 1 has editorial bias and insider bias, which represents foundational concerns in the subject of power exercise in editorial gatekeeping. Cluster 2 has journal prestige, citations, and journal ranking, which reflect the prestige economy of academic publishing, with journal ranking and manuscript citation treated as academic currency.

Overall, the motor theme quadrant highlights key discourses around systemic reform, transparency, and justice in academic publishing, with clusters emphasizing critiques of traditional peer review bias, advocacy for open science and editorial accountability, and calls for fairness and equity in publication practices. The niche theme centers on the empirical examination of editorial decision-making, focusing on how bibliometric tools and impact metrics quantify review reliability and predictive value. The emerging or declining theme captures fragmented but relevant topics such as gatekeeping, reporting bias, and prestige bias. Finally, the basic theme captures foundational concerns such as the exercise of editorial power through bias and insider dynamics, and the prestige economy of scholarly publishing where journal rank and citation metrics drive academic value.

Prestige-Driven Metrics and Research Evaluation

Figure 8 displays the thematic map of this topic. There is one keyword cluster between the motor theme and basic theme with keywords of journal impact factor, research evaluation, bibliometrics, citations, and citation analysis. This hybrid position could indicate a conceptual foundation that is both central and relatively mature. It reflects the metric-driven logic of academic prestige. While widely used as tools for performance assessment, these indicators have come under scrutiny for oversimplifying research quality. Their placement suggests they are deeply rooted in practice (basic) but are also thematically influential and increasingly critiqued (motor).

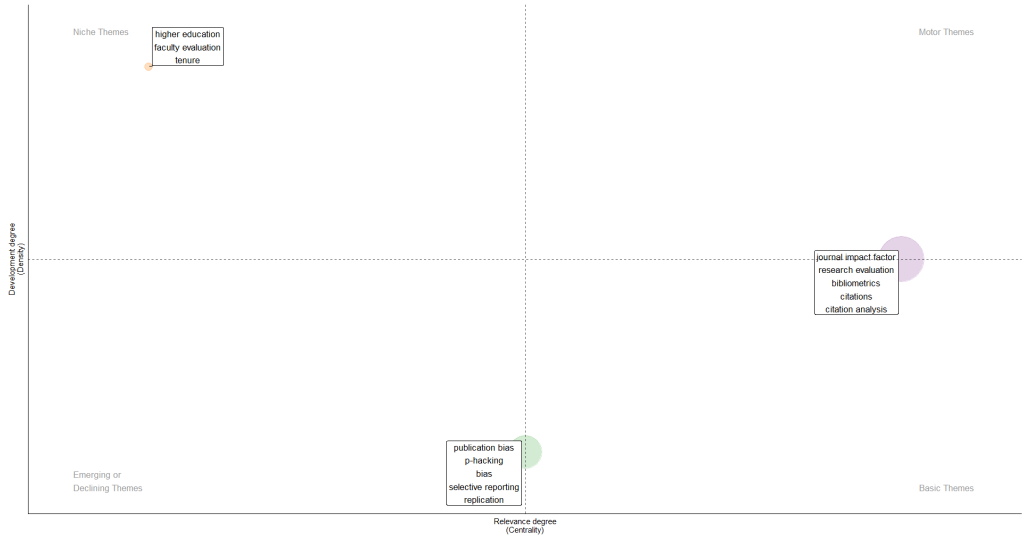

Figure 8. Thematic Map of the Prestige-driven Metrics and Research Evaluation Topic.

There is one keyword cluster between the emerging/declining theme and basic theme with keywords of publication bias, p-hacking, bias, selective reporting, and replication. This suggests ongoing but possibly fluctuating attention to the topic. This cluster indicates topics that are recurrent but not yet fully integrated into the central interest. These keywords represent how evaluation pressures lead to biased practices and compromised research integrity. Their position suggests ongoing, cyclical interest that rises with the replication crisis, but may lack sustained theoretical development.

The niche theme quadrant has a keyword cluster with higher education, faculty evaluation, and tenure. The niche theme quadrant reflects a specialized but well-developed area of research that connects evaluation metrics with institutional structures, particularly in higher education. The topic suggests prestige metrics shape academic labor, career advancement, and organizational decisions.

 doi:10.20944/preprints202507.2644.v1

Overall, the thematic structure reveals interconnected conversations around prestige and evaluation in academia. The motor-basic hybrid theme suggests the dominance of performance indicators in shaping scholarly worth, highlighting both their entrenched role in research assessment and the growing scrutiny over their reductionist influence. The emerging-basic hybrid theme captures concerns about how institutional pressures may incentivize questionable research practices. Meanwhile, the niche theme focuses on how institutional policies, particularly in academic employment and advancement, are increasingly governed by metric-driven evaluations, illustrating the connection between numerical prestige and organizational decision-making.

Barrier and Equity Issues in Research Accessibility

Figure 9 displays the thematic map of this topic. In the motor theme quadrant, two keyword clusters and one isolated keyword are identified. Cluster 1 includes keywords such as article processing charge, gold open access, Directory of Open Access Journals (DOAJ), transformative agreement, and Plan S. This cluster highlights the economic and regulatory frameworks of open access. APCs, transformative agreements, and mandates like Plan S that are central to ongoing debates about who pays for access, and who benefits from it. Cluster 2 contains ethics, higher education, open data, data sharing, and metadata. The isolated keyword in this quadrant is cost, showing financial concerns as a key but less connected element. This cluster connects data practices and institutional ethics, emphasizing responsible data sharing and openness as key enablers of equitable research dissemination, particularly within higher education contexts. The isolated keyword represents a general concern in knowledge and research access and dissemination.

Figure 9. Thematic Map of the Barrier and Equity Issues in Research Accessibility Topic.

The niche theme quadrant contains one small keyword cluster with diversity and disparities. This small cluster signals growing interest in how access inequalities reflect broader social disparities, such as race, gender, or geography. Additionally, four isolated keywords of hybrid journal, collection development, funding, and barriers appear as standalone thematic points, each forming its own cluster and suggesting areas of specialized or underdeveloped inquiry. The isolated keywords may represent research access issues that are not yet integrated with other bigger topics.

The emerging or declining theme quadrant has one keyword cluster with keywords such as databases and grey literature, indicating shifting attention toward alternative research sources and dissemination methods. The cluster points to alternative and less visible forms of knowledge dissemination, often used in the Global South or in non-academic communities to bypass the traditional notions of "legitimate" publication. Two additional isolated keywords of knowledge equity and collaboration form separate thematic nodes. This suggests that while knowledge equity and collaboration

 doi:10.20944/preprints202507.2644.v1

are being paid attention to in the grand scheme of research accessibility, they still lack a consolidated research base or theoretical anchoring to the bigger topics.

The basic theme quadrant is the most populated, featuring four clusters and one isolated keyword. Cluster 1 includes open access, scholarly publishing, academic libraries, predatory journals, and scientific journals. It represents the infrastructure of access discourse, referencing institutions, publication types, and challenges related to legitimacy (e.g., predatory journals). Cluster 2 has keywords such as Sci-Hub, copyright, piracy, black open access, and shadow libraries. It reflects pirate or informal access routes, showing how users circumvent paywalls, often ethically and legally contested, yet seen as necessary in under-resourced regions. Cluster 3 includes equity and research accessibility. This cluster links access with fairness; these terms articulate the moral imperative behind open access, though their basic positioning suggests the lack of developmental traction. Cluster 4 contains digital library and education. They represent library-based and educational avenues for access, often seen as enablers of digital inclusion. The isolated keyword in this quadrant is research, indicating its broad relevance yet disconnected thematic linkage.

Overall, the thematic map on barrier and equity issues in research accessibility reveals a landscape of structural, ethical, and practical concerns. The motor theme quadrant emphasizes the economic and institutional dimensions of open access, highlighting debates over funding models and the ethics of equitable data sharing. The niche theme focuses on specialized discussions on social disparities in access, as well as underexplored infrastructural and financial challenges that remain thematically isolated. Within the emerging or declining theme shows discussions on alternative dissemination channels and grassroots knowledge practices, though these remain conceptually scattered and undertheorized. Meanwhile, the basic theme presents the conversation in core infrastructures, both formal and informal, of research accessibility, encompassing institutional roles and piracy issues that underlie access discourse.

5.2.4. Time Sliced Analysis: Prestige-Driven Metrics and Research Evaluation Before and After DORA

Keyword Co-Occurrence Network Before DORA

Figure 10 displays the keyword co-occurrence networks of this topic before DORA. The largest cluster (green), has 22 keywords and is centered around the keyword journal impact factor, which appears as the most prominent node. This cluster also includes related terms such as bibliometrics, citations, research evaluation, journal ranking, citation analysis, journals, bias, publications, ranking, publishing, peer review, bibliometric indicators, h-index, scientometrics, methodology, Google scholar, self-citation, journal quality, research evaluation, journal impact factor, and impact. Together, these keywords indicate a strong thematic focus on metrics-based assessment and the widespread influence of impact factor as a dominant evaluative tool. This cluster forms the intellectual backbone of the pre-DORA discourse, reinforcing how impact factor-based evaluation was normalized as the dominant standard. It reflects both the tools and culture of a performance-oriented academic system: one where research was valued primarily for its visibility and citation count, and where impact metrics defined prestige and quality. Terms like self-citation and ranking hint at growing awareness of metric manipulation, though critique appears embryonic at this stage.

 doi:10.20944/preprints202507.2644.v1

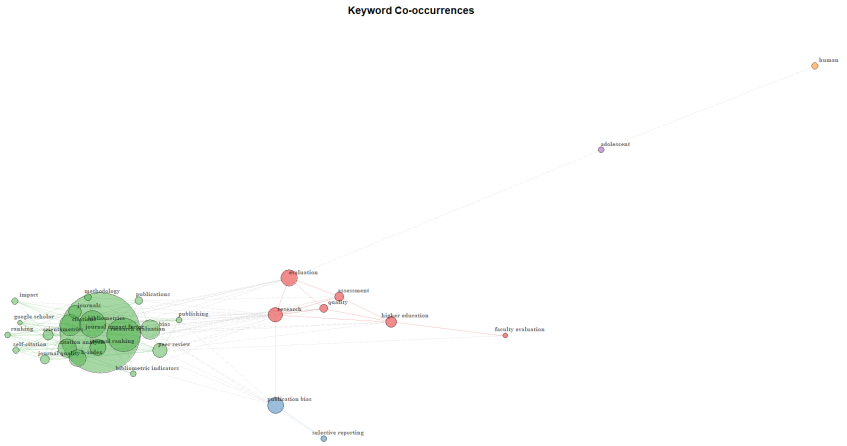

Figure 10. Keyword Co-occurrence Networks of the Prestige-driven Metrics and Research Evaluation Topic: pre-DORA.

The second-largest cluster, shown in red, includes six keywords: evaluation, research, quality, assessment, higher education, and faculty evaluation. This cluster reflects institutional-level concerns around performance evaluation in academic settings. This cluster links the use of bibliometric measures to institutional practices, particularly in faculty reviews, promotions, and higher education governance. It emphasizes the real-world consequences of metric reliance, indicating that scholarly metrics were not just abstract tools, but deeply tied to career structures and academic hierarchies.

A smaller cluster, in blue, contains the keywords publication bias and selective reporting, pointing to early concerns about research transparency and integrity. This smaller cluster captures early concerns about the limitations and distortions of metrics-driven publishing, particularly how pressure to produce "impactful" results may encourage non-transparent or biased research practices. Additionally, there are two isolated keywords, adolescent and human, each forming their own separate cluster. These may represent noise or less thematically integrated terms in the context of the overall topic.

Overall, the pre-DORA landscape is dominated by a strong, centralized focus on traditional bibliometric indicators, especially the Journal Impact Factor. This reflects a scholarly environment where quantitative metrics were deeply embedded in research assessment practices, and where critical discourse around these measures was only just beginning to emerge. The landscape appears hierarchical, metric-driven, and institutionally focused, with limited representation of alternative evaluative models or systemic critiques.

MCA-Based Conceptual Clustering Before DORA

Figure 11 displays the MCA keyword clusters for this topic before DORA. The largest cluster, shown in blue, includes a wide range of keywords that reflect the central focus on research assessment and bibliometric indicators. Prominent terms in this cluster include assessment, quality assessment, reporting, replication, research, science policy, validity, quality control, authorship, impact, peer review, evaluation, methodology, SSCI, journals, digital libraries, citation index, editorial policies, survey, journal impact factors, journal citation reports, journal ranking, ranking, h-index, journal quality, PageRank, bibliometric indicators, citations, citation analysis, eigenfactor, scientometrics, research quality, journal metrics, self-citation, journal evaluation, research funding, and scientific quality. This cluster forms the intellectual and operational core of pre-DORA discourse, reflecting a deep entrenchment of bibliometric logic in how research value was assessed. The emphasis on diverse indicators (e.g., h-index, eigenfactor, PageRank) and ranking tools (journal citation reports, SSCI, journal quality) reflects a quantitative mindset that prioritized measurability and performance. The inclusion of terms like authorship and methodology shows some early interest in the structure

 doi:10.20944/preprints202507.2644.v1

of scholarly production, though these appear secondary to the core logic of metrics. The presence of scientific quality and research funding signals how evaluative metrics influenced both reputation and resource allocation, reinforcing prestige hierarchies in academic publishing.

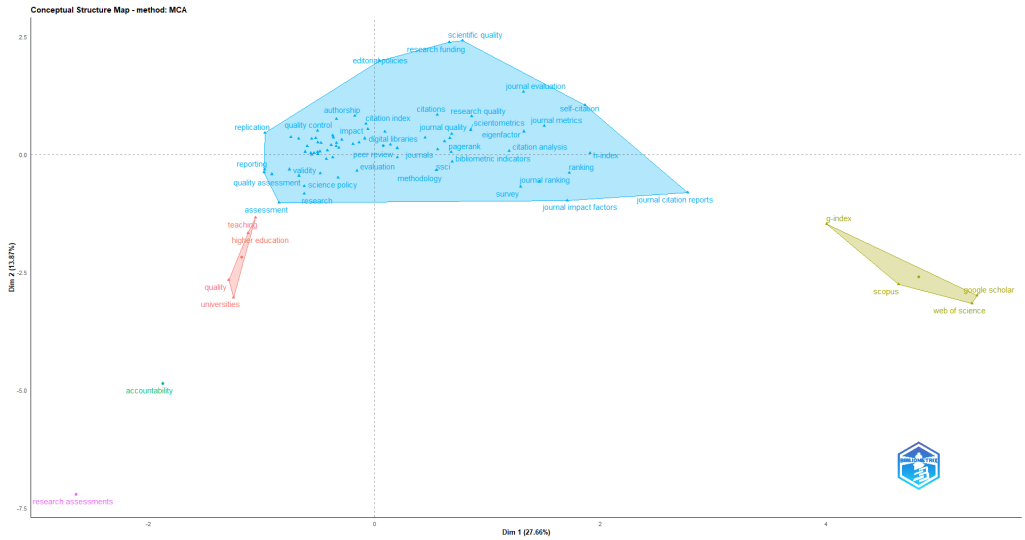

Figure 11. MCA Keyword Cluster of the Prestige-driven Metrics and Research Evaluation Topic: pre-DORA

The second cluster, shown in yellow, includes the keywords g-index, Scopus, Web of Science, and Google Scholar, highlighting discussions around citation databases and indexing tools used in evaluating research output. This cluster represents the technical infrastructure of bibliometric evaluation. The focus is on the data sources and indexing platforms that underpin metric generation. These tools shaped which publications counted, how citations were measured, and ultimately, which forms of knowledge were made visible or valued.

The third cluster, shown in red, features keywords such as teaching, higher education, quality, and universities, suggesting a thematic focus on the intersection of research evaluation with academic institutions and education. This cluster reflects how research evaluation intersected with educational institutions, particularly in the matter of quality assurance, institutional ranking, and accountability measures in universities. The inclusion of teaching suggests that research performance was increasingly tied to broader evaluations of academic productivity, often at the expense of teaching-focused roles. This cluster reveals early traces of the managerial logic of academic audit cultures. Additionally, there is one isolated keyword, accountability, which forms its own cluster. This may indicate an emerging or underexplored concept within the broader discourse on evaluation and metrics. The isolation of accountability indicates that while the idea was present, it had not yet become integrated into the mainstream evaluative discourse. Its marginal status suggests that system-level critiques or demands for responsible metrics were still emerging, laying the groundwork for post-DORA conceptual shifts.

