OPEN PRAXIS

‘GenAI’ Literature Search Tools and Scholarly Diversity: An Algorithmic Ethnographical Analysis

RESEARCH ARTICLES

KATHY M. CHANDLER
KATY JORDAN
ISHAQ AL-NAABI
PANAGIOTA TZANNI
LEONE GATELY

*Author affiliations can be found in the back matter of this article

INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION

ABSTRACT

The rapid integration of Generative AI tools into academic research raises critical questions about their impact on epistemic justice in scholarly knowledge dissemination. This study examines whether ‘GenAI’ literature search tools amplify or lessen any biases compared to ‘traditional’ literature search databases. Adopting algorithmic ethnography, we analysed 800 search results consisting of the top 20 journal articles in four topics in higher education across 10 literature search platforms: five ‘GenAI’ tools and five ‘traditional’ databases. Metrics included first-author gender, geographic affiliation, numbers of citations and journal impact factors. We found that the search results from ‘GenAI’ literature search tools exhibited substantially greater geographic diversity but showed no consistent advantage in terms of gender balance. Both ‘GenAI’ literature search tools and ‘traditional’ literature search databases prioritised high-citation numbers and high-JIF publications. This paper contributes to the literature that looks beyond perceptions of ‘GenAI’ literature search tools and examines their outputs. The insights gained are useful for researchers and scholars making decisions regarding which platforms to use when conducting searches of academic literature and for those advising them.

CORRESPONDING AUTHOR:
Kathy M. Chandler
Lancaster University, UK
k.m.chandler@lancaster.ac.uk

KEYWORDS:
Generative AI; literature search tools; databases; algorithmic ethnography; epistemic injustice; diversity; bias; gender; geography; higher education

TO CITE THIS ARTICLE:
Chandler, K. M., Jordan, K., Al-Naabi, I., Tzanni, P., & Gately, L. (2026). ‘GenAI’ Literature Search Tools and Scholarly Diversity: An Algorithmic Ethnographical Analysis. *Open Praxis*, *18*(2), pp. 291–309. DOI: https://doi.org/10.55982/openpraxis.18.2.970

INTRODUCTION

This study arose following a conversation between two of the authors of this paper during doctoral supervision in June 2024. The doctoral researcher is undertaking a part-time structured PhD in E-Research and Technology Enhanced Learning. Having already successfully completed a structured set of modules over two years with the support of tutors and peers in their international cohort, they are now engaged in a period of intensive individual research with the support of their supervisor. Supervision meetings take place monthly via videoconferencing and every aspect of the doctoral research is discussed in depth.

The supervisor questioned the doctoral researcher's decision to use only 'Generative AI' ('GenAI') tools to search for literature on their chosen topic and suggested it would be appropriate to use some well-known, 'traditional' literature search databases instead or at least in addition. The doctoral researcher mentioned that an effort was made to search for papers within the 'traditional' databases, but the results were less relevant to the research topic and the searches more time-consuming compared to 'GenAI' tools, which make the process more streamlined. In the absence of university guidance to researchers on the use of GenAI tools for research activities at this time, the focus of the conversation then turned to the ethical implications of researchers using such tools and whether the 'GenAI' literature search tools might be more inherently biased towards research conducted by male researchers in the USA and Europe than the 'traditional' literature search databases. We decided to test this notion with the help of academic colleagues and devised the study reported in this paper.

While ethical concerns about Generative AI (GenAI) in academia, such as training data limitations, language inequities and systemic biases have been well investigated (see for example, Anis & French, 2023; Dhingra et al., 2023; Yan, Sha et al., 2024), empirical research that examines 'GenAI' literature search tools remains critically scarce. Current research focuses on efficiency gains (Kumar & Gunn, 2024; Whitfield & Hofmann, 2023) or output quality (Bjelobaba et al., 2024). However, it does not attempt to quantitatively analyse how these tools contribute to scholarly visibility across gender and geographic lines. This area is particularly urgent given evidence that doctoral students using 'GenAI' search tools report zero ethical considerations about representation (Kumar and Gunn, 2024). Hence, without investigation, we risk automating the marginalisation of Global Majority, female, and early-career scholars under the guise of algorithmic neutrality, a phenomenon Kay et al. (2025) term as algorithmic epistemic injustice.

We should highlight from the outset that so-called 'traditional' academic databases are far from neutral. The choice of which database to use for a systematic literature review can have a marked effect on the results, with one study finding less than 6% commonality in articles retrieved from the same search performed across three 'traditional' databases (Wanyama et al., 2022). 'Traditional' databases already encode epistemic injustices through opaque "relevance" rankings (Jordan & Tsai, 2024). These rankings influence academics' perceptions. An eye tracking experiment has demonstrated that higher ranking results are more often trusted, even when those ranked lower are more relevant to the search topic (Pan et al, 2007).

