Research Evaluation, 2026, **35**, rvag004
https://doi.org/10.1093/reseval/rvag004
Article

OXFORD

Artificial intelligence in academic practices and policy discourses across 'Big 5' publishers

Gergely Ferenc Lendvai[1,*] and Aczél Petra[2]

[1]Science Strategy Office, Ludovika University of Public Service, Ludovika tér 2, Hungary
[2]Management Campus, Széchenyi István University, Egyetem tér 1, Hungary

*Corresponding author. Science Strategy Office, Ludovika University of Public Service, Ludovika tér 2, Budapest, 1083, Hungary. E-mail: Lendvai.gergely.ferenc@uni-nke.hu

Abstract

The present study investigates how the five largest academic publishers (Elsevier, Springer, Wiley, Taylor & Francis, and SAGE) are responding to the epistemic and procedural challenges posed by generative AI through formal policy frameworks. Situated within ongoing debates about the boundaries of authorship and the governance of AI-generated content, our research aims to critically assess the discursive and regulatory contours of publishers' authorship guidelines (PGs). We employed a multi-method design that combines qualitative coding, semantic network analysis, and comparative matrix visualization to examine the official policy texts collected from each publisher's website. Findings reveal a foundational consensus across all five publishers in prohibiting AI systems from being credited as authors and in mandating disclosure of AI usage. However, beyond this shared baseline, marked divergences emerge in the scope, specificity, and normative framing of AI policies. Co-occurrence and semantic analyses underline the centrality of 'authorship', 'ethics', and 'accountability' in AI discourse. Structural similarity measures further reveal alignment among Wiley, Elsevier, and Taylor & Francis, with Springer as a clear outlier. Our results point to an unsettled regulatory landscape where policies serve not only as instruments of governance but also as performative assertions of institutional identity and legitimacy. Consequently, the fragmented field of PG highlights the need for harmonized, inclusive, and enforceable frameworks that recognize both the potential and risks of AI in scholarly communication.

Keywords: AI authorship; academic publishing; policy analysis; academic writing and AI; semantic network analysis.

1. Introduction

In 1967, Roland Barthes famously declared 'the death of the author'. A paradigm-shifting concept, Barthes challenged the idea that meaning resides in the intentions of a singular, sovereign creator, namely, the author of a respective text. In a different context, Barthes' thought echo through today's writing ecosystem with the rise of generative artificial intelligence (commonly referred to as AI), however, with a vastly new interpretation, or much rather, provocation that acquires an unorthodox materiality: what happens when the 'author' is no longer human at all?

It is undeniable that popular AI-based tools, such as ChatGPT, generate text without consciousness or intention, yet their outputs may, indeed, circulate as coherent, persuasive academic prose. This collapse (or transformation) of authorship into computation reanimates Barthes' critique not only as a philosophical and ethical dilemma but now as a policy problem, too. Authorship in academic publishing has been the forefront of research ethics for decades, if not centuries. Questions on who may become an author, authorship positions and preferences, group and corresponding authorships, and unethical authorship attributions look back at a plethora of research materials (Bennett and Taylor 2003; Pontille et al. 2004; Anderson and Boden 2008; Marušić et al. 2011; Hosseini and Gordijn 2020; Singhal and Kalra 2021; Vasilevsky et al. 2021; Hosseini et al. 2025). Nonetheless, today's publishers debate not solely about whether authorship is stable and ethically adequate but about an even more quintessential question that has rarely been argued over before; who (or *what*) can be called an author (see Dwivedi et al. 2023 and cf Tennant et al. 2019). This is a 'generative' turn in both research ethics and evaluation and thus invites, or more adequately, forces the academic publishing industry to confront, explicitly, what has long been philosophically unsettled: what constitutes authorship, and who holds the power to define it.

It is almost platitude-like to mention that the emergence of AI has introduced profound epistemic, ethical, and procedural challenges to academic publishing. These technologies are not merely computational assistants as they are capable of simulating language use in ways that approximate scholarly voice, structure, and even argumentation style (Khalifa and Albadawy 2024; Cheng et al. 2025; Lendvai 2025). Recontextualising the issue from a practical standpoint, for instance, what has taken an excellent research assistant or scholar to outline weeks or months to write and research now only takes a few seconds or at most, minutes in the case of more complex analyses with the power of generative AI software. As such, these tools unsettle long-standing assumptions about what it means to author a text, to be accountable for intellectual content, and to participate in the production of knowledge.

Academic publishers, sometimes regarded as (potential) gatekeepers in the scholarly communication ecosystem (Coser 1975; Primack et al. 2019; Tutuncu 2024), have begun to respond to this disruption through the development of formal AI authorship policies. Historically, these policies were benchmark documents in *Academia*. They dictated how a paper should be prepared, how authors should be accredited, and outlined essential information about unethical and

© The Author(s) 2026. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

prohibited research practices. Today, however, most academic publishing policies were enriched with a new perspective: the use of AI for writing (for more on this see Gendron et al. 2022; Lendvai 2025). Typically, these novel stipulations aim to delineate acceptable and unacceptable uses of generative AI in the research and publication process. However, the scope, specificity, and regulatory ambition of these documents vary remarkably across publication houses with some publishers articulating detailed, multi-layered governance frameworks, while others offer brief, generic statements that gesture toward ethical responsibility without substantive guidance (see Lin 2024; Bhavsar et al. 2025). This uneven policy terrain (Lin 2024) generates confusion among authors, editors, and reviewers alike, and also creates a 'Catch-22-esque' dilemma where overregulation essentially results in less material and slowed down publication process, while underregulation open the (flood) gates of research generated and fabricated by AI. Given the reproducibility problem, often referred to as 'replication crisis', mentioned by a multitude of scholars (Schooler 2014; Anvari and Lakens 2018; Hillary and Medaglia 2020; Mede et al. 2021; Korbmacher et al. 2023), most prominently by Ioannidis (2005) and more recently, Szabo (2025), AI policies not only have to prevent but also assess potential falsified information in papers, 'AI hallucinations', and potential plagiarism while ensuring that their reputation remains undamaged to attract impactful research being published.

The significance of how academic publishers react and reflect to AI as a disruptive technology in Academia cannot be overstated. Since AI tools are no longer confined to experimental research or technical domains and are being used to draft, refine, and sometimes even generate entire scholarly articles across disciplines, the lines between human and non-human authorship are blurred which raises both philosophical and procedural dilemmas that traditional publishing standards were and are not equipped to resolve (Bin-Nashwan et al. 2023; Dwivedi et al. 2023; Islam and Islam 2024). Therefore, in the present research we aim to undertake the task to analyze these developments as a call for a critical reexamination of the foundational categories of authorship, originality, and intellectual labor upon which academic publishing has historically relied.

The choice of academic publishers and their policies are not accidental. Despite growing public debate and institutional concerns raised both internally and externally, academic research on how publishers are shaping their AI governance policies remains limited. Though research grows exponentially with foundational works on the potential dangers of AI and academic publishing, the issue of AI disclosure in papers, the role of AI guidelines, policy frameworks, and the handling of chatbots in the publication process (Lund and Naheem 2024; Perkins and Roe 2023; Ganjavi et al. 2024; Gendron et al. 2022; Lin 2024; Bhavsar et al. 2025; Resnik and Hosseini 2025), the literature still remains rather scarce with existing discussions mostly focusing on conceptual frameworks (see Gil de Zúñiga et al. (2024) for a foundational perspective in this regard), general ethical concerns, technical considerations, specific aspects of AI usage in writing, or speculative legal debates, often ignoring the specific textual and structural forms through which policies are being articulated. Moreover, while some fields have explored AI's implications for students as well as non-native English speakers, the development of scholarly writing style, its role in the peer review process, or its consideration in the context of research integrity (Leung et al. 2023; Bauchner and Rivara 2024; Cummings et al. 2024; Mehta et al. 2024; Balalle and Pannilage 2025; Chen and Gong 2025; Kim et al. 2025; Parker et al. 2025), there has been little systematic analysis of how publishers themselves are responding—both discursively and procedurally—to the rise and potentially future dominance of generative AI (Ganjavi et al. 2024; Kolbinger et al. 2024; Lund and Naheem 2024). Subsequently, the lack of comparative, empirically grounded research on policy content, thematic priorities, and structural similarities across publishers represents a significant gap in the literature.