The MCA map of the pre-DORA period reveals a highly concentrated and metrics-driven conceptual structure, with a dominant center revolving around traditional bibliometric indicators and journal-level metrics. There is a lack of critical, reflexive, or alternative evaluation approaches. The analysis reflects a field in which evaluation practices were largely unchallenged and institutionally embedded, emphasizing quantification and ranking over qualitative or systemic concerns.

Thematic Map Before DORA

Figure 12 displays the thematic map of this topic before DORA. In the motor theme quadrant, there is one prominent keyword cluster containing journal impact factor, research evaluation, bibliometrics, citations, and citation analysis. This cluster represents a well-developed and influential conceptual cluster where metrics such as the impact factor and citations are not only central but treated as valid proxies of research quality and influence. The presence in the motor quadrant shows these ideas were widely accepted, frequently used, and deeply embedded in academic decision-making.

 doi:10.20944/preprints202507.2644.v1

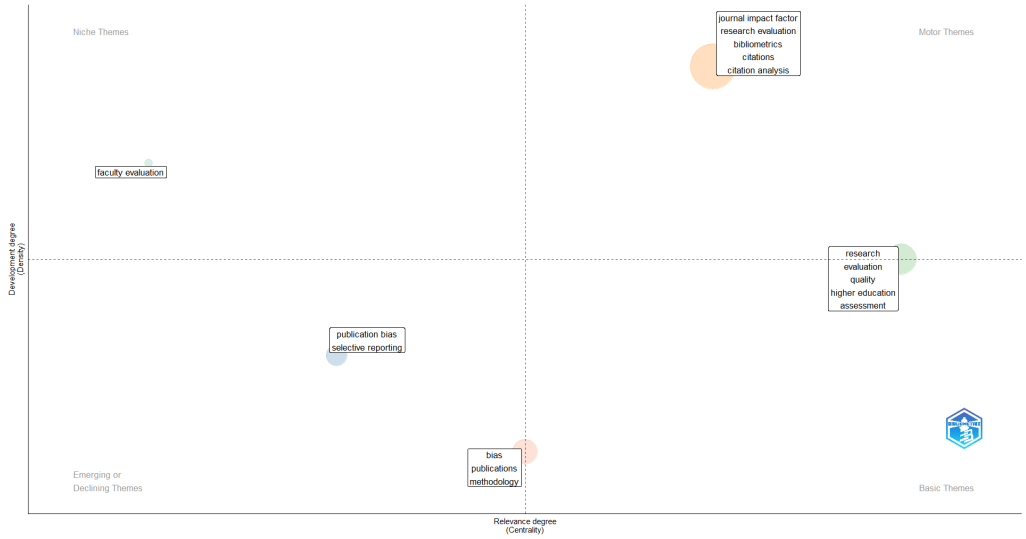

Figure 12. Thematic Map of the Prestige-driven Metrics and Research Evaluation Topic: pre-DORA.

The niche theme quadrant has one isolated keyword of faculty evaluation. This isolated keyword points to early but limited discourse on individual-level assessment, such as tenure and promotion practices. Its niche status implies that while relevant, faculty-level implications of metrics had not yet been broadly explored.

The emerging or declining theme quadrant has one keyword cluster of publication bias and selective reporting, reflecting concerns around research integrity that may be gaining or losing traction. This cluster captures concerns related to questionable research practices that were beginning to enter the scholarly conversation. Its position as emerging or declining suggests these issues were either low in visibility or not yet seen as integral to the discussion on metrics and evaluation. This foreshadows later movements like DORA, which would more forcefully integrate integrity and openness into reform narratives.

Additionally, there is a keyword cluster positioned between the motor theme and basic theme quadrants, which includes research, evaluation, quality, higher education, and assessment, bridging foundational and more engaged themes. This bridge cluster reflects how metric-based evaluations are prevalent in academia, influencing institutional decisions related to teaching, funding, and academic productivity. While still grounded as foundational matter (i.e., basic theme), the cluster's position near the motor quadrant suggests a transition from basic conceptual ideas to system-level implementation, particularly within higher education systems.

Another cluster is located between the emerging or declining theme and basic theme quadrants, containing the keywords bias, publications, and methodology, pointing to ongoing discussions about research practices and potential biases. This intermediary cluster reflects the developing critique of the evaluative status quo, suggesting a growing awareness that research practices and methodological rigor could be distorted or undervalued by metric-driven incentives.

The thematic map represents uncritical acceptance of citation-based metrics. The motor theme quadrant is centered on the normalized use of bibliometric indicators as proxies of research quality, revealing the extent to which these measures were embedded in academic governance. The niche theme highlights the limited early attention to how such metrics impacted faculty evaluation. In the emerging or declining theme quadrant, initial concerns around research integrity and selective reporting were beginning to surface. The motor and basic hybrid theme captures the transition from foundational ideas to institutional implementation of evaluative metrics, while the emerging and basic hybrid theme presents a growing critique of methodological distortion under metric-driven pressures.

 doi:10.20944/preprints202507.2644.v1

Keyword Co-Occurrence Network After DORA

Figure 13 displays the keyword co-occurrence networks of this topic after DORA. The biggest cluster (green) has 15 words, with journal impact factor as the biggest node, research evaluation as the second biggest node, and smaller nodes of research impact, gender bias, altmetrics, journal ranking, citations, bibliometrics, impact, h-index, bibliometric analysis, scopus, scientometrics, citation analysis, and social media. This cluster remains about in traditional bibliometric indicators (e.g., impact factor, h-index), but shows clear signs of expansion and diversification. The inclusion of altmetrics, social media, and gender bias indicates that quantitative evaluation is no longer limited to citation counts, but increasingly considers broader dimensions of research visibility and equity. This change could reflect DORA's influence in encouraging the adoption of alternative indicators that better capture societal and interdisciplinary value.

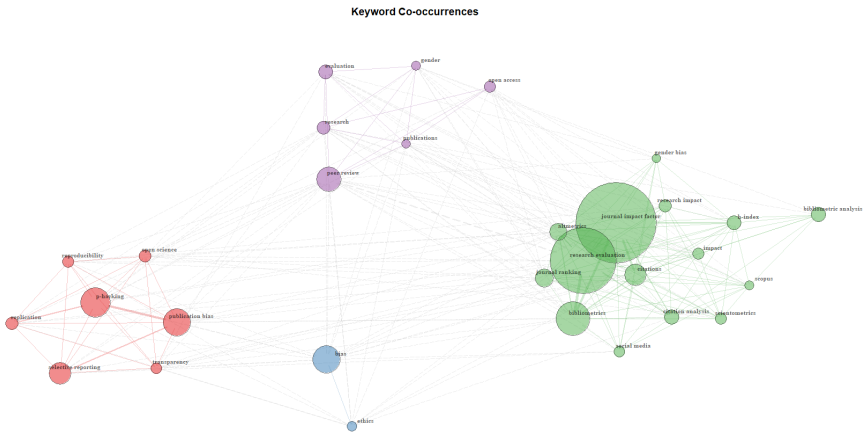

Figure 13. Keyword Co-occurrence Networks of the Prestige-driven Metrics and Research Evaluation Topic: post-DORA.

The red cluster has 7 words of reproducibility, open science, p-hacking, replication, publication bias, selective reporting, and transparency. This cluster represents a coherent and maturing thematic area centered on research transparency and integrity. The clustering of open science with reproducibility and publication bias suggests a post-DORA alignment between evaluative reform and scientific reform movements. These issues have moved from peripheral critiques (pre-DORA) into a central conceptual network, illustrating growing consensus around the need for ethical, transparent, and replicable research practices.

The purple cluster has 6 words of evaluation, research, gender, peer review, publications, and open access. This cluster reflects the integration of equity and openness into evaluative discussions. The presence of gender, peer review, and open access together points to an expanding evaluative framework that increasingly includes questions of fairness, representation, and accessibility. These are likely downstream effects of DORA's advocacy for responsible and inclusive assessment practices, pushing evaluative thinking beyond narrow performance metrics.

The smallest cluster (blue) has 2 words of bias and ethics, which highlights ongoing ethical concerns. This cluster presents the persistent ethical questions at the heart of research evaluation at the broader picture aside from the main discussion around research integrity (i.e., the red cluster).

The post-DORA co-occurrence network reflects a broader and more multidimensional landscape of research evaluation. While journal impact factor remains the most dominant node, the inclusion of keywords like altmetrics, gender bias, social media, and open science suggests a decentralization of evaluative authority and the emergence of reformist and inclusive discourses. Compared to the pre-DORA period, there is greater thematic integration of critical issues such as transparency,

 doi:10.20944/preprints202507.2644.v1

reproducibility, and ethical concerns. This conceptual shift points to the influence of DORA in challenging entrenched metrics and advocating for a more holistic and responsible assessment culture.

MCA-Based Conceptual Clustering After DORA

Figure 14 displays the keyword clusters from MCA of this topic after DORA. The biggest cluster (blue) has prominent keywords such as internationalization, article, criteria, social media, impact, internal validity, and questionable research practice, reflecting broad methodological and evaluative concerns. This dominant cluster captures broad concerns about evaluative validity and credibility, including internal validity, criteria for assessment, and questionable practices. The inclusion of social media and impact reflects efforts to redefine influence and outreach, while internationalization signals global alignment of assessment reforms. This cluster likely represents the mainstream conceptual shift toward evaluating the robustness, not just reach, of research outputs.

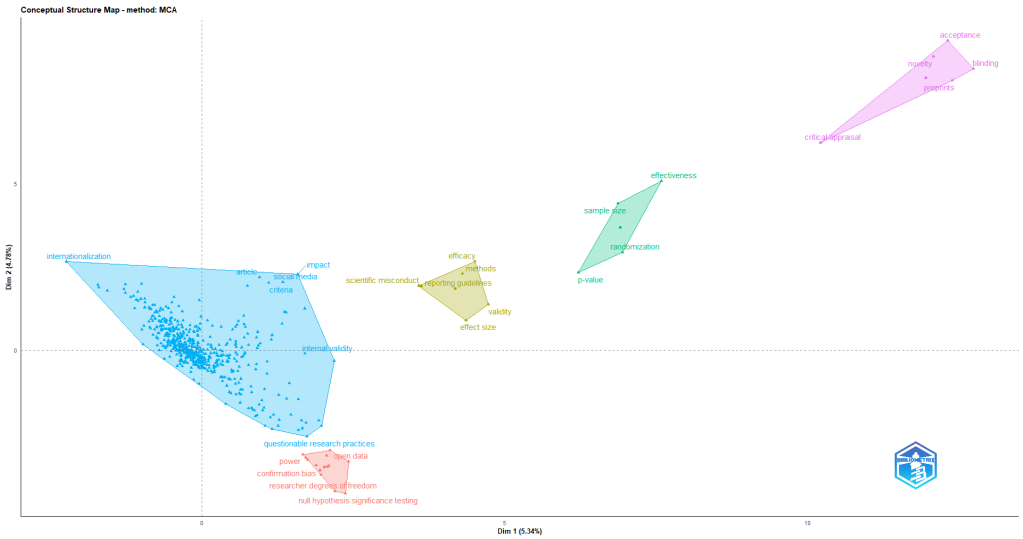

Figure 14. MCA Keyword Cluster of the Prestige-driven Metrics and Research Evaluation Topic: post-DORA.

The second cluster (red) has prominent keywords of open data, power, confirmation bias, researcher, degree of freedom, null hypothesis, and significance testing. This relates to research transparency and statistical rigor. This cluster reflects a concern with transparency and statistical reliability. Keywords like open data, confirmation bias, and degree of freedom point to debates around researcher flexibility and its influence on results, while significance testing and power relate to rigorous statistical standards. This aligns with DORA's push for accountable and open research practices, and supports broader reforms in experimental design and data reporting.

The third cluster (yellow) has prominent keywords of efficacy, methods, reporting guidelines, scientific misconduct, validity, and effect size. This cluster is strongly focused on methodological standards and ethical accountability. The co-location of reporting guidelines, effect size, and scientific misconduct suggests a post-DORA emphasis on reproducible and trustworthy evidence. It reflects increasing pressure on researchers and journals to adopt reporting norms (e.g., CONSORT, PRISMA) to ensure transparent and ethically sound dissemination.

The fourth cluster (green) has prominent keywords of sample size, effectiveness, randomization, and p-value. This compact cluster signals technical and methodological standards for empirical research, especially in experimental and clinical contexts. These terms show the post-DORA trend of valuing well-designed studies over selectively reported or underpowered ones, reinforcing calls for rigor over sensationalism.

The fifth cluster (purple) has prominent keywords of critical appraisal, novelty, preprints, blinding, and acceptance, emphasizing evaluation and innovation in research dissemination. This cluster captures the emerging discourse on innovative research dissemination and evaluative judgment. The presence of preprints and novelty suggests a growing acceptance of alternative publishing formats,

 doi:10.20944/preprints202507.2644.v1

while critical appraisal and blinding indicate efforts to systematically and fairly evaluate quality. It also reflects shifting norms in peer review and editorial practice.

The post-DORA MCA reveals a fragmented but thematically richer landscape, where evaluation is no longer viewed purely through quantitative performance indicators. Instead, the dimensions of research rigor, transparency, innovation, and methodological scrutiny are foregrounded across multiple clusters. The structural diversity of clusters, ranging from statistical validity to dissemination innovation, reflects DORA's influence in stimulating a more responsible and multidimensional discourse on how research should be assessed.

Thematic Map After DORA

Figure 15 displays the thematic map of this topic after DORA. The niche theme quadrant has two keyword clusters. Cluster 1 has the keywords of bias, risk of bias, methodology, clinical trial, and spin. This represents critical methodological scrutiny, especially around clinical trials. Terms like risk of bias reflect growing sophistication in evaluating how research is framed and interpreted, likely influenced by movements toward better reporting and reproducibility. It signals a mature but domain-specific focus. Cluster 2 has the keywords of tenure, promotion, and diversity. This cluster reflects the ongoing intersection of research evaluation with institutional equity. Although the topic continues to be niche, the addition of diversity suggests a positive interest in increasing representation.

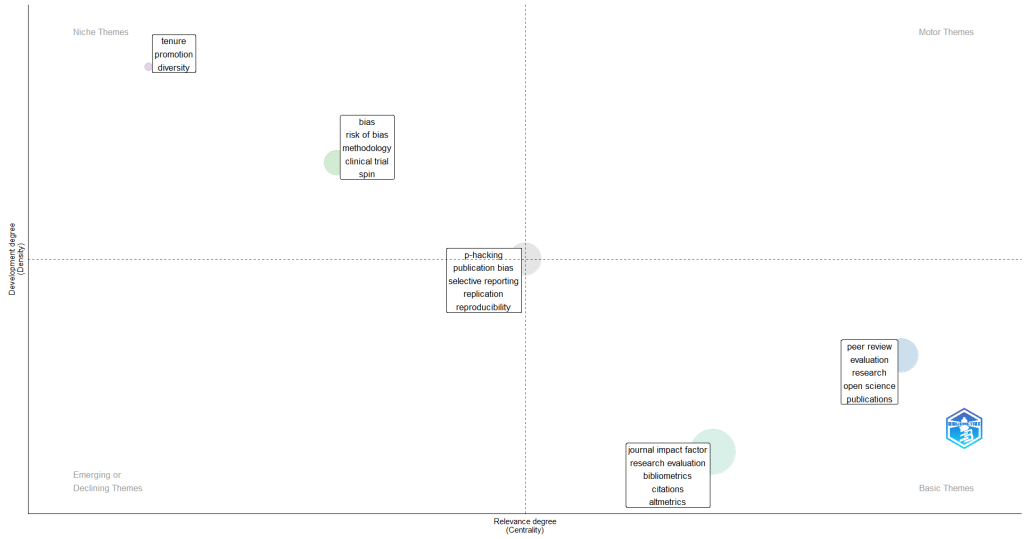

Figure 15. Thematic Map of the Prestige-driven Metrics and Research Evaluation Topic: post-DORA.

The basic theme quadrant has two keyword clusters. Cluster 1 has the keywords of peer review, evaluation, research, open science, and publications. This illustrates the fundamental building blocks of research assessment. Peer review and publication processes are paired with open science principles, suggesting that openness is now part of the foundational discourse on evaluation. It reflects the post-DORA normalization of concepts like openness and transparency as integral to the publishing system. Cluster 2 has the keywords of journal impact factor, research evaluation, bibliometrics, citations, and altmetrics. The cluster represents traditional and alternative quantitative assessment tools, including altmetrics, now recognized alongside older metrics. Despite DORA's critique of JIF, its position in the basic theme quadrant signals that metric-based evaluation remains deeply embedded, though it now coexists with competing approaches.