Researchers should consider power dynamics, cultural competence, and the ways in which they are culturally situated in terms of their own power and privilege when conducting research (Fassinger & Morrow, 2013). In this paper, we operationally define social justice as equitable representation of marginalised voices, measured through gender parity, geographic diversity, and reduced reliance on prestige metrics, such as journal impact factors (JIF) and citations that favour dominant groups in academic research and scholarly work.

We question herewith whether 'GenAI' literature search tools replicate or redress social biases around gender and geography in the field of educational research, responding to the call for empirical studies that examine such biases in depth (Li et al., 2025). Our inquiry is grounded in epistemic injustice theory (Fricker, 2008; Dotson, 2014), which examines how power imbalances distort whose knowledge is deemed credible. When 'GenAI' literature tools algorithmically rank scholarly work, they risk amplifying three forms of injustice: testimonial injustice, hermeneutical injustice and algorithmic epistemic injustice, which are explained in detail later in the paper. Our work is guided by the question: How do 'GenAI' literature search tools represent the gender and geographic diversity of scholarly work in comparison to 'traditional' literature search databases? As universities worldwide race to develop institutional AI research policies, our work is significant for higher education. Higher education institutions might use the findings to influence policy and guide their decisions about subscriptions to 'GenAI' literature search tools.

The 'GenAI' literature tools examined in this paper are the following: Consensus, Elicit, Research Rabbit, Scopus AI and Semantic Scholar. Consensus AI is a search engine that uses AI to extract insights from scientific papers, providing evidence-based answers to research questions (Consensus, n.d.). Elicit uses AI algorithms to automate literature reviews, summarising key findings and identifying relevant studies based on user queries (Elicit Help Center, n.d.), whilst Research Rabbit discovers and organises academic papers through AI-driven recommendations (ResearchRabbit, 2025), creating visual networks of related work. Scopus AI, powered by Elsevier's Scopus database, uses GenAI to summarise research trends, suggest relevant articles, and answer questions about scholarly content (Elsevier, 2025). Semantic Scholar claims to employ advanced Natural Language Processing to analyse and rank academic papers, offering smart citations and highlighting influential research (Semantic Scholar, n.d.). Together, these tools claim to streamline the research process by enhancing discovery, summarisation, and analysis, saving researchers time and improving the efficiency of literature reviews and knowledge synthesis. Each tool integrates GenAI uniquely, catering to different needs in academic and scientific exploration. These 'GenAI' literature search tools were used to generate data but were not used for the following narrative literature review, which used 'traditional' databases.

LITERATURE

Whilst there has been a great deal of recent interest in the topic of GenAI for academic research purposes generally, the literature published to date in the specific area of 'GenAI' literature search tools is somewhat limited. The quality of outputs was a concern in a study to examine the role of GenAI tools as knowledge brokers in mathematics education (Rycroft-Smith & Macey, 2025). In examining whether four different GenAI tools provide accurate information when promoted to summarise research on the topic of Venn diagrams, two literature search-specific tools, Consensus and Elicit, produced accurate, existing references, whilst the other 'GenAI' tools tested, Claude and ChatGPT 4.0, hallucinated references (Rycroft-Smith & Macey, 2025).

Much of the other extant literature in this area relates to users' perceptions. A perception often held about 'GenAI' literature search tools is that they increase efficiency (Kumar & Gunn, 2024; Whitfield & Hofmann, 2023) but evidence to support this suggestion is inconsistent. This claim of enhanced efficiency is supported by the results of a study of ChatGPT-assisted systematic reviews (Spillias et al., 2024) but another empirical study in which groups of researchers were asked to find medical literature in response to a specific question found that groups given access to iris.ai did not produce significantly more results in the same timeframe than the control group who used Google Scholar, Web of Science, and PubMed (Schoeb et al., 2020). In education, however, enhanced efficiency is not necessarily always a positive outcome. This is because it is only achieved with a comparable loss of individual agency in the process (White, 2025).

A study analysing the perceptions of 26 doctoral researchers in Educational Technology at one university in the USA following an assignment in which they were encouraged to explore a range of 'GenAI' literature search tools found that the participants perceived the tools' benefits to be facilitating discovery and making connections between papers, increasing efficiency, and making the literature review process less intimidating (Kumar & Gunn, 2024). The participants reflected, however, that the tools should not replace 'traditional' literature search databases but be supplementary, as they had concerns about the quality of outputs, the tools' inferior usability or features, and concerns around privacy. These concerns reflect those identified around the use of GenAI more generally within academia, along with whether all students have the digital literacy necessary to use such tools well and whether critical engagement with ideas will be diminished by their use (Bozkurt, 2025; Mozelius & Humble, 2024; Whitfield & Hofmann, 2023).