One of the most important articles on AI policies by publishers which also happens to be a key source of inspiration for this research emerges from Goyanes et al.'s (2025) paper who examined how leading Q1 journals across several social science disciplines are adapting their author guidelines to the emergence of generative AI. The authors conducted a thorough analysis of all guidelines and found that while most journals emphasize transparency, there is little to no reference to data privacy issues and references to AI use in literature reviews, for instance, are almost completely missing. The present study seeks to add to Goyanes et al.'s (2025) findings and address the aforementioned gap by examining the AI authorship policies and guidelines (hereinafter referred to as 'PG's) of the five largest academic publishers: Elsevier, Springer, Wiley, Taylor & Francis, and SAGE. These publishers are often called the 'Big 5' given their prominence and unparalleled success in academic publishing both in terms of commercial profitability and prestige (Larivière et al. 2015; Butler et al. 2023). Furthermore, these publishers represent not only a substantial portion of global academic output but also serve as influential standard-setters for editorial and authorship practices (Larivière et al. 2015; Butler et al. 2023). Consequently, analyzing the content, structure, and framing of their AI-related authorship policies serve as an excellent way to map the emerging contours of regulatory discourse in the field of academic publishing, especially with regard to the most prestigious publishers. Rather than approaching these policies as purely legal or administrative documents, the present study treats them as discursive artifacts, as in texts that reveal institutional priorities, potential 'anxieties', and epistemic boundaries. Our key objective is to ask how policies frame the relationship between human and non-human authorship, how they distribute responsibility and trust, and how they invoke ethical and procedural standards to maintain institutional legitimacy. While authorship forms the conceptual anchor of this study, our analysis also extends to editorial and peer review contexts, since AI governance in publishing increasingly shapes not only questions of attribution but also the evaluation and certification of scholarly work. In other words, the goal of the present article is not only to 'catalog' *what* policies say, but instead, to understand *how* they say it, and what that says about the publishing industry's broader response to AI. In doing so, our paper aims to move beyond surface-level comparison and offers a textured view of how AI is being conceptualized, managed, and problematized in scholarly publishing. Crucially, as a premise to our analysis, our study does not assume that more detailed policies are inherently better. Instead, it explores what kinds of values are being prioritized, what forms of responsibility are being imagined, and which aspects of the publication process remain unregulated or undertheorized.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

In order to analyze the PGs of the Big 5 publishers, we propose the following research questions (RQs):

- RQ1: To what extent do the major academic publishers, in our case, the Big 5 publishers, converge or diverge in their policy responses to the use of generative AI in scholarly publishing?
- RQ2: How are core themes such as authorship, accountability, and AI tool usage framed across different publisher policies?
- RQ3: What underlying structural or conceptual similarities and differences exist among publishers' AI policies?

Through the assessment of these questions, our study's objective is to contribute to an informed, critical, and empirically grounded understanding of the evolving landscape of AI governance in academic publishing. Additionally, we also aim to offer a snapshot of where policy stands today as well as a framework for thinking about where it might, or should, go next. Lastly, beyond filling a critical gap in publishing and policy studies, the article also aims to contribute to the field of research evaluation. Policies that govern AI in authorship, editorial work, and peer review directly affect how scholarly contributions are attributed, how integrity is safeguarded, and how research quality is judged, therefore, via comparing the 'titans' in publication, the Big 5 publishers, the study provides evidence of how fragmented AI governance can shape the conditions under which research outputs are evaluated and recognized across the academic system. Adding to this line of argumentation, if we accept the notion that research evaluation—at least partly—depends on stable and transparent attribution systems, it becomes essential to investigate how AI policies vary substantially in their precision and scope, and how equipped they are to make judgments about authorial responsibility. Furthermore, it is equally important to also recognize that publisher policies do not operate in isolation. On the contrary, they are active agents who shape the normative environment in which evaluators, reviewers, and institutions assign credit and legitimacy. Thus, the study's objective is also to speak directly to the infrastructures of evaluation, where policy choices at the publisher level cascade into how research performance is assessed and rewarded.

2. Materials and methods

In the present study, we employed a multi-method textual and computational analysis to examine how the five leading academic publishers articulate their policies regarding AI in the context of scholarly publishing. The methodology used integrates qualitative coding, semantic network modeling, and comparative matrix visualization in order to comprehensively uncover thematic structures, discursive emphases, and inter-publisher differences in framing AI's role and acceptable usage in editorial, authorial, and peer review contexts.

2.1. Data sources and preparation

We collected the policy data using the guidelines presented on each publisher's website on 1 June 2025 and then rechecked whether changes have been made on 22 June 2025, 29 June 2025, and lastly on 2 July 2025. This procedural consideration was instrumental given that most publishers omitted signaling a date of publication for their PGs. We found no differences during our checks. Official AI policy texts were retrieved from the websites of Elsevier, Springer Nature, Wiley, Taylor & Francis, and SAGE (the links can be found in Appendix 1). We only searched for AI policies and outlines found on a singular page on the respective publishers' website and did not check whether different journals operated by these publishers have different AI policies. The policies and guidelines included policy pages, author instructions, editorial guidance, and FAQ sections concerning the use of generative AI tools such as ChatGPT, Bard, and similar technologies in the research and publication lifecycle. As these materials were often embedded in layered webpage structures or collapsible panels, each policy was manually copied and pasted into a text file to preserve sectioning and header formatting. A total of five Word files were created—one for each publisher. Having identified the five PGs, the manually exported files were converted into plain text (.txt) format using UTF-8 encoding to ensure compatibility with downstream text analysis tools. No content alterations were made during the conversion process. As the data consists of publicly accessible policy statements, no ethical approval was required for the examinations.

2.2. Qualitative Content-Thematic Analysis (QCTA)

Since one of our primary goals was to investigate the key themes present in the AI policies, a qualitative content analysis was performed (similarly to Perkins and Roe's (2023) research—though without AI-assistance). Each policy text was parsed into smaller units, typically sentences or semantically discrete clauses, and analyzed using inductive thematic coding. The coding was conducted manually and interactively in R. To group similarities and differences, nine salient code categories were identified through iterative reading and memoing (Table 1).

Each sentence or segment was assigned to one individual or if applicable multiple codes to account for conceptual overlap. Overlapping codes were reviewed for redundancy and interpretive coherence. The aforementioned samples were included to illustrate each code within each publisher's document though it is to be mentioned that they may be altered to demonstrate them as example rather than exact excerpts. Both authors independently annotated segments and then reviewed results collaboratively.

2.3. Descriptive analysis and lexical statistics

From the coded dataset, three descriptive metrics were computed per publisher.

First, total word counts, and unique word counts were calculated after removing punctuation and stop words. Type-token ratios were also estimated to assess lexical diversity. Second, code density was calculated as the number of code assignments per 100 words of policy text. We used 100 words as a benchmark as this offered a normalized measure of how 'policy-heavy' or 'regulation-intensive' each publisher's stance was. Third, thematic breadth was measured by counting the number of unique codes present in each publisher's policy, providing insight into the scope of regulatory concerns addressed. Text preprocessing was conducted using the *tidytext*, *dplyr*, and *stringr* packages in R. All textual segments were lowercased, stripped of punctuation, and tokenized. Data cleaning ensured that analytical metrics were not inflated by repeated boilerplate language or document metadata.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

Table 1. The nine salient code categories.

Theme	Description	Example sentences and expressions based on the five investigated PGs
Authorship	Rules about whether AI can be listed as an author	*AI tools cannot be listed as authors …*
Disclosure	Whether use of AI must be disclosed and how	*Authors must disclose any AI use …*
Permissible use	Specific activities allowed (e.g. language editing, image generation)	*Use of AI for grammar correction is allowed*
Prohibited use	Forbidden uses (e.g. idea generation, data analysis)	*Do not input confidential data into AI …*
Peer review guidance	Rules for reviewers using AI tools	*Reviewers must not use AI without permission*
Editorial use	Guidance for editors, peer review automation	*Editorial decisions may not be automated …*
Accountability	Who is responsible for AI-generated content	*Authors are accountable for AI usage*
Ethical framing	Underlying values (e.g. transparency, integrity, trust). This category also includes potential intellectual property issues.	*In the spirit of academic honesty …*
Regulatory specificities and general statements (together as: specificity)	Degree of detail in policy (vague guidance vs. strict definitions), potential risks, and general statements	*AI use is generally discouraged vs. must …; Some issues regarding the usage of AI includes: …; [XY publisher] welcomes the development of AI tools …*

2.4. Semantic network analysis (SNA) and similarity analysis

To understand the latent structure of policy discourse, a semantic network was constructed from the combined corpus of all five AI policies. SNAs are frequently used to review policy papers and guidelines, especially their frames, themes, and trends they might show (Shim et al., 2015; Park et al., 2019; Danowski et al., 2023), therefore, we included this analysis to be applied to publication policies which have not been done yet. After tokenization and cleaning, lemmatization was performed to retain linguistically valid word forms (e.g. 'generative', 'generate', and 'generation' were unified as 'generate'). This method was preferred overstemming to preserve semantic interpretability in the network outputs. Word co-occurrences were computed within a sliding window of 5 words to capture syntagmatic associations while limiting spurious connections (see Nanni et al. 2013). Only word pairs that occurred together at least three times were retained. The resulting co-occurrence matrix was transformed into a graph object with nodes representing lemmatized terms and edges representing weighted co-occurrences.

Several network-level and node-level metrics were computed. This included degree centrality, betweenness centrality, and eigenvector centrality, metrics that are all used to assess the importance and influence of terms within the network (Barabasi 2016). Community detection was also performed to reveal clusters of semantically related concepts. In the network, node sizes were scaled by degree and color-coded by community.

Additionally, pairwise Jaccard similarity coefficients were calculated using binary code presence to estimate thematic overlap between publishers. These were plotted as a network.