There is one cluster positioned between the four quadrants (motor, niche, emerging/declining, and basic themes) with keywords of p-hacking, publication bias, selective reporting, replication, and reproducibility. This position highlights its fluctuating importance which may be in the transition period. The cluster represents the core of the reproducibility crisis, linking ethical misconduct with methodological and evaluative reform.

 doi:10.20944/preprints202507.2644.v1

The thematic map reflects both continuity and transformation of the discourse. The basic theme quadrant presents foundational conversation around publication practices and peer review, while recognizing a coexistence between legacy metrics and alternative indicators; This suggests that, despite critiques, quantitative assessments remain deeply institutionalized. The niche theme has two specialized areas: one engages in domain-specific methodological critiques, while the other examines how evaluation intersects with institutional equity, marking growing awareness of representation and fairness. The cross-quadrant cluster centered on the reproducibility crisis focuses on conversations about research integrity and methodological rigor. Its transitional position indicates that concerns once peripheral are now increasingly integrated into the broader discourse on reforming research assessment in a post-DORA landscape.

Thematic Evolution After DORA

Figure 16 displays the thematic evolution diagram map of this topic before and after the emergence of DORA. During the pre-DORA period (2000-2013), there are 4 prominent themes of bias, journal impact factor, publication bias, and research. These themes suggest a period dominated by concerns about the dominance of simplistic metrics and growing awareness of distortions in research dissemination (like publication bias), but limited in terms of actionable reform tools.

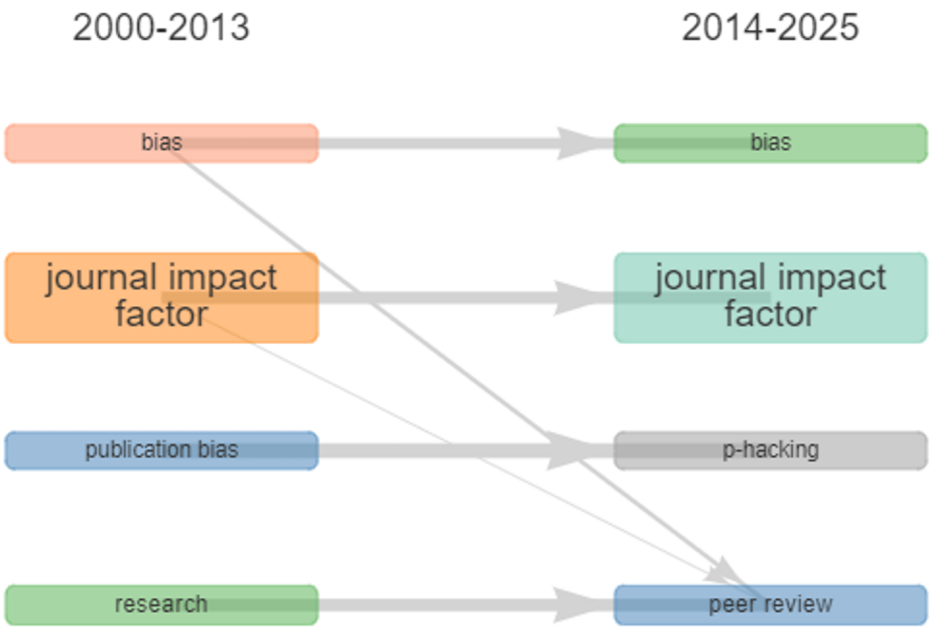

Figure 16. Thematic Evolution of the Prestige-driven Metrics and Research Evaluation Topic.

During the post-DORA period (2014-2025), there are 4 prominent themes of bias, journal impact factor, p-hacking, and peer review. These keywords reflect a shift toward introspection, transparency, and the mechanics of scientific integrity, likely by the influence of DORA and similar reform initiatives (i.e., Leiden manifesto).

In terms of linkage, using the latter period as benchmark, bias is linked with bias, indicating that structural and evaluative bias remains a persistent issue in both periods, showing little conceptual disruption. Journal impact factor is linked with journal impact factor, showing that despite DORA's critique, journal impact factor retains conceptual centrality, illustrating resistance to change in evaluation culture. P-hacking is linked with publication bias, showing how initial concerns about what gets published evolved into deeper critiques of how results are produced, a sign of increasing focus on methodological integrity. Peer review is linked with bias, journal impact factor, and research, highlighting its central role in shaping the evaluation landscape. Peer review now functions as a convergence point for many longstanding concerns; it has become the ground where biases, metrics,

 doi:10.20944/preprints202507.2644.v1

and research legitimacy all intersect. Its evolution from implicit to explicit theme reflects growing scrutiny and reform effort.

In the pre-DORA period (2000–2013), discourse centered on simplistic metrics and distortions in publishing practices, reflecting concerns over bias and publication visibility without strong reform mechanisms. In the post-DORA period (2014–2025), themes moved toward transparency and methodological integrity, with greater emphasis on practices like p-hacking and the central role of peer review. Persistent linkages such as the enduring influence of journal impact factor and structural bias illustrate continuity of the topic. Meanwhile, the emergence of peer review as a thematic hub signals its evolution into a focal point where long-standing issues around evaluation, legitimacy, and bias now converge.

5.2.5. Time Sliced Analysis: Barrier and Equity Issues in Research Accessibility Before and After Sci-Hub

Keyword Co-Occurrence Network BEFORE Sci-Hub

Figure 17 displays the keyword co-occurrence networks of this topic before Sci-hub. The red cluster contains 4 large nodes representing the keywords of scholarly publishing, academic libraries, internet, and electronic journals, with scholarly publishing as the most prominent. These terms depict the traditional infrastructure of academic dissemination prior to the emergence of shadow libraries. This thematic grouping reflects concerns related to institutional access to content, the shift toward digital resources, and the reliance on libraries and journal subscriptions.

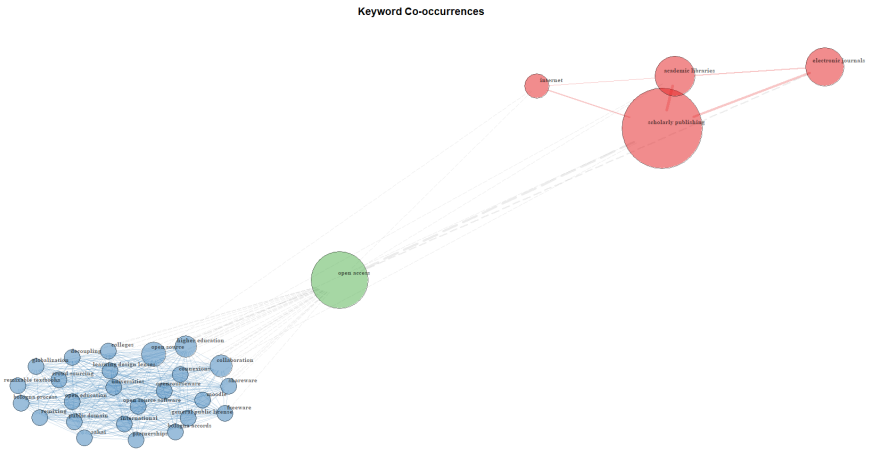

Figure 17. Keyword Co-occurrence Networks of the Barrier and Equity Issues in Research Accessibility Topic: pre-Sci-hub.

The blue cluster consists of 25 smaller nodes (keywords) of remixable textbooks, bologna process, remixing, globalization, crowdsourcing, open education, public domain, sakai, decoupling, colleges, learning design lenses, universities, open source software, open source, international, partnerships, bologna accords, general public license, opencourseware, connexions, moodle, freeware, shareware, collaboration, and higher education. This cluster reflects a conceptual space around open education, open source software, and collaborative platforms such as Moodle, Sakai, and Opencourseware. It also includes more ideological or global framing elements like the Bologna Process, globalization, and the public domain. This cluster reflects a broader, systemic discussion of knowledge democratization, with an emphasis on remixability, international cooperation, and decentralized innovation in education.

Additionally, there is one isolated keyword, open access, which serves as a linking node between these two clusters. The term open access stands alone between the two clusters, serving as a potential conceptual bridge. Its isolated position at this stage signals that while open access was recognized,

 doi:10.20944/preprints202507.2644.v1

it had not yet been fully integrated into either the institutional publishing infrastructure or the open educational ecosystems.

Before the emergence of Sci-Hub, the discourse on barriers and equity in research accessibility was somewhat fragmented. Institutional models of publishing (with access barriers rooted in subscription models) and educational movements toward openness (grounded in pedagogy and collaboration) existed in parallel but were not strongly integrated. Open access, though conceptually critical, had not been incorporated into the roader conversation.

MCA-Based Conceptual Clustering Before Sci-Hub

Figure 18 displays the keyword clusters from MCA of this topic before Sci-hub. The biggest cluster (blue) has prominent keywords such as electronic journals, academic libraries, scholarly publishing, and higher education. This cluster captures the dominant structure of knowledge dissemination before Sci-Hub that relies heavily on institutional access mechanisms. The emphasis on academic libraries and electronic journals signals a focus on digital transformation within established systems.

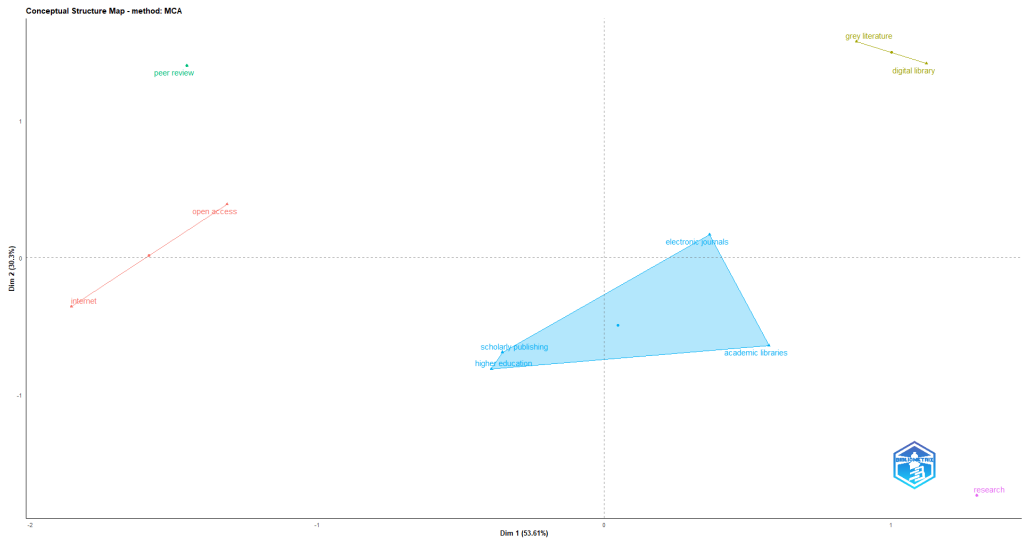

Figure 18. MCA Keyword Cluster of the Barrier and Equity Issues in Research Accessibility Topic: pre-Sci-hub.

There are two smaller clusters of similar size. The yellow cluster has prominent keywords of grey literature and digital library. It reflects emerging alternatives outside mainstream publishing. These terms represent non-traditional, often publicly accessible forms of scholarly communication. Their presence could represent growing recognition of the importance of diversified content and non-commercial dissemination strategies.

The red cluster has prominent keywords of internet and open access. The red cluster reflects early discussions of technological democratization in scholarly publishing. While small, it signifies the conceptual foundation of the open access movement, connecting the internet's infrastructure with broader aspirations for equitable access. However, its limited size and separation from larger clusters indicate that open access had not yet been mainstreamed into systemic academic discourse.

There is also an isolated keyword cluster represented by peer review. The term peer review stands as an isolated node, thematically disconnected from the rest. This suggests that while peer review was clearly part of scholarly publishing, it was not yet being conceptually linked with issues of access, equity, or systemic reform.

The MCA output portrays a landscape representing institutional infrastructure, with limited integration between traditional academic systems and emerging access movements. Academic libraries and digital journals dominate the discourse, while open access appears as a disconnected but emerging idea, not yet central to mainstream thinking. This fragmentation reflects a conceptual field not yet unified under shared concerns of equity or resistance to paywalls.

 doi:10.20944/preprints202507.2644.v1

Thematic Map Before Sci-hub

Figure 19 displays the thematic map of this topic before Sci-hub. The motor theme quadrant has three keyword clusters. Cluster 1 has scholarly journals, commercial publishers, and knowledge production. It reflects a critical focus on the traditional knowledge economy, where commercial publishers play a central role in controlling the dissemination of scholarly outputs. The tension between scholarly work and its commodification is apparent, and the positioning as a motor theme suggests that debates around access and control were already central, even before Sci-Hub's disruption. Cluster 2 has databases and education. It reflects the infrastructure side of access (i.e., how educational systems rely on database subscriptions). Cluster 3 has cost, commercial publisher, freedom of information, licensing, and private sector. The cluster represents a more explicit critique of the financial and legal barriers in academic publishing. This cluster shows how cost and licensing regimes, dominated by the private sector, were already key friction points, which is an issue Sci-Hub would later exploit and popularize.

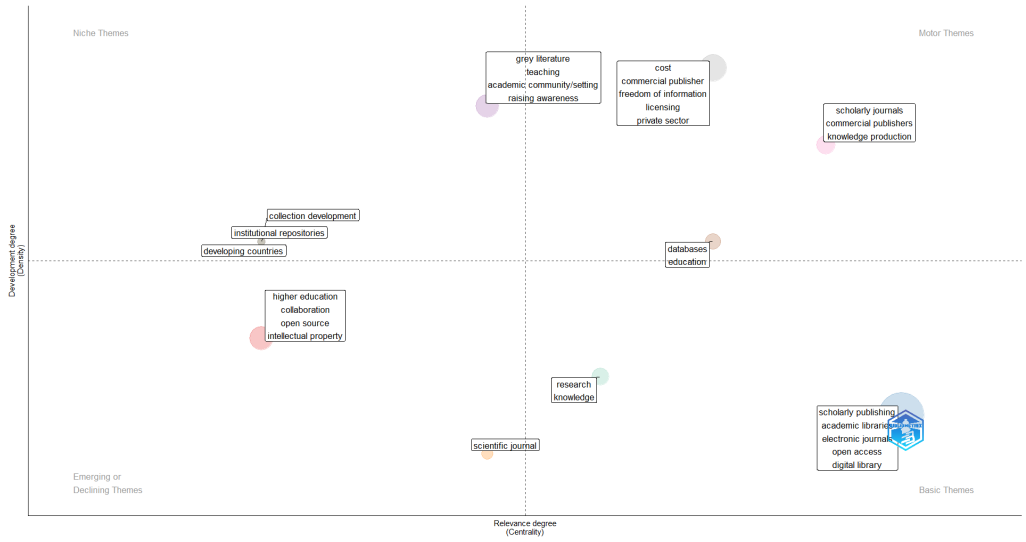

Figure 19. Thematic Map of the Barrier and Equity Issues in Research Accessibility Topic: pre-Sci-hub.

The niche theme quadrant has one keyword cluster with the keywords of grey literature, teaching, academic community/setting, and raising awareness. The cluster reflects educational reform and advocacy efforts focused on teaching about alternative forms of literature and access. It shows that grey literature was valued for its accessibility and practical relevance, even if it was not yet mainstream. This quadrant also has isolated keywords forming their own cluster of: collection, institutional repositories, and developing countries. The fragmented keywords represent region-specific and infrastructural challenges, especially for the Global South. The presence of "developing countries" as isolated hints at underdeveloped discourse on global access inequality.

The emerging or declining theme has one keyword cluster with the keywords of higher education, collaboration, open source, and intellectual property. The cluster suggests an early interest in open-source collaboration and the legal frameworks surrounding access. These were possibly emerging debates,which is still marginal but pointing toward the changes Sci-Hub would catalyze. This quadrant also has isolated keywords of scientific journal, forming their own cluster. This isolated keyword could represent a default object of study that is not integrated in a broader conversation.