2.5. Tools and reproducibility

All data processing and analysis were performed using R. Scripts were modularized and are fully reproducible. All datasets, including annotated codes, co-occurrence matrices, network metrics, and matrix outputs, are available upon request. Visualization settings were standardized for interpretability and reproducibility.

3. Results

The QCTA performed both differences and similarities in foci. The distribution of AI policy themes across the Big 5 presented distinct regulatory priorities and communicative strategies. Wiley stood out as the most comprehensive and verbose, with a total word count of 6,097 and the highest absolute frequencies in Authorship (35), Accountability (23), and Ethics (25). Elsevier, though much shorter at 2,288 words, also placed heavy emphasis on Authorship (25), Editorial Use (12), and Peer Review Guidance (13), suggesting a technically oriented stance. In contrast, Springer and Taylor & Francis offer more compact and selective policies, with 693 and 1,206 words, respectively, and fewer thematic mentions overall. Though Taylor & Francis still scores relatively high on Ethics (8) and Editorial Use (10). SAGE adopts a balanced position, with modest policy length (747 words) and mid-range coverage across most themes, including Disclosure (9), Authorship (9), and Editorial Use (8), indicating moderate institutional attention. Type-token ratio (TTR) values also help differentiate policy styles: Springer (0.553) and Taylor & Francis (0.453) display higher lexical diversity, suggesting concise but varied language, while Wiley's TTR (0.313) reflects repetitiveness due to its length and thematic focus. SAGE's TTR (0.446) and Elsevier's (0.269) reflect this balance between density and thematic recurrence. Thematically, Authorship is addressed by all five publishers, with Wiley and Elsevier dominating as mentioned above. Accountability is likewise widespread, especially in Wiley (23) and Taylor & Francis (9), reinforcing the idea of human oversight. Disclosure is present in all policies but emphasized most in Wiley (11) and SAGE (9) pointing to a shared expectation for transparency. Peer Review Guidance is inconsistently covered, prominent in Elsevier (13) and SAGE (7), but not overly accentuated in Springer (2), implying uneven regulation of reviewer behavior. Permissible Use is clearly defined in Wiley (11) and Springer (3); however, Prohibited Use remains low across the board, highest only in Wiley (7) where a plethora of examples are mentioned. The lack of detailed description may also suggest that there is a lack of consensus on prohibited use of AI and publishers may opt not to overregulate this practice to evade creating potential loopholes. Compared to peer reviewing, Editorial Use is much more

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

emphasized, particularly in Elsevier (12), Taylor & Francis (10), and SAGE (8), signaling emerging concern about AI use in editorial workflows. Ethical Framing, a normative layer often implied rather than explicitly coded, is especially dense in Wiley (25), while Springer and SAGE are more moderate at 4 mentions each. Finally, the Specificity code, measuring regulatory precision, is highest by far in Wiley (29), followed by Taylor & Francis (5), indicating that only some publishers go beyond general statements to offer concrete instructions. This combination of frequency, lexical variation, and length points to Wiley as the most expansive and directive policy, while Springer remains the leanest and most abstract. The wide variance in both coverage and expression style suggests an evolving policy landscape where standardization is still absent (see Koplin 2023; Khalifa and Albadawy 2024; Mugambiwa 2024; Chen and Gong 2025; Gao et al. 2025; Lendvai 2025; Van Niekerk et al. 2025). Together, these patterns highlight the fragmented and sometimes contradictory nature of current AI policy development among the Big Five academic publishers (Table 2).

Since we were intrigued to better understand distribution in a more comprehensive manner, we conducted a Z-score analysis (for details, see Andrade 2021) which was aimed to reveal distinct editorial tendencies among the five academic publishers. Wiley stands out for overemphasizing Accountability ($Z = 1.56$), Authorship ($Z = 1.33$), and Ethics ($Z = 1.52$), while, in contrast, Elsevier is the dominant contributor to Peer Review ($Z = 1.56$), the latter indicating a sharper focus on regulating AI tools during manuscript evaluation. Implying a preference for clearly stated limitations and definitions. Taylor and Francis show relatively elevated concern for Prohibited Use ($Z = 1.24$) and Specificity ($Z = 1.01$). Meanwhile, SAGE and Springer consistently fall below average in nearly all categories, with Springer's Z-scores dipping particularly low for Authorship ($Z = -1.48$) and Disclosure ($Z = -1.14$), suggesting limited policy elaboration. Notably, Permissible Use is significantly highlighted only by Wiley ($Z = 1.56$), while most others underrepresent it (e.g. SAGE: $Z = -0.82$). The Z-score of -1.26 for Specificity at SAGE also points to vague or less directive language in its AI guidance. We visualized the distribution in percentages using a streamlined graph to show the differences (Fig. 1).

To see the interconnectedness of the nine unique codes we employed a co-occurrence matrix analysis where we handled all five PGs together. The network as a whole is undirected and weighted, with a total of 36 unique edges across the 9 codes representing the nine coded themes. The average edge weight is ~8.28, though the distribution is skewed, with a few very strong edges and many weak ones. In terms of overall node strength (sum of edge weights), the most prominent codes are Authorship (117), Accountability (99), and Ethics (93), reinforcing their centrality in AI policy discourse. The least central are Permissible Use (24) and Prohibited Use (25), indicating that fewer publishers address these operational aspects in depth.

The co-occurrence network of AI policy codes shows that Authorship is the most connected node, with a degree of 8 and a total edge weight (sum of co-occurrences) of 117. The most frequent code pair is Authorship and Accountability, with a co-occurrence value of 27, followed by Authorship and Ethics (22), and Accountability and Ethics (21). These high-weight edges suggest that publishers often discuss AI authorship together with responsibility and ethical considerations. Disclosure also has strong links, particularly with Ethics (13) and Accountability (12), pointing to the importance of ethical framing in disclosure requirements.

Peer Review and Editorial Use co-occur 11 times, suggesting that AI guidance in peer review is often linked to editorial roles and responsibilities. Specificity appears less central in absolute degree (connected to 7 other codes), but still maintains moderate edge weights, such as with Authorship (15) and Accountability (13), indicating that publishers offering more specific policies often tie this specificity to rules on authorship and responsibility. Permissible Use and Prohibited Use are the least connected themes, both with fewer and weaker connections: Permissible Use co-occurs most with Authorship (6) and Accountability (6), while Prohibited Use connects weakly to Accountability, Ethics, and Peer Review (all 5) (Fig. 2).

The semantic network constructed from the Big Five academic publishers' AI policy documents reveals a rich structure of conceptual interrelations (Fig. 3). The final edge list comprises co-occurrences between terms, with the strongest link observed between 'ai' and 'use' ($n = 114$), followed by 'ai' and 'tool' ($n = 108$) and 'generative' and 'ai' ($n = 72$). Though these may seem natural given the scope of our research, their prominence points to strong associations and also suggest a dominant thematic cluster around the deployment and nature of AI tools. Network-level statistics indicate a moderately dense graph, with key concepts forming both tightly knit communities and bridge-like connectors across clusters. The most central node by all metrics is 'ai,' having the highest degree (265), betweenness centrality (30,153.44), and eigenvector centrality (1.000). These results outline that

Table 2. Comprehensive table of AI PGs based on the 9 codes, word count, and TTR.

Codes	Elsevier	SAGE	Springer	Taylor and Francis	Wiley
Accountability	8	3	1	9	23
Authorship	25	9	2	10	35
Disclosure	6	9	2	1	11
Editorial_Use	12	8	3	10	6
Ethics	5	4	4	8	25
Peer_Review	13	7	2	5	4
Permissible_Use	5	0	3	0	11
Prohibited_Use	1	1	0	5	7
Specificity	3	0	1	5	29
Total word count	2,288	747	693	1,206	6,097
UniqueWords	616	333	383	546	1,911
TypeTokenRatio	0.269	0.446	0.553	0.453	0.313

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

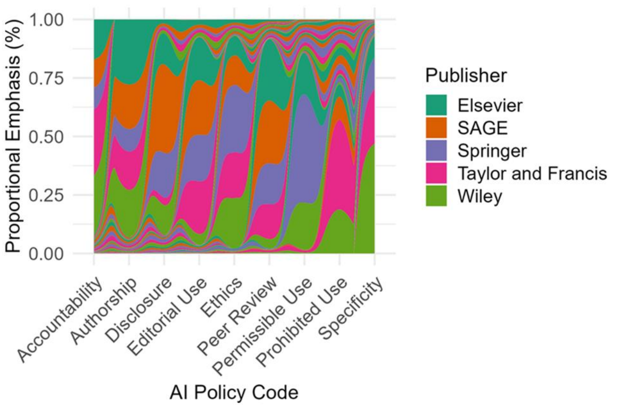

Figure 1. Distribution of the nine thematic codes across Big 5 publishers.

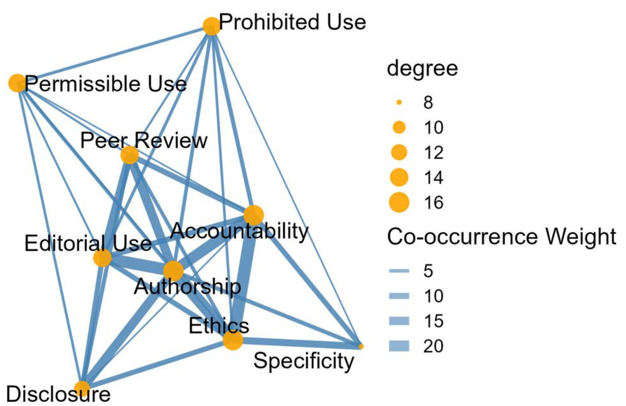

Figure 2. Co-occurrence matrix analysis network of the nine codes.

the discussions of AI lie at the heart of policy discourse, serving as a hub that links otherwise distinct concepts.

Other high-centrality nodes include 'use' (degree = 188; betweenness = 12,220.68) and 'author' (degree = 153), both pointing to frequent discussions surrounding authorship and practical applications which confirm the aforementioned prevalence of authorship-related questions in PGs. The centrality statistics for coded thematic categories support this segmentation: 'Authorship' holds the highest eigenvector value (1.0), suggesting it anchors policy discourse semantically and structurally.

The Jaccard Similarity Matrix further clarifies alignment patterns: Elsevier and Wiley exhibit perfect similarity (1.000), suggesting shared structural policy frameworks. Taylor & Francis also aligns closely with Elsevier and Wiley (both 0.8889), forming a high-similarity cluster. In contrast, Springer shows the lowest average similarity, with its weakest overlap being with SAGE (0.6667), reinforcing its outlier status. Despite moderate scores, SAGE has strong similarity with Taylor & Francis (0.875) which suggest a significant conceptual overlap despite quantitative gaps. To sum up, the results highlight a bifurcation in policy maturity, with Wiley and Elsevier leading in scope and coherence, while Springer trails in both coverage and convergence (Fig. 4).

4. Discussion

4.1. Summary of the results

This study examined how the five largest academic publishers, namely, Elsevier, Springer, Wiley, Taylor & Francis, and SAGE, are responding to the emergence of AI in scholarly communication through formal authorship policy statements. In addressing RQ1, the analysis finds a foundational consensus, as all five publishers explicitly prohibit AI systems from being named as authors and require disclosure of AI-generated content. This convergence, however, is rather unbalanced as deeper policy structures diverge markedly in both thematic coverage and regulatory tone. Variations across the five examined policies point to the fact that while the publishing industry recognizes the need to act, it is doing so through disparate institutional logics, ranging from normative leadership to cautious deferral.

In response to RQ2, the thematic analysis reveals distinct priorities across publishers. Authorship, accountability, and disclosure are consistently emphasized, forming a moral and procedural core across most policies. Yet other critical areas, such as permissible versus prohibited use, peer review guidance, and editorial responsibilities, are addressed unevenly (with Wiley being a notable exception) or omitted. For instance, there are varying levels of addressing the differences between how authors and reviewers may use AI and there is a lack of formal explanation as to 'what' AI-assistance is in practice. Thematic specificity also varies: some policies use declarative, enforceable language, while others rely on suggestive or advisory tones that leave room for ambiguity. These differences reflect a more complex problem where there are not only varied levels of policy maturity but also competing conceptions of what responsible AI integration should look like in academic publishing.

Answering RQ3, structural comparisons using Jaccard similarity show strong alignment between Wiley, Elsevier, and Taylor & Francis, revealing a convergence in policy architecture despite differing tones and emphases. Springer is a clear structural outlier, with minimal thematic overlap and lower code density. Co-occurrence analysis reinforces the centrality of authorship, ethics, and accountability across policies, suggesting a shared ethical grammar even amid policy fragmentation. Yet the underrepresentation of themes like Prohibited Use and Permissible Use points to uncertainty or strategic avoidance in tackling controversial edge cases.

To sum up the results in a more narrative sense, despite general consensus on key issues such as authorship and disclosure, deeper divergences in regulatory vision and implementation are clear and apparent. The comparative analysis shows that publishers are not simply reacting to technological change but actively negotiating their institutional identities and values through policy language. Where some treat AI as a technical tool to be integrated with careful procedural control, others frame it as a threat to core epistemic norms. The presence of detailed provisions (or absence thereof) offers differences in perceived urgency, editorial philosophy, and risk tolerance. Importantly, the language of accountability and ethics often appears without operational clarity. By this we mainly mean that although concerns about the enforceability of stated principles are usually present in the texts, there is a great amount of ambiguity as to how editors and reviewers should execute the stipulations set forth. Furthermore, given that AI detection tools are still highly prone to drawing incorrect conclusions (Elkhatat et al. 2023)—if they are even available to reviewers—it may also be problematic that a potential AI-use by a reviewer cannot be empirically proven. This may result in further complications with regard to the enforceability of these principles and gives space to potential biases.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

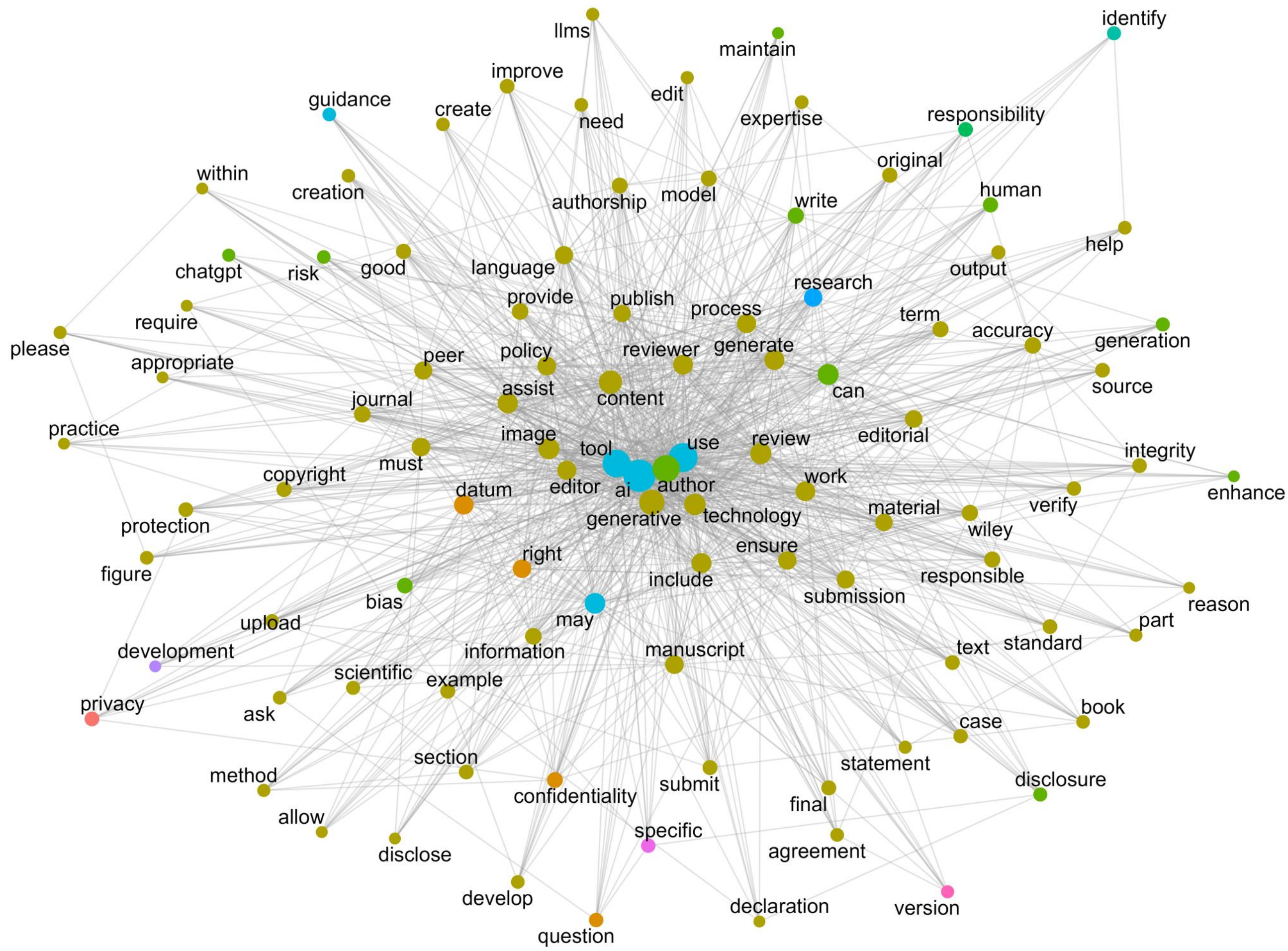

Figure 3. Semantic network of AI PGs. Top 100 terms were included and lemmatized.