The basic theme has two keyword clusters. Cluster 1 has scholarly publishing, academic libraries, electronic journals, open access, and digital library. This cluster represents the technical and institutional mechanism of the access conversation. Cluster 2 has keywords of research and knowledge. It represents general foundational terms of the discussion, which is important but not deeply connected to the subject matter.

 doi:10.20944/preprints202507.2644.v1

The thematic map reflects a landscape of the pre-disrupted research access status-quo. The motor themes capture entrenched debates around the commodification of scholarly knowledge, systemic dependence on institutional infrastructures, and mounting criticism of cost and legal barriers, which indicates that access and control were central concerns even prior to Sci-Hub's emergence. The niche theme points to pockets of advocacy and educational reform, where alternative literatures and access models, particularly those relevant to underserved regions, were promoted but remained marginal. Within the emerging or declining theme, early attention to open-source solutions and intellectual property frameworks foreshadowed structural shifts that would soon redefine access debates. Meanwhile, the basic theme anchors the discourse in institutional practices and conceptual foundations, emphasizing the role of libraries, digital platforms, and scholarly communication as core components of the evolving access ecosystem.

Keyword Co-Occurrence Network After Sci-Hub

Figure 20 displays the keyword co-occurrence networks of this topic after Sci-hub. The biggest cluster (blue) has 16 words of open access (the biggest node), scholarly publishing (second biggest node), article processing charge (third biggest node), and smaller nodes of directory of open access journals (doaj), predatory journals, journal impact factor, peer review, gold open access, bibliometrics, predatory publishing, scopus, scientific journal, institutional repositories, scholarly journals, hybrid journals, and transformative agreements. This cluster represents the formal infrastructure and economics of open access publishing, showing how the landscape diversified and became more contested after Sci-Hub's rise. The largest cluster (i.e., open access) reflects the formalization and diversification of the open access ecosystem. Prominent keywords such as "scholarly publishing," "article processing charge (APC)," and "directory of open access (doaj)" indicate that the conversation has moved beyond whether open access should exist to how it is structured and financed. The inclusion of "journal impact factor" within this cluster also shows the continuing entanglement between access and prestige-oriented evaluation systems. The growing focus on APCs highlights an equity paradox in which the paywall to read may be dissolving, but new financial barriers to publishing are emerging, especially for underfunded researchers.

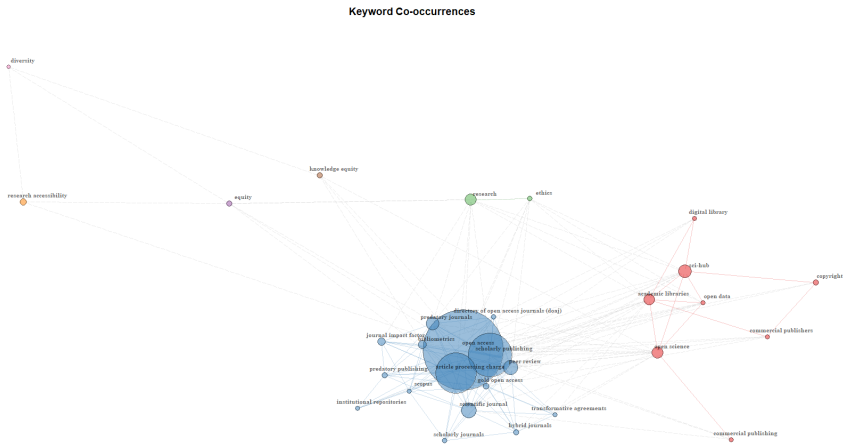

Figure 20. Keyword Co-occurrence Networks of the Barrier and Equity Issues in Research Accessibility Topic: post-Sci-hub.

The smaller red cluster has 8 keywords of digital library, sci-hub, copyright, open data, academic libraries, commercial publishers, open science, and commercial publishing. This cluster reflects tensions between traditional publishing structures and disruptive forces. Sci-Hub, as a node in this cluster, could represent a pivotal disruption that challenged the legitimacy of paywalled knowledge and catalyzed broader debates about copyright, ownership, and access. This cluster captures both the

 doi:10.20944/preprints202507.2644.v1

institutional resistance to and the ethical justification for alternative models of scholarly dissemination. The appearance of "open data" and "open science" alongside these terms signals a broader epistemic shift, one where knowledge production is increasingly framed as a public good rather than a privatized commodity.

The smallest green cluster has 2 keywords of research and ethics. This cluster represents an emerging normative framing in which access to knowledge is treated not only as a technical or economic issue, but as an ethical imperative. There are 4 isolated keywords, each forming their own cluster. Those words are knowledge equity, equity, research accessibility, and diversity. These terms, while not yet structurally integrated into the core of the network, their standalone presence points to a growing awareness of justice-oriented dimensions of access. These concepts are gaining discursive traction, likely in policy and advocacy spaces, even if they remain underdeveloped in the more metrics- or infrastructure-focused literature.

The post–Sci-Hub landscape reflects a significant broadening of concerns around research accessibility, from infrastructural debates to ethical and political critiques. Sci-Hub did not just expose cracks in the system; it provoked new lines of inquiry about who benefits from current publishing models, who is excluded, and how global knowledge infrastructures can be made more equitable.

MCA-Based Conceptual Clustering After Sci-Hub

Figure 21 displays the keyword clusters from MCA of this topic after Sci-hub. The biggest cluster (blue) has prominent keywords such as subscriptions, licensing, publishing model, creative commons, mega journals, green, self-archiving, gold oa, metrics, altmetrics, impact, india, bibliometrics, citations, beall's list, publishers, h-index, elsevier, springer, repository, black open access, research evaluation, databases, grey literature, data sharing, intellectual property, information literacy, piracy, shadow libraries. This cluster maps the wide range of legal and economic mechanisms through which access to scholarly knowledge is structured (e.g., "subscriptions," "licensing," "publishing model".) The inclusion of both "green" and "gold OA," "self-archiving," "mega journals," and "repositories" indicates a growing diversity in open access pathways. Meanwhile, the co-occurrence of "Beall's list," "piracy," "shadow libraries," "black open access," and commercial publishers like "Elsevier" and "Springer" speaks to increasing awareness of both legitimate and illicit alternatives to traditional publishing, many of which were amplified in visibility and usage following Sci-Hub's rise. Keywords such as "impact," "metrics," "altmetrics," "citations," "h-index," and "research evaluation" also show how evaluation systems remain deeply intertwined with accessibility issues; This raises concerns about how visibility, prestige, and access co-produce inequality in research access.

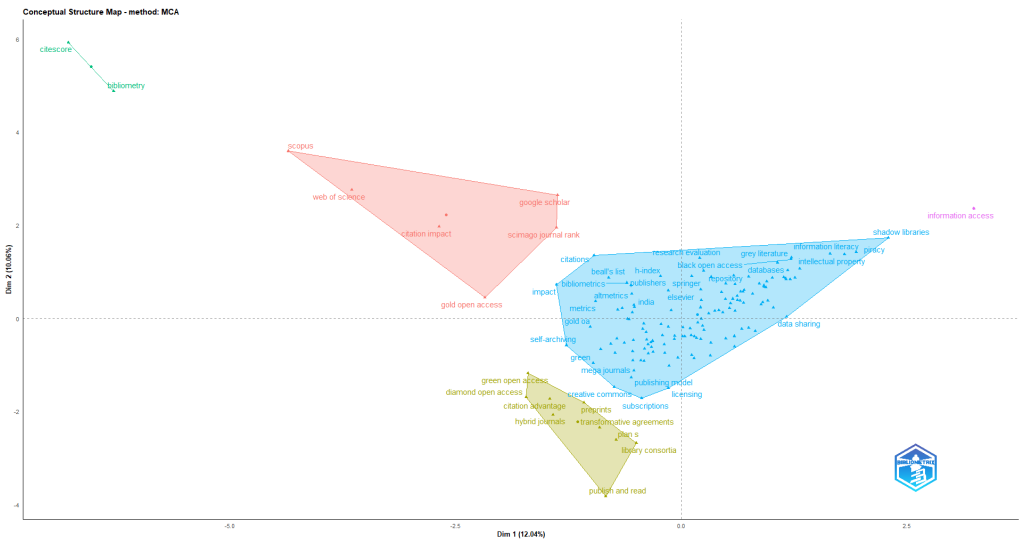

Figure 21. MCA Keyword Cluster of the Barrier and Equity Issues in Research Accessibility Topic: post-Sci-hub.

 doi:10.20944/preprints202507.2644.v1

The second biggest cluster (red) has prominent keywords such as scopus, web of science, citation impact, gold open access, scimago journal rank, and google scholar. This cluster situates access debates within the context of established indexing services and impact measurement systems. The keywords imply that accessibility isn't only about cost or copyright, but also about discoverability and institutional legitimacy. By linking open access to prestige-based metrics and citation systems, this cluster highlights ongoing tensions between openness and entrenched systems of academic valuation.

The third biggest cluster (yellow) has prominent keywords such as green open access, diamond open access, citation advantage, hybrid journals, preprints, transformative agreements, plan S, library consortia, and publish and read. This cluster implies policy-level and collaborative dimensions of open access infrastructure. Keywords such as "green open access," "citation advantage," and "preprints" reflect the operational complexities of open publishing models. The presence of policy-oriented terms like "transformative agreements," "Plan S," "publish and read," and "library consortia" suggests increased institutional and governmental engagement in shaping access models. This cluster illustrates a shift from grassroots or individual-level access work (e.g., Sci-Hub use or institutional repositories) toward systemic, policy-driven solutions intended to restructure scholarly communication globally.

The fourth biggest cluster (green) has prominent keywords such as cite score and bibliometry. This cluster focuses on specific evaluative tools and bibliometric techniques, showing continued analytical interest in how research outputs are assessed. There is one isolated keyword of information access. This keyword is a small but conceptually important node. Its isolation suggests that, despite being conceptually central to the entire discussion, "information access" may not yet be fully integrated into the broader technical, evaluative, and policy frameworks dominating post–Sci-Hub discourse. However, its presence on the map indicates that fundamental questions about access remain active, even if fragmented across different lines of inquiry.

The post–Sci-Hub MCA shows that access-related scholarship has evolved from simple dichotomies of open versus closed into a multi-dimensional terrain shaped by economics, policy, legality, technological infrastructure, and evaluative paradigms. Sci-Hub may have catalyzed the urgency of these discussions, but the current landscape reflects a maturing and increasingly institutionalized engagement with the complexities of access and equity in scholarly communication.

Thematic Map After Sci-Hub

Figure 22 displays the thematic map of this topic after Sci-hub. The motor theme quadrant has five keyword clusters. Cluster 1 has open access, scholarly publishing, article processing charge, academic libraries, and gold open access. It represents the ongoing conversations about open access infrastructure, funding models, and library roles within academic institutions. This theme underscores how open access has moved from being a disruptive idea to a normalized and structured part of scholarly publishing. Cluster 2 has predatory journals, scientific journal, peer review, bibliometrics, and journal impact factor. It represents concerns about quality control and research evaluation in an increasingly open ecosystem. It reflects the tension between expanding access and maintaining rigorous standards (e.g., predatory journal). Cluster 3 has open science, ethics, open data, institutional repository, and china. It represents growing internationalization and ethical dimensions of research dissemination, especially regarding data transparency and national participation in open movements. Cluster 4 has electronic journals, library consortia, and collection development. It highlights technical and collaborative elements for scholarly resource management in the digital age. Cluster 5 has citations, collaboration, funding, and publications. It reflects how scholarly visibility and access are shaped by collaborative structures and funding policies.

 doi:10.20944/preprints202507.2644.v1

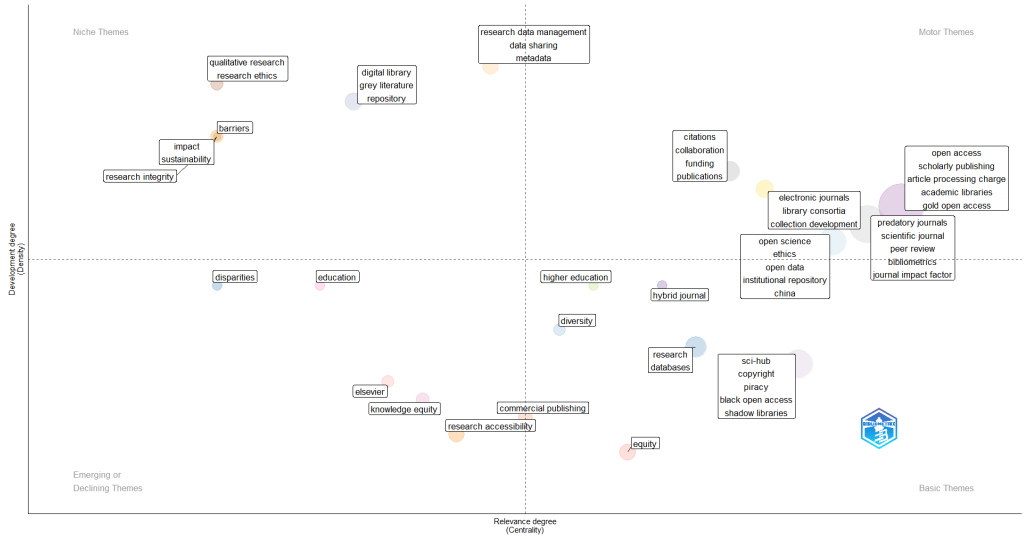

Figure 22. Thematic Map of the Barrier and Equity Issues in Research Accessibility Topic: post-Sci-hub.

The niche theme quadrant has four keyword clusters. Cluster 1 has research data management, data sharing, and metadata. It represents the growing importance of data infrastructure and its transparency in scholarly accessibility conversations. Cluster 2 has digital library, grey literature, and repository. It represents conversation about alternative or non-mainstream knowledge dissemination systems. Cluster 3 has qualitative research and research ethics. It represents methodological and ethical subthemes that are less central but still relevant to the broader conversation of research. Cluster 4 has impact and sustainability. It represents long-term structural concerns about the viability of prestige metric (e.g. impact) in the scientific dissemination ecosystem. This quadrant also has isolated keywords, each forming their own cluster of barriers, and integrity. These two words suggest areas that may be relevant to the topic but thematic integration remains limited.

The emerging or declining theme quadrant has several isolated keywords, each forming their own cluster of: disparities, education, elsevier, knowledge equity, and research accessibility. The isolated terms in this quadrant represent topics that are relevant to the overall conversation (e.g., elsevier as the commercial publisher or knowledge equity as a symbolic representation of the entrenched commercial publishing system) but have yet to be integrated with the subject matter.

The basic theme quadrant has two keyword clusters. Cluster 1 has sci-hub, copyright, piracy, black open access, and shadow libraries. It represents the core of post–Sci-Hub discourse. These terms directly reflect the realities and controversies of alternative access methods, illustrating how the rise of shadow libraries has redefined debates around legality, ethics, and access equity. Cluster 2 has research, and databases. It represents enduring, foundational discussions around how scholarly information is stored, retrieved, and managed. This quadrant also has isolated keywords, each forming their own cluster of: hybrid journal, higher education, diversity, and equity. These keywords represent issues that are recognized but remain peripheral in more structured debates.

There is also one isolated keyword, forming their own cluster of: commercial publishing positioned between the emerging or declining theme quadrant and basic theme quadrant. This position indicates the transitional standing of this topic in the broader literature landscape.

The thematic map reflects a maturing but increasingly fragmented discourse. The motor themes reflect how open access has evolved into an institutionalized norm, encompassing debates on funding models, quality assurance, international participation, and collaborative infrastructure. The niche themes reveal sustained but specialized interest in data transparency, nontraditional dissemination pathways, and ethical concerns, suggesting depth but limited connectivity to mainstream discussions. The emerging or declining themes consist largely of isolated concepts that, while relevant, remain theoretically diffuse and under-integrated. Meanwhile, the basic and basic-emerging hybrid themes present foundational conversations shaped by shadow libraries and copyright disputes, affirming

 doi:10.20944/preprints202507.2644.v1

the centrality of alternative access in the current publishing ecosystem. Overall, the map signals an increasingly complex terrain, where institutional reform, grassroots disruption, and ethical pluralism intersect but not yet harmonize.

Thematic Evolution After Sci-Hub

Figure 23 displays the thematic evolution diagram map of this topic before and after the emergence of Sci-hub. During the pre-Sci-hub period (2000-2011), there are 9 prominent themes of collection development, databases, grey literature, higher education, institutional repositories, research, scholarly journals, scholarly publishing, and scientific journal. These keywords indicate how academic content is curated, stored, and disseminated. This era reflects the groundwork of digital transition, where institutions focused on building and managing academic resources, often within the confines of subscription-based access and traditional library systems of higher education institutions.