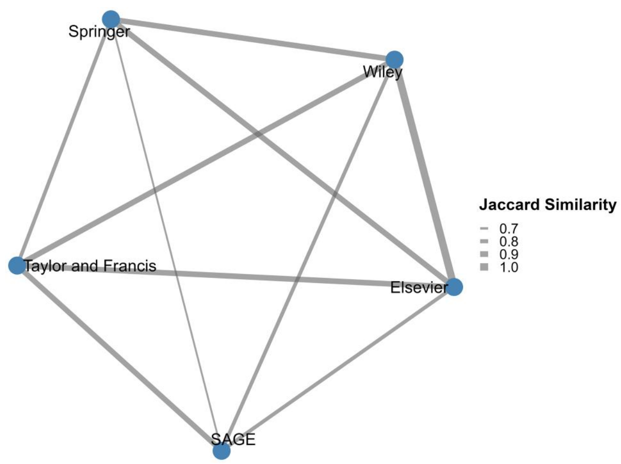

Figure 4. Similarity network based on the Jaccard Similarity Matrix of the Big 5 PGs texts.

Subsequently, the observed inconsistencies suggest a regulatory environment that is still coalescing, shaped as much by internal editorial politics as by external legal or ethical pressures.

To present a comprehensive list of content in the PGs, we created a summary table. The summary table is aimed to demonstrate that formal consensus exists only on baseline ethical positions, such as denying AI authorship and mandating disclosure. Not only did we summarize our findings presented in this study but we also outlined specific aspects that are ambiguous (not outlined explicitly or in great detail) (Table 3).

4.2. Where to go now?—Implications and critical reflections on PGs

The current wave of AI authorship policies signals a growing institutional need to respond to technological disruption; however, the responses themselves often mask more than they reveal. This is not to describe that publishers do not appear proactive in reacting to AI. Alss, many policies function more as symbolic gestures rather than enforceable instruments. This dilemma raises a fundamental question: are these policies designed to regulate actual practice or to perform institutional credibility?

Several documents adopt an ethics-heavy rhetoric while offering a few actionable protocols, relying on the moral vocabulary of trust and accountability without operational grounding. Such language may satisfy reputational concerns, but it does little to guide day-to-day editorial decisions or clarify borderline cases (cf Laher 2025). As mentioned before, there is also a striking absence of enforcement mechanisms. Policies do, indeed, specify what authors should or should not do (Laher 2025) but at the same time, they rarely state what will happen if these directives are ignored or violated.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

Table 3. Summary table of key criteria among Big 5 publishers.

Criterion	Elsevier	Taylor and Francis	Sage	Springer	Wiley
Explicitly prohibits AI as author (not fulfilling criteria)	Yes	Yes	Yes	Yes	Yes
Specifies AI tools cannot meet authorship criteria	Yes	Yes	Yes	Yes	Yes
Requires disclosure of AI tool usage	Yes	Yes	Yes	Yes	Yes
Mentions standard authorship guidelines (e.g. cope and not publisher-specific)	Yes (Elsevier guidelines, also RELX)	No	Yes (Sage guidelines and COPE)	No, though links are available to publishing ethics guidelines	No, though links are available to publishing ethics guidelines
Separates guidelines for authors vs. reviewers vs. editors	Yes	Yes	Yes	Yes, partly (editors are not mentioned in detail)	Not directly
AI use allowed for writing/linguistic assistance	Yes	Yes	Yes	Yes	Yes
AI use allowed for peer review assistance	Ambiguous. Uploading the manuscript is not allowed, assistance shall be reported	Ambiguous. Uploading the manuscript is not allowed, assistance shall be reported	Ambiguous. Reviewers are responsible for AI-assistance, uploading files is generally prohibited, language polishing is acceptable	Ambiguous. Uploading the manuscript is not allowed, assistance shall be reported	Not mentioned explicitly
Requires human accountability for AI-generated content (authors, reviewers, editors)	Yes	Yes	Yes	Yes	Yes
Mentions legal responsibility (e.g. copyright, liability)	Yes, briefly	Yes	Yes, briefly for regarding peer reviewing	Briefly, in particular regarding AI generated images	Yes, in great detail
Tone/style	Legalistic, procedural	Moderate, policy-driven	General, abstract	Minimalist, deferential	Ethical, instructive

In this sense, publishers appear to have partly outsourced both trust and responsibility to the author and reviewer community, assuming good faith where procedural checks may be warranted. Equally concerning is the vague boundary between AI assistance and authorship (see the assistance and generation problem presented in Laher 2025). While most policies prohibit listing AI as an author, they offer little guidance on when AI-generated content becomes substantial enough to compromise authorship claims. We are not claiming that this problem is to be solved by academic publishers; as a matter of fact, even copyright law and intellectual property legislation seem to be struggling to find a clear and comprehensible way to formulate this problem (Bukhari et al. 2023; Watiktinnakorn et al. 2023; Kuai 2024; Gaffar and Albarashdi 2025; Thambaiya et al. 2025), especially, since the copyright laws differ extensively state-by-state and region-by-region (Zhuk 2024; Quintais 2025). The lack of harmonsied rules, however, create a gap, or, using legal jargon, a 'grey zone' where significant intellectual labor can be delegated to AI without violating any stated rule, especially if disclosure is vague or partial. Moreover, by focusing heavily on generative AI, many policies overlook other AI applications such as citation generation, automated summarization, or idea scaffolding (Lendvai 2025). These uses, though less visible, have just as much potential to shape knowledge production, therefore, the appearingly narrow focus on generative AI and large language models like ChatGPT may blind policy frameworks to broader epistemic shifts underway.

However, a deeper tension lies in the policies' implicit definition of authorship. After reviewing the five policies, we can state that these PGs assume that authorship is tied to intentionality, accountability, and intellectual contribution but these values are themselves socially constructed and historically contingent. This a highly positivist and traditional view which is certainly challenged by the use of AI in academic writing. In attempting to 'protect' authorship from AI, publishers inadvertently reify a notion of authorship that may already be out of sync with collaborative, interdisciplinary, and increasingly mediated scholarly practices. Using a more practical approach, if AI becomes integrated into legitimate workflows, the distinction between tool and co-contributor becomes increasingly difficult to defend. Subsequently, the fear of AI authorship, then, may reflect not just ethical concern but also anxiety about destabilizing the current gatekeeping architecture of academic publishing. There are also inconsistencies in how responsibility is distributed. Some policies speak as though responsibility ends at the level of the author, with little reflection on editorial and institutional complicity. Yet editors, peer reviewers, and even publishers themselves are adopting AI tools more and more whether for plagiarism detection, desk rejection triage, or metadata

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

tagging (see Leung et al. 2023; Mollaki 2024). Since acknowledgement is more explicitly articulated for authors, and less so for reviewers and editors own AI use risk, publishers may risk creating a two-tiered system where AI is suspect when used by authors but acceptable when deployed by institutional actors. This asymmetry further undermines the ethical coherence of the policies and raises questions about transparency and power, too. The absence of clear standards for how reviewers and editors may or may not use AI opens the door to opacity, bias, and unequal application. Naturally, it is outlined in almost all PGs that reviewers may not load manuscript to AI tools. However, this action cannot be traced at all and leaves a pivotal gap in the already fragile reviewing procedure. Let's propose an example. If a reviewer does, indeed, upload a paper that they have to review to generate ideas about the peer review, is it considered generation or just simply assistance? Moreover, how and more importantly, who, can challenge whether a reviewer or even an editor used AI to review the manuscript submitted? If the answer is the author or authors, by the nature of the procedure of academic publishing, they have to prove their assumptions and that the AI-generated or AI-assisted content violated the peer reviewing principles—a task that is seemingly impossible to execute perfectly and without any reasonable doubts. PGs also fail to comprehensively address scenarios in which confidential material may be leaked to third-party systems and thereby raising data privacy and intellectual property concerns. This omission may not be accidental either as it allows publishers to retain AI-powered efficiency while holding authors to higher ethical standards. The politics of AI governance, in other words, are not just about the rules themselves but about who gets to 'break them'. Perhaps most troubling is the reactive nature of these policies. They seek to control behavior *after* a problem has emerged rather than proactively shaping the terms of engagement. Prevention, of course, cannot be stipulated efficiently since a paper has to be submitted before checking whether AI has been employed during the research. Nonetheless, the PGs generally miss a critical opportunity to articulate a positive vision for how AI could be ethically and constructively integrated into academic workflows (cf Wiley's PG). Instead of simply listing prohibitions, policies could imagine a framework for human–AI collaboration that preserves integrity while embracing innovation. Framing AI mainly as a threat with benefits mostly being mentioned regarding language improvement, publishers may inadvertently discourage experimentation and stigmatize legitimate use cases. This risk is particularly acute for researchers from under-resourced institutions who may rely on AI tools to meet editorial standards set forth by more privileged peers.

Most policies also fail to address the geopolitical and linguistic dimensions of AI integration. Indeed, PGs outline the positive aspects for linguistic improvements through AI tools but for scholars writing in non-native English or working outside the Global North, these technologies can serve as crucial forms of access. Without culturally sensitive provisions, blanket restrictions on certain AI use may inadvertently reinforce existing inequalities. A policy that forbids AI-assisted editing, for instance, might disproportionately disadvantage scholars from linguistically marginalized backgrounds and in this way, even seemingly neutral rules can entrench structural inequities. It is also problematic as to how to define linguistic assistance. Is, for example, a full translation of a manuscript assistance or generation given the aforementioned rigid understanding of authorship? How can authors assure that their language improvement measures conducted with AI do not cast a shadow of a doubt over a potential misuse of technology? These questions are crucial since a declaration at the end of the paper, which is proposed by all PGs, serves very little in answering them. Finally, the performative clarity of the policies belies the epistemic instability they are meant to manage. Generative AI challenges not just technical processes but fundamental assumptions about originality, labor, and scholarly voice. It prompts a rethinking of what it means to 'write', to 'contribute', and to 'own' knowledge. In attempting to domesticate this disruption through policy, publishers walk a fine line between safeguarding integrity and resisting transformation. The challenge going forward is not merely to draw better boundaries, but to ask more radical questions about the future of knowledge production in an age of algorithmic co-authorship. If current policies are any indication, the publishing industry is still hesitating to fully confront this paradigm shift.

To end the subsection on the problematization and critical remarks, we propose that the findings of this study also carry important implications for the field of research evaluation. Among a myriad of other criteria and parameters, research evaluation (in a general sense) heavily relies on stable and transparent systems of attribution, accountability, and integrity—whether it be in authorship, editorial processes, or peer reviewing. From this perspective, the current, fragmented state of AI governance among the Big 5 publishers as revealed in our results is not just an editorial issue but a *direct challenge* to the infrastructures of evaluation since questions concerning the attribution of authorship, the disclosure of AI use, or the ethics of peer review all constitute the very mechanisms by which evaluators make judgments about scholarly quality. Substantiating this claim and adding to the disclosure requirements, one crucial implication stems from the fact that disclosure requirements vary in granularity. Uniformity cannot be required, naturally—after all, referencing the Latin aphorism, 'varietas delectat', it is acceptable and agreeable that each publisher—being private companies after all—regulates AI use as they seem fit. However, the lack of minimal uniformity in conformity requirements may lead evaluators to examine research on uncertain grounds—or much rather, ground rules. By this we mean that the evidentiary record of authorial contribution may differ across outputs, complicating assessments of originality, labor, and intellectual responsibility. Let us give an example. If an author outlines in the acknowledgment section that 'generative AI has been used to enhance language' may be in line with most PGs. However, it does not help evaluators in understanding what enhancement meant; grammar checks (e.g. correct S-V agreements), spelling mistake corrections, synonym findings (e.g. to evade repetition), or on a larger scale, sentence restructurings (correction of awkward sentencing), use of advanced academic language (e.g. introduction of more formal formulations or even jargon potentially unbeknown by the author), or even a potential spill-over where enhancement results in the generation of new ideas (for instance, AI outputs alternative arguments, conceptual framings, or examples that go beyond surface-level editing). Though these nuances are certainly of importance to the editors of a journal or even the reader, for research evaluation, this ambiguity is nothing short of critical significance since the same disclosure

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

statement could describe both minimal copy-editing and substantive intellectual input; however, evaluators have no common criteria to interpret the difference. Looking a bit outside of the box of publishing PGs, for instance, in high(er)-stakes contexts such as tenure review or grant allocation, such opacity may even risk disadvantaging some scholars while benefiting others, depending on how evaluators '*read*' the disclosure. The absence of minimal standardization thus transfers interpretive burden to evaluators who may apply idiosyncratic or even inconsistent thresholds of what constitutes acceptable AI use. Let us also not forget that the current detection systems are hardly reliable for evaluators to use as anchors. As mentioned briefly earlier, these tools frequently produce false positives, disproportionately flagging texts from non-native English speakers, while simultaneously having a harder time to detect more sophisticated AI outputs (for more on these issues see e.g. Liang et al. 2023; Giray 2024). Now, for a field that is inherently dependent on examining comparability of different outputs, the combination of vague disclosure rules and flawed detection technologies poses a serious challenge. Naturally, addressing these gaps through standardized, transparent disclosure practices is just the first step and as detection tools become thoroughly more sophisticated, it may be speculated that PGs will rely more on their capabilities. Yet even with technological progress, we propose that the responsibility cannot and shall not be outsourced to machines alone. As long as evaluation depends on human judgment (even partially), guided by clear and consistent rules, without minimal harmonization of PGs, the legitimacy of research evaluation risks being undermined by inconsistent practices, inequitable outcomes, and diminished trust in the systems that govern scholarly recognition.

4.3. Future agenda and policy recommendations

It should be noted that while this study analyzes policy texts rather than direct empirical data on research evaluation outcomes, the interpretations offered here aim to outline potential directions for future empirical inquiry rather than assert causal claims. Thus, we propose that policy development must move beyond reactive boundary-setting toward proactive design of responsible AI integration frameworks. Future policies should articulate not only what is prohibited or permitted, but also under what conditions, and with what safeguards. Clearer guidance is needed around thresholds of AI contribution that necessitate disclosure or disqualify authorship, especially in collaborative or interdisciplinary projects. Publishers should also address editorial and peer review contexts explicitly, acknowledging their own use of AI tools to avoid double standards. Developing transparent audit mechanisms, such as AI-use declarations or metadata tagging, could support accountability without adding punitive burdens. Policy language should be adaptable and not static; therefore, it should be capable of evolving alongside rapidly shifting technological capabilities. A shared cross-publisher standard or consortium could help unify core principles while leaving room for contextual differences. As mentioned earlier, culturally sensitive provisions are also essential to avoid excluding or penalizing under-resourced scholars who may rely on AI for linguistic or structural support. The inclusion of researcher and reviewer perspectives in future policy formation could increase legitimacy and practical usability. Methodologically, it would also be important to add further layers to the analysis. For instance, socio-semantic networks (cf Basov et al. 2020) or large-scale policy analyses can outline structural differences more clearly with sound empirical background. Finally, it would be essential for publishers to better catalog and archive AI policies. During our research, we found that there is a lack of clear publication dates and there is no accessibility to earlier versions of AI policies. These measures can be easily implemented and would show more transparency for all actors in the publication process.

Finally, to contribute to the research evaluation aspect we introduced earlier, we also propose that future policies should articulate not only what is prohibited or permitted, but also under what conditions and with what safeguards. For instance, much clearer guidelines are needed to preserve the integrity of peer review, a process central to research evaluation. This is, of course, an extremely hard task even more so as empirical evidence points to great complexities in identifying AI-use during the process (see Yu et al. (2025) and also earlier results from Walters (2023) in a more general text detection regard). However, if we position ourselves in the place of an author who received a presumably AI-generated or -assisted review (an experience that may well resonate with the reader of this paper), the evaluative stakes become clear: the credibility of assessment depends on transparency in how reviews themselves are produced. From such a position, a number of dilemmas arise. How could an author distinguish between a review that is partially AI-assisted—for instance, one that uses AI for grammar polishing and one in which substantive judgments are largely generated by an algorithm? If the review seems generic or inconsistent, should the author feel 'entitled' to challenge its legitimacy, and if so, what procedures or appeals would be available? The guidelines seem to be almost silent on this matter. Moreover, raising such concerns may create professional risks. It may not be an overstatement that questioning a review may very well be perceived as 'antagonistic' toward the reviewer in question, or even potentially to the editor, and may as well be damaging to collegial relations. Authors may thus feel trapped in a paradoxical situation we mentioned in the introduction. They can choose between accepting a review they suspect to be 'AI-shaped' or initiating a conflict that is difficult, if not impossible to prove beyond reasonable doubt. These problems also highlight issues of accountability, an issue frequently discussed in PGs per our results (cf the results in Li et al. (2024)) yet mostly in the case of authorship and authorial content generation. If a review contains questionable reasoning or misapplied concepts, who bears responsibility when parts of it were produced with AI assistance: the reviewer, the editor, or the publisher who allowed it? From the author's perspective, the lack of clarity in responsibility chains not only complicates the response to a single review but also undermines trust in the evaluative system more broadly. Though the proverbial 'stone of wisdom' has not yet been beheld by anyone, we argue that one possible avenue to address these dilemmas would be the gradual introduction of audit mechanisms. These might take the form of voluntary and specified AI-use declarations by reviewers or experimental metadata tagging systems that indicate when AI has been employed in the evaluation process. Of course, such measures would not be simple to design nor to enforce since they raise concerns about confidentiality, additional workload, and the possibility of eroding trust and collegiality upon which peer review depends. Yet even modest steps toward greater transparency could provide authors and evaluators with a minimal

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

evidentiary record of AI involvement, reducing the uncertainty that currently surrounds evaluative judgments. Importantly and from a more holistic perspective, these mechanisms should not be limited to the example of peer review. Similar principles could apply to other stages of the publication cycle, for instance, in the case of editorial triage or copyediting where AI may play an 'invisible' assisting role in shaping how research is presented, judged, and ultimately evaluated. Again, while far from definitive solutions, such practices could mark an incremental move toward balancing the legitimate demand for disclosure with the equally important need to preserve the functionality and credibility of scholarly evaluation.

5. Limitations

This study is limited to a single temporal snapshot of publisher policies which may evolve rapidly in response to legal, technological, and reputational pressures. We analyzed only publicly available documents and cannot account for internal practices, unpublished guidelines, or enforcement mechanisms which may serve with more details about AI governance in academic writing and reviewing. Moreover, since the thematic coding relies on interpretive judgment, which, while systematically applied, may reflect some degree of subjective framing. Additionally, the analysis does not cover smaller publishers, open-access platforms, or disciplinary journals, which may have different orientations toward AI use.