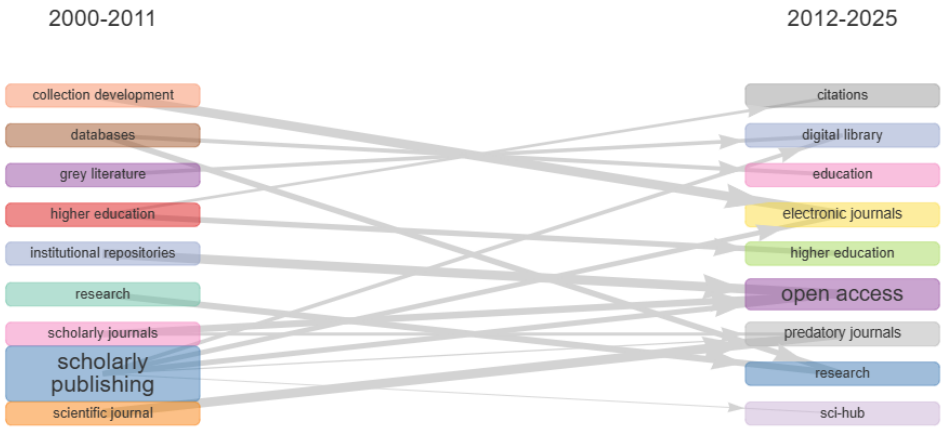

Figure 23. Thematic Evolution of the Barrier and Equity Issues in Research Accessibility Topic.

During the post-Sci-hub period (2012-2025), there are 9 prominent themes of citations, digital library, education, electronic journals, higher education, open access, predatory journals, research, and sci-hub. This change represents a transformation toward accessibility, openness, and critical evaluation of publishing systems. The emergence of new themes such as open access, predatory journals, Sci-Hub, and electronic journals points to a more direct engagement with the systemic barriers to research accessibility. The prominence of Sci-Hub itself as a standalone theme in this period is a testament to its catalytic impact on global discussions about equity, piracy, and the ethics of information access. Additionally, recurring themes such as higher education, research, digital libraries, and education show the continuity of institutional involvement, but these are now situated within a broader debate on how access intersects with impact, legitimacy, and evaluation.

In terms of linkage, using the latter period as a benchmark, citations is linked with higher education. This link reflects the growing role of bibliometrics and citation-based evaluation in shaping how access influences academic visibility and impact. Digital library is linked with grey literature and scholarly publishing. This link points to the expansion of informal or alternative dissemination channels in the wake of access limitations. Education is linked with databases. This link alludes to the broader connection of scientific data storage and management. Electronic journals is linked with collection development and scholarly publishing. This link illustrates the digitization of publishing and its impact on how academic content is curated. Higher education is linked with higher education. This link shows the persistent involvement of higher education in the conversation.

Open access is linked with institutional repositories, scholarly journals, and scholarly publishing. This link highlights its roots in pre-Sci-Hub infrastructures but also signaling its post-Sci-Hub mainstreaming of knowledge accessibility. Predatory journals is linked with scholarly journals and scientific journal. This link indicates an evolution from trusted dissemination venues to contested

 doi:10.20944/preprints202507.2644.v1

and potentially exploitative models, likely an unintended side effect of scholarly publishing norms. Research is linked with databases and research. This link also represents the broader connection of scholarly work and its channel to visibility. Scihub is linked with scholarly publishing. This link most importantly shows its disruptive position, both as a critique of the paywalled status quo and as a symptom of unmet access needs.

Overall, the thematic evolution captures a critical transition in the discourse: from building academic access structures to radical disruption in the discourse on research accessibility. In the pre–Sci-Hub period (2000–2011), scholarly conversations were grounded in institutional curation, emphasizing traditional publishing, repositories, and library systems; This reflects a subscription-bound model of controlled academic dissemination. By contrast, the post–Sci-Hub period (2012–2025) marks a pivot toward openness, evaluation, and critique. Themes such as open access, predatory journals, and Sci-Hub itself point to intensified scrutiny of publishing hierarchies and access inequalities. The emergence of Sci-Hub as a dominant theme underscores its catalytic role in challenging the paywall paradigm. Linkages across the periods reveal both continuity and disruption: while infrastructures like higher education and digital libraries persist, new connections, such as between citations and academic visibility or between open access and institutional repository signal a transformed landscape. Ultimately, the diagram captures an evolving discourse that moves from institutional stewardship to contested, decentralized models of knowledge access.

5.3. Literature Landscape Through Result Integration

Table 5 juxtaposes key characteristics of the literature landscape of the three topics, as informed by the integration of both performance- and conceptual analysis results.

Table 5. Key Characteristics of the Literature Landscape.

	Insights from performance analysis	Insights from conceptual analysis	Key characteristic of the discourse from integration
Topic 1	Highest growth across topics, indicating a steadily increasing scholarly interest. The discourse is not only empirical but also reflective and opinion-driven US is the dominant country	Focus on peer review systems and editorial practices Clusters reveal discourses on editorial bias, inclusion, and power asymmetries in the scholarly communication process. Recurring keywords are "peer review," "publication bias," "prestige bias"	Discourse is critical and reform-oriented The main theme is transparency, bias, power in editorial systems Links to epistemic justice and legitimacy in academia
Topic 2	Largest number of literature among all topics. Active discussion as seen from the volume of conference proceeding and citation rate. US is the dominant country Sources represent a mix of methodological and policy-focused journals.	Focus on impact factors, rankings, citation metrics Clusters reveal the discourse of scholarly metrics, its limitations, and the following problem. Recurring keywords are "impact factor," "research evaluation," "metrics"	Critique of metric-based academic culture The main theme is performance pressures, questionable research practices, scholarly metricization Mix of critical reassessment and institutional entrenchment
Topic 3	Relatively lowest scholarly interest as seen from annual growth rate and citation rate. US is the dominant country Sources focus on the library and information science area.	Focus on access inequities, APCs, copyright, predatory publishing Clusters reveal concerns of economic and legal barriers to research accessibility, as well as grassroot responses. Recurring keywords are "open access," "APC," "predatory publishing"	Equity and ethics at the center of the discourse The main theme is institutional reform, funding, informal infrastructures Tensions between open science ideals and exclusionary research access

5.3.1. Gatekeeping and Editorial Bias

The discourse on gatekeeping and editorial bias in scholarly publishing has gained traction over the past two decades, as evidenced by a robust annual growth rate of 10.7% in the topic's literature, despite it constituting the smallest dataset among the analyzed topics. This upward trend reflects an increasing scholarly engagement with the mechanisms and consequences of editorial decision-making. The predominance of journal articles, supplemented by a substantial number of review pieces and editorials, indicates a discourse that is both empirically grounded and normatively reflective, aligning with broader concerns about the credibility, fairness, legitimacy, and transparency of peer review systems.

Conceptual analysis reinforces this characterization by revealing a landscape structured around the intersection of institutional, ethical, and systemic dimensions of editorial practice and its entailing issue. Central themes revolve around the operational dynamics of peer review and editorial judgment, while distinct clusters address socio-structural issues such as identity-based bias, inclusion efforts, and the asymmetric power relations between editors and authors. The thematic map further delineates the discourse into motor themes advocating reform and transparency, and niche inquiries into bibliometric accountability, and foundational concerns (i.e., basic theme) about prestige-based gatekeeping and insider bias. Recurring keywords, such as "peer review," "publication bias," and "prestige bias", signal enduring anxieties over the equity and credibility of scholarly publishing. Geographically, the United States emerges as a dominant site of authorship and discourse production, reflecting its influential role in shaping editorial norms. Collectively, the conceptual and performance analyses

 doi:10.20944/preprints202507.2644.v1

depict a discourse increasingly preoccupied with interrogating the structures, practices, and values that underpin editorial gatekeeping and its implications for epistemic justice in academia.

5.3.2. Prestige-Driven Metrics and Research Evaluation

The discourse surrounding prestige-driven metrics and research evaluation has emerged as a prominent and expansive area of scholarly inquiry, as evidenced by the dataset's substantial volume. This volume, coupled with relatively high citation rate, could imply the intensifying global examination of performance indicators such as journal impact factors, citation counts, and institutional rankings. The substantial presence of reviews and conference papers indicate active discussion of the topic. The dominant national affiliation of corresponding authorship in the United States, further reflects concentrated scholarly and policy interest in systems of academic valuation. Top publication venues align with the field's methodological and policy-oriented engagement.

Conceptually, the discourse focuses on the pervasive influence of quantitative research indicators within academic systems, and the consequences of their institutionalization. Keyword co-occurrence patterns reveal that core concerns revolve around the use of impact factors, peer review, and other metrics within higher education and publication systems, as well as their entanglement with research biases and integrity threats such as p-hacking and selective reporting. The MCA clustering highlights similar areas of inquiry, with the addition of innovation in research and acceptance processes. The thematic map further clarifies the structure of this discourse: the motor-basic hybrid theme captures both the dominance and critical reassessment of performance metrics in defining scholarly worth, while the emerging-basic cluster identifies growing concern over the institutional pressures that may foster questionable research practices. The niche theme points to the metricization of academic labor, where hiring, promotion, and funding decisions are increasingly tethered to quantitative evaluations of prestige. Together, these findings illustrate a discourse that critically examines the epistemic, methodological, and institutional consequences of a metric-oriented academic culture.

5.3.3. Barrier and Equity Issues in Research Accessibility

The discourse on barriers to research accessibility and equity in knowledge dissemination, while showing the slowest annual growth rate among the three topics, still conceptually reflects a growing scholarly engagement with the structural and normative questions surrounding who gets to access and contribute to academic knowledge. While the United States again led in corresponding authorship, the global south likely features prominently in the thematic scope due to the direct impact of access inequities on researchers in resource-limited contexts. Key publishing venues focus on scholarly communication and library science. Top keywords reflect ongoing tensions between formal publishing infrastructures and alternative models of access.

Conceptual analysis reveals a discourse fundamentally concerned with the ethical, structural, and economic dynamics that underpin academic publishing. The co-occurrence network highlights tensions between commercial publishing interests, access mechanisms, and the pursuit of scholarly inclusivity. Central themes include financial barriers such as article processing charges, ethical dilemmas like predatory publishing and copyright restrictions, and the infrastructural roles of libraries and repositories. The MCA clustering points out the systemic nature of access barriers, linking them to entrenched prestige hierarchies and global disparities in publishing capacity. The thematic map presents a discourse distributed across several conceptual terrains: the motor theme quadrant emphasizes institutional and financial debates over equitable access and open science funding models; the niche themes point to underexplored disparities and infrastructural challenges; and the emerging themes points toward grassroots dissemination practices and informal infrastructures, albeit in a fragmented manner. The basic themes reinforce the foundational role of both formal and informal systems in shaping how, and for whom, academic knowledge circulates. Together, these findings portray a field grappling with the tension between democratizing ideals and exclusionary research access, revealing the urgent need for structural reform in how access to knowledge is ethically justified.

 doi:10.20944/preprints202507.2644.v1

5.4. Key Literature Changes Through Result Integration

5.4.1. Prestige-Driven Metrics and Research Evaluation Before and After DORA

Table 6 juxtaposes key characteristics of the literature landscape of the prestige-driven metrics and research evaluation topic BEFORE and AFTER the emergence of San Francisco DORA, by integrating insights from both performance- and conceptual analysis results. The discourse on prestige-driven metrics and research evaluation has undergone a significant conceptual and structural transformation across the pre- and post-DORA periods, as evident from the shift from metric-centric focus to more reflexive and reform-oriented critique.

Table 6. Key Literature of the Prestige-driven Metrics and Research Evaluation Topic

	Insights from performance analysis	Insights from conceptual analysis
Pre-DORA	Rapid publication growth High number of conference proceedings, implying topic engagement Dominated by USA	Evaluation centered on quantification and hierarchy Journal impact factor viewed as legitimate proxy for value Early concerns (e.g., selective reporting) were marginal Critiques were emergent and peripheral
Post-DORA	Relatively lower engagement from growth rate and number of conference proceedings. More diversified discussion as seen from the distribution of document type. Dominated by China	Reflexive and multidimensional discourse Focus on methodological integrity, ethics, and equity Research integrity and peer review now central Emphasis on rethinking research evaluation
Key change from Integrated insight	Shift from a metrics-dominated, performance-driven model to a more reflexive, ethically grounded, and pluralistic evaluation paradigm. Metrics are now contextualized within broader concerns of transparency, fairness, and methodological rigor. Keyword shifts to conversation in examining the processes and incentives that shape scientific outputs.	

In the pre-DORA era (2000–2013), scholarly engagement was characterized by rapid growth and a relatively narrow conceptual focus. The landscape was dominated by traditional bibliometric indicators, particularly the Journal Impact Factor, which served as both a practical and symbolic anchor for institutional research assessment. Thematic clustering during this period reflected a hierarchical and quantitatively-driven evaluative practice, wherein citation-based metrics were largely accepted as legitimate proxies for scholarly value. Although early concerns about methodological distortion and selective reporting had begun to surface, these remained peripheral to a dominant narrative that privileged performance over process. The thematic map of this period captured a landscape of normative entrenchment, where evaluation was institutionalized, quantification was uncontested, and critiques of the system were largely emergent rather than mainstream.

Following the publication of the San Francisco DORA in 2013, the discourse experienced both expansion and conceptual reorientation. While publication volume grew considerably in the post-DORA period (2014–2025), it did so with a slower trajectory, suggesting a maturation and possible saturation of the field. More significantly, the thematic architecture of the post-DORA landscape became increasingly multidimensional. The discourse became widespread as seen from the shift in dominant country and diversified document type. Traditional metrics like impact factor remained influential, but they now coexisted with an expanding vocabulary of reform, including altmetrics, open science, research transparency, and p-hacking. Conceptual analysis revealed a diversification of evaluative concerns, as the once-dominant bibliometric discourse was challenged by emerging clusters emphasizing methodological integrity, ethical scrutiny, and institutional equity. Thematic maps demonstrated how previously marginal concerns, such as research integrity and reproducibility, gained centrality and began to reshape the evaluative conversation. Peer review, in particular, evolved from a supporting evaluative mechanism into a foundational theme, reflecting an ongoing examination and scrutiny.

The thematic evolution across periods implies both continuity and change in the literature. On one hand, the enduring prominence of journal impact factors and persistent concerns about structural bias illustrate the deep institutional embeddedness of prestige-based metrics. On the other hand, the post-DORA discourse reflects a broadening of what constitutes legitimate research assessment. Rather than a rejection of quantitative indicators, the post-DORA era is characterized by their recontextualization, where metrics are no longer seen as ends in themselves, but as tools that must be scrutinized within broader ethical, methodological, and social frameworks. DORA's influence is most evident in the discursive shift toward transparency, equity, and rigor, as well as in the decentralization of evaluative authority across methodological, epistemic, and geographic lines. This

 doi:10.20944/preprints202507.2644.v1

transition from a metrics-dominated paradigm to a more holistic and contested evaluative landscape marks a significant inflection point in the scholarly discourse on research assessment.

5.4.2. Barrier and Equity Issues in Research Accessibility Before and After Sci-Hub

Table 7 juxtaposes key characteristics of the literature landscape of the barrier and equity issues in research accessibility topic BEFORE and AFTER the emergence of Sci-hub, as suggested by the integration of both performance- and conceptual analysis results. The emergence of Sci-Hub changed the discourse surrounding barriers and equity in research accessibility, transforming a relatively institution-bound conversation into a contested, politically relevant, and increasingly decentralized debate.

Table 7. Key Literature of the Barrier and Equity Issues in Research Accessibility Topic.

	Insights from performance analysis	Insights from conceptual analysis
Pre-Sci-hub	Relatively smaller starting point interest from high growth but low citation rate Dominated by USA Sources are library-oriented field	Framed as infrastructure and digitization issue Themes focused on institutional roles (e.g., libraries) Open access was fragmented and peripheral Debate largely within traditional publishing structures Keywords imply discussion around the infrastructure of research access.
Post-Sci-hub	Relatively higher engagement with the topic as seen from nearly doubled citation rate albeit lower growth rate. Dominated by USA Sources include science evaluation and communication venues.	Shift to systemic, ethical, and legal critique Central themes are APCs, piracy, predatory journals, open access Motor themes now include reform and inclusivity Expanded discourse on grassroots dissemination and data ethics Keywords imply economic and political-oriented discourse, with Sci-hub as a key object.
Key change + Integrated insight	Shift from technical/institutional access issues to a politicized, systemic critique of publishing norms. The field now engages with ethical, economic, and global dimensions of access, driven by Sci-Hub's role in exposing structural inequities and legitimizing informal knowledge sharing.	

In the pre–Sci-Hub period (2000–2011), scholarly attention to access issues was modest, with slow but steady growth and low citation impact, signaling that the topic was largely peripheral within the academic community. The discourse was institutionally grounded, dominated by themes related to library systems, digital repositories, and electronic publishing. Conceptually, access was framed as a technical or infrastructural concern, embedded in the digital transition and rooted in institutional stewardship. Thematic maps and keyword networks reveal fragmented debates, with open access still an emerging and isolated concept, largely unintegrated into mainstream publishing norms. The focus remained on building formal structures within the traditional publishing ecosystem rather than disrupting it.

Post–Sci-Hub (2012–2025), the landscape shifted both quantitatively and conceptually. While the publication growth rate slowed down, the volume of literature increased substantially, and average citation rates nearly doubled, indicating heightened scholarly interest and recognition of the issue's systemic relevance. Sci-Hub itself emerged as a central node in both keyword networks and thematic structures, catalyzing a transition from infrastructural improvement to ethical critique and resistance. The discourse evolved to encompass economic models (e.g., article processing charges), legal challenges (e.g., copyright and piracy), and informal access mechanisms. Importantly, the vocabulary broadened: keywords such as predatory journals, Sci-Hub, and open access now indicate a more politicized framing of access inequities, with global implications.

The thematic evolution results reinforce this paradigm shift. Pre–Sci-Hub discussions were confined to custodial roles of libraries and institutions in navigating the digital transition, while post–Sci-Hub discourse interrogates the legitimacy and inclusivity of the entire publishing model. Motor themes now reflect institutionalization of open access and a turn toward systemic reform, while emerging and niche themes, though scattered, suggest expanding interest in grassroots knowledge practices, data ethics, and alternative dissemination. Continuities such as ongoing reliance on institutional repositories coexist with disruptions, most notably, the normalization of Sci-Hub as both a symptom and critique of exclusionary access models. In sum, the Sci-Hub moment did not merely introduce a new actor into the scholarly communication ecosystem, it redefined the terms of debate. The field has transitioned from discussions of how knowledge should be accessed within institutional frameworks, to who gets to access knowledge, under what conditions, and through which mechanisms. This has produced a more mature, though fragmented, discourse that integrates political economy, ethical pluralism, and infrastructural critique, marking a decisive transformation in how the academic community conceptualizes and responds to barriers in research accessibility.

 doi:10.20944/preprints202507.2644.v1

6. Discussion

This study examines how the scholarly discourse on academic publishing reform has evolved between 2000 and 2025, with particular attention to three interconnected domains: editorial gatekeeping and bias, prestige-driven metrics and evaluation, and barriers and equity issues in research accessibility. Through bibliometric analysis, we mapped thematic developments, conceptual structures, and changes in discourse over time. The results reveal both persistent concerns and discourse transformation shaped by major interventions such as the calls for research assessment reform (e.g., San Francisco DORA or Leiden Manifesto) and the emergence of Sci-hub shadow library.

6.1. Who Decides What Knowledge Is? Editorial Elitism in Scholarly Publishing

The discourse on editorial gatekeeping and bias, though constituting the smallest dataset, exhibited the fastest growth rate, indicating rising scholarly concern over the fairness and transparency of editorial decision-making. The findings reveal a conceptual field structured around institutional, ethical, and systemic dimensions of editorial practices, with clusters focusing on identity-based bias, inclusion, and power asymmetries between editors and authors. These findings align with the broader calls for epistemic justice in knowledge production [12,51].

The prominence of keywords such as "peer review," "publication bias," and "prestige bias" reflects enduring anxieties over the equity and legitimacy of scholarly communication [3,50,60]. Together, these terms point to systemic vulnerabilities in the editorial and evaluative mechanism of the publishing ecosystem. Their continued presence across decades of literature suggests that the issue is not transient complaints, but rather chronic structural flaws in how academic discourse is curated, validated, and circulated. As reform efforts gather momentum, addressing the structural roots of these anxieties becomes essential to foster a more inclusive and epistemically just academic community.

Thematic clustering of transparency and reform in editorial processes, bibliometric accountability, and prestige-based gatekeeping are consistent with the literature's call for more democratic models of publishing, such as open peer review, preprint dissemination, and post-publication peer review systems [51,52]. Moreover, foundational concerns persist regarding insider bias and elitist editorial gatekeeping, especially within journals with invite-only submission practices, which may systematically disadvantage underrepresented scholars and global perspectives [13–15].

A geographic analysis of corresponding authorship shows the concentration of discourse production in the United States. While this reflects the country's institutional prominence in global publishing networks, it also raises questions about epistemic centralization and the marginalization of voices from the Global South and non-English-speaking contexts [3,16]. The dominance of United States-based contributions likely influences editorial norms and peer review standards, suggesting a need to diversify representation and editorial board composition to better reflect global scholarly communities. In this light, editorial bias and gatekeeping are not merely procedural concerns but represent issues of epistemic justice, shaping who participates in knowledge production and how scholarly legitimacy is constructed.

6.2. Reassessing Metrics and Meaning in Research Evaluation

The discourse surrounding prestige-driven metrics and research evaluation has also emerged as a prominent and expansive field within scholarly inquiry. This is evidenced by the substantial dataset volume and relatively high citation rates that characterize this topic. Together, these indicators suggest a globally intensifying scrutiny of quantitative performance metrics such as JIF, citation counts, and institutional rankings. Furthermore, the concentrated authorship in the United States points to a strong institutional and policy-driven interest in academic valuation mechanisms, which continue to shape global publishing norms and funding priorities [19,94].

Conceptual analysis reveals that this discourse is related to the institutionalization of quantitative indicators and their consequences for how scholarly worth is defined and measured. Common keywords like "impact factor," "peer review," and "research integrity" frequently co-occur, which

 doi:10.20944/preprints202507.2644.v1

reflects core concerns about how these metrics govern academic publishing and career progression [7]. The reliance on these metrics are associated with threats to research integrity such as p-hacking, selective reporting, and outcome-driven publication practices (i.e., research that focuses on getting good results rather than answering the research objective).

Clustering patterns additionally highlight ongoing inquiries about research evaluation and reform of acceptance processes. The thematic map captures a motor-basic hybrid theme that reflects both the entrenched dominance of traditional metrics and a growing critical reassessment of their legitimacy. Emerging-basic clusters further point to rising concern about institutional pressures that encourage questionable research practices, while niche themes identify the increasing metricization of academic labor. These processes are evidenced by hiring, promotion, and funding decisions being increasingly tethered to prestige-based quantitative evaluations [55,95].

The emphasis on quantifiable outputs (e.g., JIF) has inadvertently fostered a culture of expediency and strategic optimization, often referred to as "gaming the system" [64,96]. In pursuit of rapid publication and increased visibility, scholars may feel pressured to prioritize productivity over substance, leading to ethically precarious practices. Among the most recent manifestations of this trend is the accelerated use of large language models (LLMs), such as generative AI tools, to produce draft manuscripts or complete entire papers, sometimes without disclosure [97]. This shift reflects desire to meet journal thresholds, maximize citations, and secure grant funding by creating an environment where speed, not rigor, is rewarded. In this context, LLMs become instrumental in enabling superficial compliance with academic conventions, rather than facilitating deeper intellectual engagement. This phenomenon exposes a structural vulnerability: when metrics become the dominant markers of scholarly value, they incentivize performative scholarship over epistemic integrity [7,94].

The publication of the San Francisco DORA in 2013 marks a pivotal moment in the evolution of this discourse. Before DORA, between 2000 and 2013, literature on research evaluation experienced rapid growth with a narrow conceptual scope. JIF dominated both discourse and practice, operating as a practical and symbolic anchor for institutional research assessment. Thematic clusters from this era reflect a hierarchically structured system where bibliometric indicators were widely accepted as valid proxies for scholarly excellence. Concerns about metric distortion, methodological bias, and selective reporting existed but were peripheral, lacking mainstream traction. Evaluation practices during this time were deeply institutionalized, with performance quantification serving as the uncontested norm across academia [98,99].

In contrast, the post-DORA period from 2014 to 2025 is marked by both expansion in publication output and a significant conceptual transformation. Although growth rates slowed, the thematic complexity of the discourse increased, indicating a maturing and diversifying field. Traditional metrics like JIF remained influential, but they became increasingly contested and coexisted alongside reformist vocabulary such as altmetrics, research transparency, and methodological reproducibility. Emerging clusters focused on research integrity gained centrality, and peer review shifted from a supporting function to a foundational theme of evaluative inquiry. Geographic diversification and broader document types reflected a decentralization of discourse production and evaluative norms [26,100].

Taken together, this thematic evolution suggests both continuity and change. The enduring prominence of JIF and persistent structural biases illustrate the deep institutional embeddedness of prestige-driven metrics. Simultaneously, the post-DORA literature reflects an expanded understanding of what constitutes meaningful research assessment. Rather than rejecting metrics entirely, the post-DORA era is defined by their recontextualization: metrics are no longer viewed as ends in themselves but as tools requiring ethical and methodological scrutiny. DORA's impact is evident in the discourse's shift toward transparency, equity, and rigor, as well as in the diversification of evaluative authority across epistemic, methodological, and geographic domains. This transition from a metrics-dominated paradigm to a more reflexive and contested evaluative landscape marks a critical inflection point in the scholarly understanding of academic worth and institutional legitimacy.

 doi:10.20944/preprints202507.2644.v1

6.3. Rewriting the Terms of Access: Political Economy and Equity in Scholarly Publishing

Despite having the slowest annual growth rate among the three topics analyzed, the discourse surrounding barriers to research accessibility and equity has evolved into a conceptually rich and politically important area of inquiry. This trend reflects an increasing engagement with structural, ethical, and normative questions about who gets to participate in academic knowledge creation/consumption and under what conditions. While corresponding authorship remains concentrated in the United States, thematic content suggests growing relevance to researchers in the Global South, where resource constraints and exclusionary publishing models pose challenges [101]. Leading publication venues in library and information science reinforce the topic's institutional and infrastructural anchoring, while the prevalence of keywords such as "open access," "predatory publishing," and "copyright" signals ongoing conflicts between formal publishing systems and alternative models of dissemination [71,102].

Conceptual analysis reveals that this discourse is fundamentally oriented toward exploring the political economy of access. The literature reflects an enduring tension between commercial publishing interests and the pursuit of inclusive scholarly communication. Financial barriers such as APC, ethical concerns over copyright enforcement, and the proliferation of predatory journals all point to the systemic obstacles confronting researchers outside elite publishing circuits [22,73]. Additionally, keyword co-occurrence networks and MCA clustering highlight the vital yet underexamined role of libraries and institutional repositories in mitigating access disparities. These topics occupy both foundational and niche positions in the thematic map, revealing how formal and informal infrastructures are embedded within broader debates about equity, legitimacy, and openness [103].

The emergence of Sci-Hub in 2011 marks a paradigmatic shift in how barriers to access are conceptualized and contested. Prior to Sci-Hub, literature on research accessibility was modest in scale and influence. Discussions were generally confined to institutional contexts, with themes focusing on the custodial responsibilities of libraries and the technical challenges of electronic publishing. Open access existed largely as a peripheral concern, discussed in terms of digital infrastructure rather than systemic reform.

In the post-Sci-Hub era, however, the discourse underwent both quantitative expansion and conceptual transformation. The volume of publications increased substantially, and citation rates nearly doubled, reflecting greater scholarly attention to the issue's ethical and political ramifications. Sci-Hub itself emerged as a central discursive node, symbolizing both a symptom and a critique of exclusionary publishing models [24,84,85]. Conceptually, access debates moved beyond technical improvements to interrogate the legitimacy of the academic publishing ecosystem. Motor themes now emphasize institutional engagement with open access mandates and funding models, while niche and emerging clusters map out fragmented efforts to document alternative access mechanisms, data ethics, and grassroots dissemination practices [104]. Thematic maps reveal growing interest in decentralized solutions, such as informal networks of scholarly sharing, even though it remains unevenly distributed and ethically contested. This shift has broadened the vocabulary of access discourse to include terms like "piracy," "knowledge equity," indicating a more politically relevant and socially responsive framing of the issue.

Importantly, continuity persists in the reliance on institutional repositories and formal commercialized repository systems, even as they coexist with disruptive models. The normalization of Sci-Hub has challenged academic norms around legitimacy and copyright, prompting scholars to re-evaluate the ethical justifications for paywalled knowledge [85]. The dichotomy between custodial and insurgent access strategies illustrates the tension between democratizing ideals and entrenched hierarchies in scholarly publishing. At its core, this evolving discourse demands that academic institutions address not only how knowledge is disseminated but for whom, and under what conditions it remains accessible.

Together, these findings point to a decisive transformation in how research accessibility is understood within the scholarly community. The field has shifted from institution-bound dialogues about

 doi:10.20944/preprints202507.2644.v1

infrastructure to a multidimensional critique encompassing political economy, ethical pluralism, and epistemic inclusion. Reform efforts must now grapple with the ethical imperative to make knowledge not only technically accessible but also socially and structurally equitable.

7. Conclusions

This study provides a systematic bibliometric overview of how scholarly discourse on academic publishing reform, with the overarching research question of: *How has scholarly literature on academic publishing reform evolved across the themes of editorial gatekeeping and bias, prestige-driven research metrics and evaluation, and barrier and equity issues in research accessibility over the past 25 years?* The exploration of editorial gatekeeping, prestige-driven metrics, and research accessibility reveals a scholarly ecosystem marked by entrenched structural asymmetries. Editorial bias and exclusion reflect ongoing concerns about legitimacy and transparency, particularly in how decisions about academic merit are made and whose voices are amplified. Metrics such as impact factors and citation counts, while designed to measure influence, have become instruments of performative scholarship by pressuring researchers to prioritize speed, conformity, and visibility over integrity. These same metrics often drive a culture of expediency, where generative AI tools are leveraged not to support inquiry, but to simulate it. The rising use of LLM for paper content generation illustrates how a flawed system pushes scholars toward shortcuts that satisfy institutional demands while eroding methodological rigor [64].

Barriers to research accessibility compound these challenges. Scholars from resource-limited regions confront prohibitive APC, restrictive copyright regimes, and exclusion from prestigious editorial circles. The discourse increasingly recognizes that academic publishing is not just a matter of technical dissemination, but of ethical inclusion. The emergence of Sci-Hub reframed the debate around access, elevating it from an infrastructural issue to one of political resistance and systemic critique [24,82]. Post–Sci-Hub discussions interrogate who has the right to access knowledge, under what conditions, and through which mechanisms, thereby bringing discourses about grassroots dissemination and ethical pluralism to the forefront.

These three systemic problems are fundamentally interlinked. First, gatekeeping, metrics, and access do not operate independently, but feed into one another to perpetuate academic hierarchies and serve elite interests. Together they form a system that privileges conformity, rewards institutions over individuals, and marginalizes alternative epistemologies [4]. Second, the dominance of Global North scholarship risks marginalizing voices from contexts most affected by inequities in publishing systems. Third, the change in discourses as seen from the findings reflects a cultural shift toward reflexivity, transparency, and ethical accountability, though structural inertia remains a significant barrier. These findings align with theoretical perspectives that conceptualize academic publishing as a complex system in which power, prestige, and access privilege co-produce patterns of inclusion and exclusion [69,105]. What may seem like isolated scholarly experiences are in fact parts of a larger, shared struggle; a struggle faced not only by current researchers but also by future scholars and the general public who rely on credible and inclusive academic knowledge.

The study carries implications for ongoing debates about evaluation reform, open science, and equity. For policy-makers and institutions, the findings point to the importance of moving beyond the metricization of academic labor toward evaluative practices that account for methodological rigor, transparency, and social value. Specifically, research and tenureship evaluation should prioritize the societal relevance, methodological rigor, and ethical impact of scholarship rather than its performance on prestige-driven metrics or the venue it is published in; this shift recognizes that the true value of research lies not in how many people "like" it or where it appears in, but in its capacity to contribute to collective understanding.