Lastly, it should be noted that while this study analyzes policy texts rather than direct empirical data on research evaluation outcomes, the interpretations offered here aim to outline potential directions for future empirical inquiry rather than assert causal claims. Thus, we invite future research to reflect on the issues mentioned in this paper with more robust correlative analyses to further the discussion on AI practices in academic publishing and knowledge production.

Acknowledgements

None declared.

Funding

Gergely Ferenc Lendvai supported by the EKOP-2025-NKE-3-003 university research scholarship program of the Ministry for Culture and Innovation from the source of the national research, development and innovation fund.

References

Anderson, P. A., and Boden, S. D. (2008) 'Ethical Considerations of Authorship', *The International Journal of Spine Surgery*, 2: 155–8. https://doi.org/10.1016/sasj-2008-comment1

Andrade, C. (2021) 'Z Scores, Standard Scores, and Composite Test Scores Explained', *Indian Journal of Psychological Medicine*, 43: 555–7. https://doi.org/10.1177/02537176211046525

Anvari, F., and Lakens, D. (2018) 'The Replicability Crisis and Public Trust in Psychological Science', *Comprehensive Results in Social Psychology*, 3: 266–86. https://doi.org/10.1080/23743603.2019.1684822

Balalle, H., and Pannilage, S. (2025) 'Reassessing Academic Integrity in the Age of AI: A Systematic Literature Review on AI and Academic Integrity', *Social Sciences & Humanities Open*, 11: 101299. https://doi.org/10.1016/j.ssaho.2025.101299

Barabasi, A.-L. (2016). *Network Science*.

Basov, N., Breiger, R., and Hellsten, I. (2020) 'Socio-Semantic and Other Dualities', *Poetics*, 78: 101433. https://doi.org/10.1016/j.poetic.2020.101433

Bauchner, H., and Rivara, F. P. (2024) 'Use of Artificial Intelligence and the Future of Peer Review', *Health Affairs Scholar*, 2: qxae058. https://doi.org/10.1093/haschl/qxae058

Bennett, D. M., and Taylor, D. M. (2003) 'Unethical Practices in Authorship of Scientific Papers', *Emergency Medicine*, 15: 263–70. https://doi.org/10.1046/j.1442-2026.2003.00432.x

Bhavsar, Daivat. et al. (2025) 'Policies on Artificial Intelligence Chatbots Among Academic Publishers: A Cross-Sectional Audit', *Research Integrity and Peer Review*, 10: 1. https://doi.org/10.1186/s41073-025-00158-y

Bin-Nashwan, S. A., Sadallah, M., and Bouteraa, M. (2023) 'Use of ChatGPT in Academia: Academic Integrity Hangs in the Balance', *Technology in Society*, 75: 102370. https://doi.org/10.1016/j.techsoc.2023.102370

Bukhari, S. W. R., Hassan, S. U., and Aleem, Y. (2023) 'Impact of Artificial Intelligence on Copyright Law: Challenges and Prospects', *Journal of Law & Social Studies*, 5: 647–56. https://doi.org/10.52279/jlss.05.04.647656

Butler, L.-A. et al. (2023) 'The Oligopoly's Shift to Open Access: How the Big Five Academic Publishers Profit from Article Processing Charges', *Quantitative Science Studies*, 4: 778–99. https://doi.org/10.1162/qss_a_00272

Chen, C., and Gong, Y. (2025) 'The Role of AI-Assisted Learning in Academic Writing: A Mixed-Methods Study on Chinese as a Second Language Students', *Education Sciences*, 15: 141. https://doi.org/10.3390/educsci15020141

Cheng, A., Calhoun, A., and Reedy, G. (2025) 'Artificial Intelligence-Assisted Academic Writing: Recommendations for Ethical Use', *Advances in Simulation*, 10: 1–9. https://doi.org/10.1186/s41077-025-00350-6

Coser, L. A. (1975) 'Publishers as Gatekeepers of Ideas', *The Annals of the American Academy of Political and Social Science*, 421: 14–22. https://doi.org/10.1177/000271627542100103

Cummings, R. E., Monroe, S. M., and Watkins, M. (2024) 'Generative AI in First-Year Writing: An Early Analysis of Affordances, Limitations, and a Framework for the Future', *Computers & Composition/Computers and Composition*, 71: 102827. https://doi.org/10.1016/j.compcom.2024.102827

Danowski, J. A. et al. (2023) 'Policy Semantic Networks Associated with ICT Utilization in Africa', *Social Network Analysis and Mining*, 13: 1–17. https://doi.org/10.1007/s13278-023-01068-x

Dwivedi, Yogesh K. et al. (2023) 'Opinion Paper: 'So What If ChatGPT Wrote It?' Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy', *International Journal of Information Management*, 71: 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Elkhatat, A. M., Elsaid, K., and Almeer, S. (2023) 'Evaluating the Efficacy of AI Content Detection Tools in Differentiating Between Human and AI-Generated Text', *International Journal for Educational Integrity*, 19: 1–16. https://doi.org/10.1007/s40979-023-00140-5

Gaffar, H., and Albarashdi, S. (2025) 'Copyright Protection for AI-Generated Works: Exploring Originality and Ownership in a Digital Landscape', *Asian Journal of International Law, 1–24*, 15: 23–46. https://doi.org/10.1017/s2044251323000735

Ganjavi, Conner. et al. (2024) 'Publishers' and Journals' Instructions to Authors on Use of Generative Artificial Intelligence in Academic and Scientific Publishing: bibliometric Analysis', *BMJ*, 384: e077192. https://doi.org/10.1136/bmj-2023-077192

Gao, R. et al. (2025) 'Legal Regulation of AI-Assisted Academic Writing: Challenges, Frameworks, and Pathways', *Frontiers in Artificial Intelligence*, 8: 1546064. https://doi.org/10.3389/frai.2025.1546064

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

Gendron, Y., Andrew, J., and Cooper, C. (2022) 'The Perils of Artificial Intelligence in Academic Publishing', *Critical Perspectives on Accounting*, 87: 102411. https://doi.org/10.1016/j.cpa.2021.102411

Gil de Zúñiga, H., Goyanes, M., and Durotoye, T. (2024) 'A Scholarly Definition of Artificial Intelligence (AI): Advancing AI as a Conceptual Framework in Communication Research', *Political Communication*, 41: 317–34. https://doi.org/10.1080/10584609.2023.2290497

Giray, L. (2024) 'The Problem with False Positives: AI Detection Unfairly Accuses Scholars of AI Plagiarism', *The Serials Librarian*, 85: 181–9. https://doi.org/10.1080/0361526x.2024.2433256

Goyanes, M., Lopezosa, C., and Piñeiro-Naval, V. (2025) 'The Use of Artificial Intelligence (AI) in Research: A Review of Author Guidelines in Leading Journals across Eight Social Science Disciplines', *Scientometrics*, 130: 3725–41. https://doi.org/10.1007/s11192-025-05377-0

Hillary, F. G., and Medaglia, J. D. (2020) 'What the Replication Crisis Means for Intervention Science', *International Journal of Psychophysiology*, 154: 3–5. https://doi.org/10.1016/j.ijpsycho.2019.05.006

Hosseini, M., and Gordijn, B. (2020) 'A Review of the Literature on Ethical Issues Related to Scientific Authorship', *Accountability in Research*, 27: 284–324. https://doi.org/10.1080/08989621.2020.1750957

Hosseini, M. et al. (2025) 'Group Authorship, an Excellent Opportunity Laced with Ethical, Legal and Technical Challenges', *Accountability in Research, 1–23*, 32: 762–84. https://doi.org/10.1080/08989621.2024.2322557

Ioannidis, J. P. A. (2005) 'Why Most Published Research Findings Are False', *PLoS Medicine*, 2 e124. https://doi.org/10.1371/journal.pmed.0020124

Islam, I., and Islam, M. N. (2024) 'Exploring the Opportunities and Challenges of ChatGPT in Academia', *Discover Education*, 3: 1–14. https://doi.org/10.1007/s44217-024-00114-w

Khalifa, M., and Albadawy, M. (2024) 'Using Artificial Intelligence in Academic Writing and Research: An Essential Productivity Tool', *Computer Methods and Programs in Biomedicine Update*, 5: 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Kim, J. et al. (2025) 'Exploring Students' Perspectives on Generative AI-Assisted Academic Writing', *Education and Information Technologies*, 30: 1265–300. https://doi.org/10.1007/s10639-024-12878-7

Kolbinger, F. R. et al. (2024) 'Reporting Guidelines in Medical Artificial Intelligence: A Systematic Review and Meta-Analysis', *Communications Medicine*, 4: /71. https://doi.org/10.1038/s43856-024-00492-0

Koplin, J; Philosophy Documentation Center (2023) 'Plagiarism, Academic Ethics, and the Utilization of Generative AI in Academic Writing', *International Journal of Applied Philosophy*, 37: 17–40. https://doi.org/10.5840/ijap2023372202