For editors, the persistence of identity-based and structural bias necessitates a reformation from "Gatekeeper" to "Mentor" through the publication mentorship program, where researchers are required to work with the editorial team and its associates to bring the manuscript to the satisfactory quality; this practice reflects the true purpose of academia: to encourage an environment of inclusive and

 doi:10.20944/preprints202507.2644.v1

continual learning (Cross et al., 2019). This mentorship spirit recognizes that all scholars were once learners who benefited from the patience and guidance of their predecessors; extending the same generosity to future researchers aligns with the very nature of academia as a community devoted to teaching and learning [106]. To deny other scholars, emerging or established, such opportunities would undermine the institution's commitment to nurturing knowledge and scholarly growth.

Publishers, on the other hand, bear a complementary responsibility to ensure that the systems supporting editorial practices are equally inclusive and representative. This involves cultivating greater diversity among editorial boards, not only in terms of gender, geography, and career stage, but also in disciplinary perspectives and lived experiences [107]. If possible, publishers should make the paper freely accessible to public as much as possible, akin to the green or diamond open access model, as these routes minimize financial and structural barriers for both authors and readers, thereby dismantling inequities in who can access, produce, and benefit from scholarly work, ensuring that research serves both academic and broader audiences.

For the broader academic community, including researchers of all career stages and research students, the politicization of access debates highlights the urgency of systemic reform that addresses not only technical barriers but also the ethical justification of knowledge distribution. Reform must happen now. Those who conform to current norms without at least questioning them risk perpetuating systemic injustices. Academia's evaluative and publishing frameworks must be critically reimagined to center equity, transparency, and integrity. Disruptive forces like DORA and Sci-Hub have opened the door for change, but deeper reform is needed. LLM misuse, publishing elitism, and restricted access are not isolated incidents; They are symptoms of a system designed to prioritize institutional prestige over public good. If we fail to act, we could allow these structures to persist and calcify. To challenge them is not merely a professional obligation; it is a moral imperative.

While bibliometric methods provide robust tools for mapping scholarly discourse, this study is limited by coverage biases in bibliographic databases, potential underrepresentation of non-English literature, and the focus on published outputs rather than lived experiences of researchers. Future research could integrate qualitative analyses of author perspectives, comparative studies of Global North and Global South publishing systems, and evaluations of the real-world impact of reform initiatives such as DORA and open access mandates. Further investigation into the intersections of the three domains may yield deeper insights into how power, prestige, and access privilege co-evolve in scholarly communication.

The future scholarly culture must reflect inclusive values. Editorial systems should diversify not only in demographic representation but also in ideological openness. Metrics must be contextualized, not idolized. Access must be reconceptualized as a public right, not a privilege granted by publishers. Reflexivity must replace routinization, and collaboration must triumph over gatekeeping. The transformation will not be simple, but it is essential. A new academic ethos, one that reclaims scholarship for the many and not the few, is within reach, if we collectively strive for it. The worth of knowledge is not inherently in its metric, but in its capacity to include, to challenge, and to develop the society it lives in.

Author Contributions: Conceptualization, T.W.; methodology, T.W.; validation, T.W.; formal analysis, T.W.; resources, S.N.; data curation, T.W.; writing—original draft preparation, T.W., S.N.; writing—review and editing, T.W., S.N.; visualization, T.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset is available upon request.

Acknowledgments: The authors acknowledge the support of librarians at the University of Alberta for support and guidance in bibliographic search.

Conflicts of Interest: No conflict of interest to the best knowledge of the authors.

 doi:10.20944/preprints202507.2644.v1

Abbreviations

The following abbreviations are used in this manuscript:

JIF	Journal Impact Factor
DORA	Declaration on Research Assessment
Altmetrics	Alternative Metrics
APC	Article Processing Charges
WoS	Web of Science
MCA	Multiple Correspondence Analysis
DOAJ	Directory of Open Access Journals
LLM	Large Language Models

Appendix A. Removed and Consolidated Keywords

For the Gatekeeping and Editorial Bias topic,the following terms were excluded: "news production", "online journalism", "journalism", "news", "newspapers", "online-news", "online news", "social media", "communication", "innovation", "content analysis", "cultural evolution", "cultural transmission", "social learning", "clinical trials", "meta-analysis", "meta-analyses", "systematic review", "COVID-19", "epidemiology", "artificial intelligence", "trial registration", "clinical trial registration", "evidence-based medicine", "protocol", "depression", “systematic reviews”, and “spine”. This was done because these terms, while potentially frequent or popular in bibliographic records, reflect broader or unrelated topics and thus dilute the thematic precision of the analysis.

Moreover, synonymous or closely related terms were merged to reduce fragmentation and better capture coherent conceptual themes. For example: "publication", "publishing", "academic publishing", and "scholarly publishing" were treated as a single concept. "journal prestige" and "prestige" were also merged. “Academic journals”, “academic journal”, “Journal” and “journals” were also merged. “Editorial decisions” and “editorial decision” were also merged. This was necessary to prevent the dispersion of conceptually similar ideas into separate clusters, which could obscure the true thematic structure of the field.

For the Prestige-driven metrics and research evaluation topic, several terms were removed to improve thematic precision and eliminate noise from unrelated or overly broad research areas. The excluded terms include: "depression", "surgery", "consort", "internet", "child", "optimization", "deep learning", "serials", "data mining", "machine learning", "schizophrenia", "uncertainty", "randomized controlled trials", "systematic review", "meta-analysis", "simulation", "management", "evidence-based practice", "education", "china", "clinical trials", "evidence-based medicine", "systematic reviews", "review", "COVID-19", "SARS-CoV-2", "PRISMA", "artificial intelligence", "meta-analyses", "clinical trial registration", "HARKing", "clinical appraisal", "big data", "clinical practice guidelines", "observational study", "correlation", "biology", "neck pain", "qualitative research", "Iran", "World Wide Web", "UK", "Australia", "United Kingdom", "marketing", "history", "modeling", "communication", "MEDLINE", "pharmacoepidemiology", and "coronavirus".

In addition, several synonymous or closely related terms were merged, which are "randomized controlled trials" and "randomized controlled trial"; "human" and "humans"; "research evaluation" and "research assessment"; "journal impact factor", "impact factors", and "impact factor"; “citations” and “citation”; “h-index” and “h index”; “journals” and “journal”; “journal ranking” and “journal rankings”; “publications” and “publication”; “altmetrics” and “alternative metrics” and “altmetric”.

For the Barrier and equity issues in research accessibility topic, the removed terms include "Wikipedia", "cancer", "pharmacology", "Ethiopia", "Nigeria", "recruitment", "COVID-19", "periodicals", "assessment", "cognitive behavioral theory", "economics", "federated search", "metasearch", "electronic books", "usage statistics", "blogs", "evaluation", "serials", "chemistry", "portals", "case study", "information technology", "information retrieval", "Web 2.0", "reference services", "guidelines", "Dublin Core", "OAI-PMH", "health disparities", "health inequity", "climate change", "International Association for

 doi:10.20944/preprints202507.2644.v1

Vegetation Science (IAVS)", "health equity", "medical education", "artificial intelligence", "serials crisis", "vegetation classification", “systematic review”, and “meta-analysis”.

Furthermore, synonymous and conceptually overlapping terms were merged as follows: "article processing charge", "APC", "article processing charges", "article processing charge (APC)", "article processing charges (APC)"; "human" and "humans"; "open access", "open access journals", "open access publishing", "OA"; "journal impact factor", "impact factors", "impact factor"; "academic libraries", "university libraries", "libraries", "library", "digital library" and "digital libraries"; "scholarly publishing", "scholarly communication", "academic publishing", "publishing", "scientific publishing", "journal publishing"; "scientific journal", "journals", "scientific journals"; "research accessibility" and "accessibility"; "electronic journals", "electronic publishing", "e-journals"; "cost" and "price"; "predatory journals", "predatory journal", "predatory"; "Directory of Open Access Journals (DOAJ)" and "DOAJ".

References

1. Bedeian, A.G. Peer review and the social construction of knowledge in the management discipline. *Academy of Management Learning & Education* **2004**, *3*, 198–216. Publisher: Academy of Management, https://doi.org/10.5465/amle.2004.13500489.
2. Hyland, K. *Academic publishing: Issues and challenges in the construction of knowledge*; Oxford University Press: UK, 2016. OCLC: 1479866814.
3. Connell, R. *Southern theory: The global dynamics of knowledge in social science*, 1 ed.; Routledge, 2020. https://doi.org/10.4324/9781003117346.
4. Fricker, M. *Epistemic injustice: Power and the ethics of knowing*, 1 ed.; Oxford University PressOxford, 2007. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001.
5. Lawrence, P.A. The politics of publication. *Nature* **2003**, *422*, 259–261.
6. Brembs, B. Reliable novelty: New should not trump true. *PLOS Biology* **2019**, *17*, e3000117. https://doi.org/10.1371/journal.pbio.3000117.
7. Brembs, B.; Button, K.; Munafò, M. Deep impact: Unintended consequences of journal rank. *Frontiers in Human Neuroscience* **2013**, *7*. https://doi.org/10.3389/fnhum.2013.00291.
8. Primack, R.B.; Regan, T.J.; Devictor, V.; Zipf, L.; Godet, L.; Loyola, R.; Maas, B.; Pakeman, R.J.; Cumming, G.S.; Bates, A.E.; et al. Are scientific editors reliable gatekeepers of the publication process? *Biological Conservation* **2019**, *238*, 108232. https://doi.org/10.1016/j.biocon.2019.108232.
9. Tennant, J.P.; Waldner, F.; Jacques, D.C.; Masuzzo, P.; Collister, L.B.; Hartgerink, C.H.J. The academic, economic and societal impacts of open access: An evidence-based review. *F1000Research* **2016**, *5*, 632. https://doi.org/https://doi.org/10.12688/f1000research.8460.3.
10. Moher, D.; Bouter, L.; Kleinert, S.; Glasziou, P.; Sham, M.H.; Barbour, V.; Coriat, A.M.; Foeger, N.; Dirnagl, U. The hong kong principles for assessing researchers: Fostering research integrity. *PLOS Biology* **2020**, *18*, e3000737. Publisher: Public Library of Science (PLoS), https://doi.org/10.1371/journal.pbio.3000737.
11. Ross, J.S.; Gross, C.P.; Desai, M.M.; Hong, Y.; Grant, A.O.; Daniels, S.R.; Hachinski, V.C.; Gibbons, R.J.; Gardner, T.J.; Krumholz, H.M. Effect of blinded peer review on abstract acceptance. *JAMA* **2006**, *295*, 1675. Publisher: American Medical Association (AMA), https://doi.org/10.1001/jama.295.14.1675.
12. Helmer, M.; Schottdorf, M.; Neef, A.; Battaglia, D. Gender bias in scholarly peer review. *eLife* **2017**, *6*. Publisher: eLife Sciences Publications, Ltd, https://doi.org/10.7554/elife.21718.
13. Kidney International. Guide for authors.
14. Nature. Other types of submissions.
15. Teixeira Da Silva, J.A.; Daly, T. Inclusive editorial policies can reduce epistemic injustice in academic authorship. *CHEST* **2024**, *166*, e22–e23. Publisher: Elsevier BV, https://doi.org/10.1016/j.chest.2024.01.052.
16. Meneghini, R.; Mugnaini, R.; Packer, A.L. International versus national oriented Brazilian scientific journals. A scientometric analysis based on SciELO and JCR-ISI databases. *Scientometrics* **2006**, *69*, 529–538. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s11192-006-0168-z.
17. Neuberger, J.; Counsell, C. Impact factors: Uses and abuses. *European Journal of Gastroenterology & Hepatology* **2002**, *14*, 209–211. https://doi.org/10.1097/00042737-200203000-00001.
18. Casadevall, A.; Fang, F.C. Causes for the persistence of impact factor mania. *mBio* **2014**, *5*. Publisher: American Society for Microbiology, https://doi.org/10.1128/mbio.00064-14.
19. Hicks, D.; Wouters, P.; Waltman, L.; De Rijcke, S.; Rafols, I. Bibliometrics: The leiden manifesto for research metrics. *Nature* **2015**, *520*, 429–431. https://doi.org/10.1038/520429a.

 doi:10.20944/preprints202507.2644.v1

20. Biagioli, M.; Lippman, A., Eds. *Gaming the metrics: Misconduct and manipulation in academic research*; The MIT Press, 2020. https://doi.org/10.7551/mitpress/11087.001.0001.
21. Lawson, S.; Gray, J.; Mauri, M. Opening the black box of scholarly communication funding: A public data infrastructure for financial flows in academic publishing. *Open Library of Humanities* **2016**, *2*, e10. Publisher: Open Library of the Humanities, https://doi.org/10.16995/olh.72.
22. Segado-Boj, F.; Martín-Quevedo, J.; Prieto-Gutiérrez, J.J. Jumping over the paywall: Strategies and motivations for scholarly piracy and other alternatives. *Information Development* **2024**, *40*, 442–460. Publisher: SAGE Publications, https://doi.org/10.1177/02666669221144429.
23. Chan, L.; Costa, S. Participation in the global knowledge commons: Challenges and opportunities for research dissemination in developing countries. *New Library World* **2005**, *106*, 141–163. Publisher: Emerald, https://doi.org/10.1108/03074800510587354.
24. Wan, S. Which nationals use sci-hub mostly? *The Serials Librarian* **2023**, *83*, 228–232. Publisher: Informa UK Limited, https://doi.org/10.1080/0361526x.2023.2193613.
25. Rossello, G.; Martinelli, A. Breach of academic values and misconduct: The case of Sci-Hub. *Scientometrics* **2024**, *129*, 5227–5263. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s11192-024-05046-8.
26. Cagan, R. San Francisco declaration on research assessment. *Disease Models & Mechanisms* **2013**, p. dmm.012955. https://doi.org/10.1242/dmm.012955.
27. Declaration on Research Assessment [DORA]. San Francisco declaration on research assessment, 2012.
28. Priem, J.; Taraborelli, D.; Groth, P.; Neylon, C. Altmetrics: A manifesto. *DigitalCommons@University of Nebraska - Lincoln* **2011**. https://doi.org/https://digitalcommons.unl.edu/scholcom/185/.
29. Nazim, M.; Bhardwaj, R.K. Open access initiatives in European countries: Analysis of trends and policies. *Digital Library Perspectives* **2023**, *39*, 371–392. Publisher: Emerald, https://doi.org/10.1108/dlp-06-2022-0051.
30. Sugimoto, C.R.; Larivière, V. *Measuring research: What everyone needs to know*; What everyone needs to know, Oxford University Press: New York, 2018.
31. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.M. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research* **2021**, *133*, 285–296. Publisher: Elsevier BV, https://doi.org/10.1016/j.jbusres.2021.04.070.
32. Padmalochanan, P. Academics and the field of academic publishing: Challenges and approaches. *Publishing Research Quarterly* **2019**, *35*, 87–107. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s12109-018-09628-2.
33. Green, T. Is open access affordable? Why current models do not work and why we need internet-era transformation of scholarly communications. *Learned Publishing* **2019**, *32*, 13–25. Publisher: Wiley, https://doi.org/10.1002/leap.1219.
34. Buljung, B.; Bongiovanni, E.; Li, Y. *Navigating the research lifecycle for the modern researcher*, 2 ed.; Colorado Pressbooks Network, 2024.
35. Xu, L.; Grant, B. Doctoral publishing and academic identity work: Two cases. *Higher Education Research & Development* **2020**, *39*, 1502–1515. Publisher: Informa UK Limited, https://doi.org/10.1080/07294360.2020.1728522.
36. Binfield, P.; Rolnik, Z.; Brown, C.; Cole, K. Academic journal publishing. *The Serials Librarian* **2008**, *54*, 141–153. Publisher: Informa UK Limited, https://doi.org/10.1080/03615260801973992.
37. De Vries, D.R.; Marschall, E.A.; Stein, R.A. Exploring the peer review process: What is it, does it work, and can it be improved? *Fisheries* **2009**, *34*, 270–279. Publisher: Oxford University Press (OUP), https://doi.org/10.1577/1548-8446-34.6.270.
38. Day, A.; Peters, J. Quality indicators in academic publishing. *Library Review* **1994**, *43*, 4–72. Publisher: Emerald, https://doi.org/10.1108/00242539410068015.
39. Yuan, J.T.; Aires, D.J.; DaCunha, M.; Funk, K.; Habashi-Daniel, A.; Moore, S.A.; Heimes, A.; Sawaf, H.; Fraga, G.R. The h-index for associate and full professors of dermatology in the united states: An epidemiologic study of scholastic production. *Cutis* **2017**, *100*, 395–398.
40. Garfield, E. The history and meaning of the journal impact factor. *JAMA* **2006**, *295*, 90. Publisher: American Medical Association (AMA), https://doi.org/10.1001/jama.295.1.90.
41. N.K., S.; Mathew K., S.; Cherukodan, S. Impact of scholarly output on university ranking. *Global Knowledge, Memory and Communication* **2018**, *67*, 154–165. Publisher: Emerald, https://doi.org/10.1108/gkmc-11-2017-0087.

 doi:10.20944/preprints202507.2644.v1

42. Bradshaw, C.J.A.; Brook, B.W. How to rank journals. *PLOS ONE* **2016**, *11*, e0149852. Publisher: Public Library of Science (PLoS), https://doi.org/10.1371/journal.pone.0149852.
43. Havey, N.; Chang, M.J. Do journals have preferences? Insights from the journal of higher education. *Innovative Higher Education* **2022**, *47*, 915–926. https://doi.org/10.1007/s10755-022-09634-5.
44. Jussim, L.; Honeycutt, N. Bias in psychology: A critical, historical and empirical review. *Swiss Psychology Open* **2024**, *4*, 5. https://doi.org/10.5334/spo.77.
45. Emerson, G.B.; Warme, W.J.; Wolf, F.M.; Heckman, J.D.; Brand, R.A.; Leopold, S.S. Testing for the presence of positive-outcome bias in peer review: A randomized controlled trial. *Archives of Internal Medicine* **2010**, *170*. Publisher: American Medical Association (AMA), https://doi.org/10.1001/archinternmed.2010.406.
46. Ioannidis, J.P.A. Why most published research findings are false. *PLoS Medicine* **2005**, *2*, e124. https://doi.org/10.1371/journal.pmed.0020124.
47. Koch-Weser, D.; Yankauer, A. The authorship and fate of international health papers submitted to the American Journal of Public Health in 1989. *American Journal of Public Health* **1993**, *83*, 1618–1620. Publisher: American Public Health Association, https://doi.org/10.2105/ajph.83.11.1618.
48. Weller, A.C. *Editorial peer review: Its strengths and weaknesses*; ASIST monograph series, Information Today: Medford, N.J, 2001.
49. Gibson, J.L. Against desk rejects! *PS: Political Science & Politics* **2021**, *54*, 673–675. Publisher: Cambridge University Press (CUP), https://doi.org/10.1017/s1049096521000524.
50. Rath, M.; Wang, P. Open peer review in the era of open science: a pilot study of researchers' perceptions. In Proceedings of the 2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Toronto, ON, Canada, jun 2017; pp. 1–2. https://doi.org/10.1109/jcdl.2017.7991608.
51. Ross-Hellauer, T. What is open peer review? A systematic review. *F1000Research* **2017**, *6*, 588. Publisher: F1000 Research Ltd, https://doi.org/10.12688/f1000research.11369.2.
52. Hunter, J. Post-publication peer review: Opening up scientific conversation. *Frontiers in Computational Neuroscience* **2012**, *6*. Publisher: Frontiers Media SA, https://doi.org/10.3389/fncom.2012.00063.
53. Universities UK. The concordat to support research integrity, 2025.
54. Wright, J.; Avouris, A.; Frost, M.; Hoffmann, S. Supporting academic freedom as a human right: Challenges and solutions in academic publishing. *The International Journal of Human Rights* **2022**, *26*, 1741–1760. Publisher: Informa UK Limited, https://doi.org/10.1080/13642987.2022.2088520.
55. McKiernan, E.C.; Schimanski, L.A.; Muñoz Nieves, C.; Matthias, L.; Niles, M.T.; Alperin, J.P. Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations. *eLife* **2019**, *8*, e47338. https://doi.org/10.7554/eLife.47338.
56. Salinas, S.; Munch, S.B. Where should I send it? Optimizing the submission decision process. *PLOS ONE* **2015**, *10*, e0115451. Publisher: Public Library of Science (PLoS), https://doi.org/10.1371/journal.pone.0115451.
57. Baccini, A.; De Nicolao, G.; Petrovich, E. Citation gaming induced by bibliometric evaluation: A country-level comparative analysis. *PLOS ONE* **2019**, *14*, e0221212. Publisher: Public Library of Science (PLoS), https://doi.org/10.1371/journal.pone.0221212.
58. D'Souza, B.; Kulkarni, S.; Cerejo, C. Authors' perspectives on academic publishing: initial observations from a large-scale global survey. *Science Editing* **2018**, *5*, 39–43. Publisher: Korean Council of Science Editors, https://doi.org/10.6087/kcse.116.
59. Wang, J.L.; Li, X.; Fan, J.R.; Yan, J.P.; Gong, Z.M.; Zhao, Y.; Wang, D.M.; Ma, L.; Ma, N.; Guo, D.M.; et al. Integrity of the editing and publishing process is the basis for improving an academic journal's Impact Factor. *World Journal of Gastroenterology* **2022**, *28*, 6168–6202. Publisher: Baishideng Publishing Group Inc., https://doi.org/10.3748/wjg.v28.i43.6168.
60. Ioannidis, J.P.A.; Maniadis, Z. Quantitative research assessment: Using metrics against gamed metrics. *Internal and Emergency Medicine* **2024**, *19*, 39–47. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s11739-023-03447-w.
61. Plevris, V. From integrity to inflation: Ethical and unethical citation practices in academic publishing. *Journal of Academic Ethics* **2025**. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s10805-025-09631-1.
62. Meho, L.I. Gaming the metrics? Bibliometric anomalies and the integrity crisis in global university rankings, 2025. Version Number: 4, https://doi.org/10.48550/ARXIV.2505.06448.
63. Rafols, I.; Leydesdorff, L.; O'Hare, A.; Nightingale, P.; Stirling, A. How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy* **2012**, *41*, 1262–1282. Publisher: Elsevier BV, https://doi.org/10.1016/j.respol.2012.03.015.

 doi:10.20944/preprints202507.2644.v1

64. Edwards, M.A.; Roy, S. Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science* **2017**, *34*, 51–61. Publisher: Mary Ann Liebert Inc, https://doi.org/10.1089/ees.2016.0223.
65. Leckert, M. (E-) valuative metrics as a contested field: A comparative analysis of the altmetrics- and the leiden manifesto. *Scientometrics* **2021**, *126*, 9869–9903. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s11192-021-04039-1.
66. Rushforth, A.; Hammarfelt, B. The rise of responsible metrics as a professional reform movement: A collective action frames account. *Quantitative Science Studies* **2023**, *4*, 879–897. Publisher: MIT Press, https://doi.org/10.1162/qss_a_00280.
67. Boury, H.; Albert, M.; Chen, R.H.C.; Chow, J.C.L.; DaCosta, R.; Hoffman, M.M.; Keshavarz, B.; Kontos, P.; McAndrews, M.P.; Protze, S.; et al. Exploring the merits of research performance measures that comply with the San Francisco declaration on research assessment and strategies to overcome barriers of adoption: Qualitative interviews with administrators and researchers. *Health Research Policy and Systems* **2023**, *21*. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1186/s12961-023-01001-w.
68. Dewidar, O.; Elmestekawy, N.; Welch, V. Improving equity, diversity, and inclusion in academia. *Research Integrity and Peer Review* **2022**, *7*, 4. https://doi.org/10.1186/s41073-022-00123-z.
69. Larivière, V.; Haustein, S.; Mongeon, P. The oligopoly of academic publishers in the digital era. *PLOS ONE* **2015**, *10*, e0127502. https://doi.org/10.1371/journal.pone.0127502.
70. van Bellen, S.; Alperin, J.P.; Larivière, V. The oligopoly of academic publishers persists in exclusive database, 2024. Version Number: 1, https://doi.org/10.48550/ARXIV.2406.17893.
71. Björk, B.C. Open access to scientific articles: A review of benefits and challenges. *Internal and Emergency Medicine* **2017**, *12*, 247–253. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s11739-017-1603-2.
72. Harwell, E.N. Academic journal publishers antitrust litigation, 2024.
73. Segado-Boj, F.; Prieto-Gutiérrez, J.; Martín-Quevedo, J. Attitudes, willingness, and resources to cover article publishing charges: The influence of age, position, income level country, discipline and open access habits. *Learned Publishing* **2022**, *35*, 489–498. Publisher: Wiley, https://doi.org/10.1002/leap.1455.
74. Solomon, D.J.; Björk, B. Publication fees in open access publishing: Sources of funding and factors influencing choice of journal. *Journal of the American Society for Information Science and Technology* **2012**, *63*, 98–107. Publisher: Wiley, https://doi.org/10.1002/asi.21660.
75. Klebel, T.; Ross-Hellauer, T. The APC-barrier and its effect on stratification in open access publishing. *Quantitative Science Studies* **2023**, *4*, 22–43. Publisher: MIT Press, https://doi.org/10.1162/qss_a_00245.
76. Budzinski, O.; Grebel, T.; Wolling, J.; Zhang, X. Drivers of article processing charges in open access. *Scientometrics* **2020**, *124*, 2185–2206. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s11192-020-03578-3.
77. Cantrell, M.H.; Swanson, J.A. Funding sources for open access article processing charges in the social sciences, arts, and humanities in the united states. *Publications* **2020**, *8*, 12. Publisher: MDPI AG, https://doi.org/10.3390/publications8010012.
78. Ellingson, M.K.; Shi, X.; Skydel, J.J.; Nyhan, K.; Lehman, R.; Ross, J.S.; Wallach, J.D. Publishing at any cost: A cross-sectional study of the amount that medical researchers spend on open access publishing each year. *BMJ Open* **2021**, *11*, e047107. Publisher: BMJ, https://doi.org/10.1136/bmjopen-2020-047107.
79. Zhang, L.; Watson, E.M. Measuring the impact of gold and green open access. *The Journal of Academic Librarianship* **2017**, *43*, 337–345. Publisher: Elsevier BV, https://doi.org/10.1016/j.acalib.2017.06.004.
80. Gouzi, F.; Tasovac, T. Gold, green, diamond: What you should know about open access publishing models authors, 2025.
81. Normand, S. Is diamond open access the future of open access? *The iJournal: Student Journal of the Faculty of Information*, **2018**, *3*, 1–7.
82. Himmelstein, D.S.; Romero, A.R.; Levernier, J.G.; Munro, T.A.; McLaughlin, S.R.; Greshake Tzovaras, B.; Greene, C.S. Sci-Hub provides access to nearly all scholarly literature. *eLife* **2018**, *7*, e32822. https://doi.org/10.7554/eLife.32822.
83. Greshake, B. Looking into Pandora's box: The content of Sci-hub and its usage. *F1000Research* **2017**, *6*, 541. https://doi.org/10.12688/f1000research.11366.1.
84. Karaganis, J., Ed. *Shadow libraries: Access to knowledge in global higher education*; International Development Research Centre, The MIT Press: Cambridge, 2018.

 doi:10.20944/preprints202507.2644.v1

85. Bohannon, J. Who's downloading pirated papers? Everyone. *Science* **2016**, *352*, 508–512. Artwork Size: 686192584 bytes Pages: 686192584 bytes.
86. Sugiharto, S. De-westernizing hegemonic knowledge in global academic publishing: Toward a politics of locality. *Journal of Multicultural Discourses* **2021**, *16*, 321–333. Publisher: Informa UK Limited, https://doi.org/10.1080/17447143.2021.2017442.
87. Schiltz, M. Science without publication paywalls: cOAlition S for the realisation of full and immediate open access. *Frontiers in Neuroscience* **2018**, *12*, 656. https://doi.org/10.3389/fnins.2018.00656.
88. Grötschel, M. Project DEAL: Plans, challenges, results, 2019.
89. Aria, M.; Cuccurullo, C. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics* **2017**, *11*, 959–975. https://doi.org/https://doi.org/10.1016/j.joi.2017.08.007.
90. Johs, H. *Multiple correspondence analysis for the social sciences*, 1 ed.; Routledge: Abingdon, Oxon; New York, NY: Routledge, 2018., 2018. https://doi.org/10.4324/9781315516257.
91. Cobo, M.; López-Herrera, A.; Herrera-Viedma, E.; Herrera, F. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *Journal of Informetrics* **2011**, *5*, 146–166. Publisher: Elsevier BV, https://doi.org/10.1016/j.joi.2010.10.002.
92. Lorenzo Lledó, G.; Scagliarini Galiano, C. Revisión bibliométrica sobre la realidad aumentada en Educación. *Revista General de Información y Documentación* **2018**, *28*. Publisher: Universidad Complutense de Madrid (UCM), https://doi.org/10.5209/rgid.60805.
93. Ossom Williamson, P.; Minter, C.I.J. Exploring PubMed as a reliable resource for scholarly communications services. *Journal of the Medical Library Association: JMLA* **2019**, *107*, 16–29. https://doi.org/10.5195/jmla.2019.433.
94. Moher, D.; Naudet, F.; Cristea, I.A.; Miedema, F.; Ioannidis, J.P.A.; Goodman, S.N. Assessing scientists for hiring, promotion, and tenure. *PLOS Biology* **2018**, *16*, e2004089. Publisher: Public Library of Science (PLoS), https://doi.org/10.1371/journal.pbio.2004089.
95. Pulverer, B. Impact fact-or fiction? *The EMBO Journal* **2013**, *32*, 1651–1652. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1038/emboj.2013.126.
96. Fire, M.; Guestrin, C. Over-optimization of academic publishing metrics: Observing Goodhart's law in action. *GigaScience* **2019**, *8*. Publisher: Oxford University Press (OUP), https://doi.org/10.1093/gigascience/giz053.
97. Sinclair, B.J. Letting ChatGPT do your science is fraudulent (and a bad idea), but AI-generated text can enhance inclusiveness in publishing. *Current Research in Insect Science* **2023**, *3*, 100057. Publisher: Elsevier BV, https://doi.org/10.1016/j.cris.2023.100057.
98. Seglen, P.O. Why the impact factor of journals should not be used for evaluating research. *BMJ* **1997**, *314*, 497–497. Publisher: BMJ, https://doi.org/10.1136/bmj.314.7079.497.
99. Van Raan, A.F.J. Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics* **2005**, *62*, 133–143. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1007/s11192-005-0008-6.
100. Ross-Hellauer, T.; Görögh, E. Guidelines for open peer review implementation. *Research Integrity and Peer Review* **2019**, *4*. Publisher: Springer Science and Business Media LLC, https://doi.org/10.1186/s41073-019-0063-9.
101. Chan, L.; Kirsop, B.; Arunachalam, S. Towards open and equitable access to research and knowledge for development. *PLoS Medicine* **2011**, *8*, e1001016. Publisher: Public Library of Science (PLoS), https://doi.org/10.1371/journal.pmed.1001016.
102. Mathew, R.P.; Patel, V.; Low, G. Predatory journals- the power of the predator versus the integrity of the honest. *Current Problems in Diagnostic Radiology* **2022**, *51*, 740–746. Publisher: Elsevier BV, https://doi.org/10.1067/j.cpradiol.2021.07.005.
103. Willinsky, J. *The access principle: The case for open access to research and scholarship*; Digital libraries and electronic publishing, MIT Press: Cambridge, Mass., 2006.
104. Posada, A.; Chen, G. Inequality in knowledge production: The integration of academic infrastructure by big publishers. In Proceedings of the 22nd International Conference on Electronic Publishing. OpenEdition Press, jun 2018. https://doi.org/10.4000/proceedings.elpub.2018.30.
105. Fyfe, A.; Coate, K.; Curry, S.; Lawson, S.; Moxham, N.; Røstvik, C.M. Untangling academic publishing: A history of the relationship between commercial interests, academic prestige and the circulation of research, 2017.
106. Bowen, W.M.; Schwartz, M. *The chief purpose of universities: Academic discourse and the diversity of ideas*; MSL Academic Endeavors: Cleveland, Ohio, 2010. OCLC: 740985550.

 doi:10.20944/preprints202507.2644.v1

107. Nahai, F. Diversity on journal editorial boards: Why it is important. *Aesthetic Surgery Journal* **2021**, *41*, 635–637. Publisher: Oxford University Press (OUP), https://doi.org/10.1093/asj/sjab078.