Korbmacher, Max. et al. (2023) 'The Replication Crisis Has Led to Positive Structural, Procedural, and Community Changes', *Communications Psychology*, 1: 3. https://doi.org/10.1038/s44271-023-00003-2

Kuai, J. (2024) 'Unravelling Copyright Dilemma of AI-Generated News and Its Implications for the Institution of Journalism: The Cases of US, EU, and China', *New Media & Society*, 26: 5150–68. https://doi.org/10.1177/14614448241251798

Laher, S. (2025) 'Using AI in Academic Writing: What's Allowed and What's Not', *South African Journal of Psychology*, 55: 155–8. https://doi.org/10.1177/00812463251338244

Larivière, V., Haustein, S., and Mongeon, P. (2015) 'The Oligopoly of Academic Publishers in the Digital Era', *PLoS ONE*, 10: e0127502. https://doi.org/10.1371/journal.pone.0127502

Lendvai, G. F. (2025) 'ChatGPT in Academic Writing: A Scientometric Analysis of Literature Published between 2022 and 2023', *Journal of Empirical Research on Human Research Ethics*, 20: 131–48. https://doi.org/10.1177/15562646251350203

Leung, T. I. et al. (2023) 'Best Practices for Using AI Tools as an Author, Peer Reviewer, or Editor', *Journal of Medical Internet Research*, 25: e51584. https://doi.org/10.2196/51584

Li, Z.-Q. et al. (2024) 'Use of Artificial Intelligence in Peer Review Among Top 100 Medical Journals', *JAMA Network Open*, 7: e2448609. https://doi.org/10.1001/jamanetworkopen.2024.48609

Liang, W. et al. (2023) 'GPT Detectors Are Biased Against Non-Native English Writers', *Patterns*, 4: 100779. https://doi.org/10.1016/j.patter.2023.100779

Lin, Z. (2024) 'Towards an AI Policy Framework in Scholarly Publishing', *Trends in Cognitive Sciences*, 28: 85–8. https://doi.org/10.1016/j.tics.2023.12.002

Lund, B. D., and Naheem, K. T. (2024) 'Can ChatGPT be an Author? A Study of Artificial Intelligence Authorship Policies in Top Academic Journals', *Learned Publishing*, 37: 13–21. https://doi.org/10.1002/leap.1582

Marušić, A., Bošnjak, L., and Jerončić, A. (2011) 'A Systematic Review of Research on the Meaning, Ethics and Practices of Authorship across Scholarly Disciplines', *PLoS ONE*, 6: e23477. https://doi.org/10.1371/journal.pone.0023477

Mede, N. G. et al. (2021) 'The 'Replication Crisis' in the Public Eye: Germans' Awareness and Perceptions of the (ir)Reproducibility of Scientific Research', *Public Understanding of Science*, 30: 91–102. https://doi.org/10.1177/0963662520954370

Mehta, V. et al. (2024) 'The Application of ChatGPT in the Peer-Reviewing Process', *Oral Oncology Reports*, 9: 100227. https://doi.org/10.1016/j.oor.2024.100227

Mollaki, V. (2024) 'Death of a Reviewer or Death of Peer Review Integrity? The Challenges of Using AI Tools in Peer Reviewing and the Need to Go Beyond Publishing Policies', *Research Ethics*, 20: 239–50. https://doi.org/10.1177/17470161231224552

Mugambiwa, S. (2024) 'Reaping the Rewards with Minimal Toil: Evaluating the Polemics of Artificial Intelligence in Academia and the Future of Academic Writing', *Edelweiss Applied Science and Technology*, 8: 3535–41. https://doi.org/10.55214/25768484.v8i6.2752

Nanni, L. et al. (2013) 'Different Approaches for Extracting Information from the Co-Occurrence Matrix', *PLoS ONE*, 8: e83554. https://doi.org/10.1371/journal.pone.0083554

Park, C., Lee, J., and Paik, D. (2019). *Identifying Policy Frames Using Semantic Network Analysis*. London: SAGE Publications Ltd eBooks. https://doi.org/10.4135/9781526479907

Parker, J. L. et al. (2025) 'Negotiating Meaning with Machines: AI's Role in Doctoral Writing Pedagogy', *International Journal of Artificial Intelligence in Education*, 35: 1218–38. https://doi.org/10.1007/s40593-024-00425-x

Perkins, M., and Roe, J. (2023) 'Academic Publisher Guidelines on AI Usage: A ChatGPT Supported Thematic Analysis', *F1000Research*, 12: 1398. https://doi.org/10.12688/f1000research.142411.2

Pontille, D., Biagioli, M., and Galison, P. (2004) 'Scientific Authorship. Credit and Intellectual Property in Science', *Revue Française De Sociologie*, 45: 374. https://doi.org/10.2307/3323164

Primack, Richard B. et al. (2019) 'Are Scientific Editors Reliable Gatekeepers of the Publication Process?', *Biological Conservation*, 238: 108232. https://doi.org/10.1016/j.biocon.2019.108232

Quintais, J. P. (2025) 'Generative AI, Copyright and the AI Act', *Computer Law & Security Review*, 56: 106107. https://doi.org/10.1016/j.clsr.2025.106107

Resnik, D. B., and Hosseini, M. (2025) 'Disclosing Artificial Intelligence Use in Scientific Research and Publication: When Should Disclosure be Mandatory, Optional, or Unnecessary?', *Accountability in Research*, 1–13. https://doi.org/10.1080/08989621.2025.2481949

Schooler, J. W. (2014) 'Metascience Could Rescue the 'Replication Crisis', *Nature*, 515: 9. https://doi.org/10.1038/515009a

Shim, J., Park, C., and Wilding, M. (2015) 'Identifying Policy Frames Through Semantic Network Analysis: An Examination of Nuclear Energy Policy Across Six Countries', *Policy Sciences*, 48: 51–83. https://doi.org/10.1007/s11077-015-9211-3

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

Singhal, S., and Kalra, B. S. (2021) 'Publication Ethics: Role and Responsibility of Authors', *Indian Journal of Gastroenterology*, 40: 65–71. https://doi.org/10.1007/s12664-020-01129-5

Szabo, C. (2025). *Unreliable: Bias, Fraud, and the Reproducibility Crisis in Biomedical Research*. Columbia University Press.

Tennant, Jonathan P. et al. (2019) 'Ten Hot Topics around Scholarly Publishing', *Publications*, 7:34. https://doi.org/10.3390/publications7020034

Thambaiya, Nirogini. et al. (2025) 'Copyright Law in the Age of AI: Analysing the AI-Generated Works and Copyright Challenges in Australia', *International Review of Law Computers & Technology*, 39: 1–26. https://doi.org/10.1080/13600869.2025.2486893

Tutuncu, L. (2024) 'Gatekeepers or Gatecrashers? The Inside Connection in Editorial Board Publications of Turkish National Journals', *Scientometrics*, 129: 957–84. https://doi.org/10.1007/s11192-023-04905-0

Van Niekerk, J., Delport, P. M. J., and Sutherland, I. (2025) 'Addressing the Use of Generative AI in Academic Writing', *Computers and Education Artificial Intelligence*, 8: 100342. https://doi.org/10.1016/j.caeai.2024.100342

Vasilevsky, Nicole A. et al. (2021) 'Is Authorship Sufficient for Today's Collaborative Research? A Call for Contributor Roles', *Accountability in Research*, 28: 23–43. https://doi.org/10.1080/08989621.2020.1779591

Walters, W. H. (2023) 'The Effectiveness of Software Designed to Detect AI-Generated Writing: A Comparison of 16 AI Text Detectors', *Open Information Science*, 7: 1–24. https://doi.org/10.1515/opis-2022-0158

Watiktinnakorn, C., Seesai, J., and Kerdvibulvech, C. (2023) 'Blurring the Lines: How AI is Redefining Artistic Ownership and Copyright', *Discover Artificial Intelligence*, 3: 1–10. https://doi.org/10.1007/s44163-023-00088-y

Yu, S. et al. (2025) 'Is Your Paper Being Reviewed by an LLM? Benchmarking AI Text Detection in Peer Review', arXiv Preprint, 1–29. https://doi.org/10.48550/arXiv.2502.19614, preprint: not peer reviewed.

Zhuk, A. (2024) 'Navigating the Legal Landscape of AI Copyright: A Comparative Analysis of EU, US, and Chinese Approaches', *AI And Ethics*, 4: 1299–306. https://doi.org/10.1007/s43681-023-00299-0

Appendix 1—PGs, last accessed on 2 July 2025

Elsevier: https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals#2-for-reviewers

Taylor & Francis: https://taylorandfrancis.com/our-policies/ai-policy/

Sage: https://www.sagepub.com/journals/editorial-policies/artificial-intelligence-policy

Springer: https://www.springer.com/gp/editorial-policies/artificial-intelligence—ai-/25428500

Wiley: https://www.wiley.com/en-us/publish/book/ai-guidelines

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026

© The Author(s) 2026. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Research Evaluation, 2026, 35, 1–13
https://doi.org/10.1093/reseval/rvag004
Article

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvag004/8461575 by Université Montpellier III - Paul-Valéry user on 05 February 2026