It is notable that the concerns identified by the participants in Kumar and Gunn's study (2024) did not include any ethical issues or concerns around the social justice implications of using GenAI. Moreover, whilst the authors themselves identify a need to understand how such tools might be used ethically, they do not expand on what they mean by 'ethically'. Similarly, whilst an examination of five 'GenAI' literature search tools, including Research Rabbit, Elicit and Consensus, identified concerns around the quality of outputs in relation to the given prompt (Bjelobaba et al., 2024), the same study does not comment on the biases associated with such tools. It does, however, identify such issues in relation to some of the other types of GenAI tools that it examines.

Social justice concerns regarding the use of GenAI for academic purposes include the associated excessive energy consumption (Selwyn, 2022), limitations around the training data used (Anis & French, 2023) and the predominance of English language-based tools, which make them far more useful to those in some societies over others and amplify existing global digital disparity (Yan, Greiff, et al., 2024; Yan, Sha, et al., 2024). However, there are also suggestions that GenAI can enhance equitability in academia through helping researchers overcome the difficulties associated with marginalised backgrounds, speaking English as a second language and a lack of access to resources that privileged academics take for granted (Anis & French, 2023; Bjelobaba et al., 2024).

Whilst some academics highlight the potential harm associated with the systemic biases around gender, race and social class that Large Language Models and GenAI tools are capable of perpetuating (Dhingra et al., 2023; Mei et al., 2023; Yan, Greiff, et al., 2024) and others call for transparent reporting of their use within research and academic writing (Bozkurt, 2024; Tlili et al., 2025), we have yet to consider these concerns in relation to 'GenAI' literature search tools.

THEORETICAL FRAMEWORK

Ideas around epistemic injustice frame this study theoretically and particularly those relating to scholarly diversity around the axes of gender and geography. Feminist philosopher Miranda Fricker identifies two kinds of epistemic injustice (Fricker, 2008), both of which are apparent in the world of academic publishing. The first of these is testimonial injustice, when the author's credibility is diminished by the readers' prejudice. Readers might like to think that we judge each piece of academic work on its own merits but inevitably, we have preconceived ideas about every journal article as we open it, based on the author's name, institutional affiliation, their country of origin, and the status of the journal in which the work is published, for example. Gender bias is also an issue in academia. Despite evidence to suggest that female authors in the field of economics write with greater clarity than their male counterparts, their articles spend up to six months longer under peer review (Hengel, 2022). Journal articles with a female lead author are less likely to be published in prestigious journals and, when they are, they are less often cited than those with male lead authors (Bendels et al., 2018).

The second type of epistemic injustice is hermeneutical injustice (Fricker, 2008), when an unfair distribution of resources means that authors are disadvantaged when trying to understand their own social experience as compared with that of a wider society. This is also apparent in the world of academic writing and publishing, where journals based in Global Majority countries face challenges compared to their richer and more prestigious counterparts in North America and Europe (Bol et al., 2023). Authors with the opportunity to study within elite institutions have privileged access to teaching and resources that enable them to succeed when submitting their work to the journals most likely to be described as 'high quality' (Wellmon & Piper, 2017). Their institutions can afford subscriptions to such journals. Their teachers have published in such journals themselves and can provide advice based on their own interactions with reviewers and editors. Those with access to such resources are more likely to have their work accepted by top journals and when they do, have institutions able to pay the necessary fees to make their work open access. The figures published by one medical journal, for example, show stark regional disparities, with acceptance rates ranging from 24.5% for papers whose authors are based in North America to 0% for papers whose authors are from Africa and the Middle East (Onken et al., 2021).

A further type of epistemic injustice is theorised by Kristie Dotson, which she terms 'third order exclusion' (Dotson, 2014). This is when there are epistemic barriers to making the author's ideas intelligible to the audience. The author's writing may be dismissed as nonsense, dangerous or ridiculed by the dominant knowers who in this case might be the community of academic journal editors and reviewers. Dominant epistemic landscapes are characterised by the affective practices of the community of knowers, and it is their ways of thinking that undermine the credibility of oppressed individuals or groups as producers of knowledge, leading the oppressed to doubt their ability as knowledge producers (Hernandez, 2023). This conceptualisation of epistemic injustice has long been reflected in the suggestion that dominant groups only listen to the oppressed if they couch their ideas in ways that make those in power comfortable (Hill Collins, 2000).

Researchers build on the ideas of Fricker and others to propose a further type of epistemic injustice, which is particularly relevant to our work in this paper: algorithmic epistemic injustice (Byrnes & Spear, 2023). These authors are concerned with the epistemic injustice associated with the rapidly increasing use of machine learning in a medical context, where individuals' knowledge, and particularly that of patients, is devalued by the authority afforded to automated systems. Their argument can also be applied within education and especially in the context of GenAI use (Kay et al, 2025). Kay et al. (2025) identify multiple ways in which GenAI negatively impacts human capacity to understand and trust oppressed and marginalised groups. For example, GenAI, they suggest, can amplify the voices of those who already dominate, thereby increasing testimonial injustice. GenAI can also heighten hermeneutical injustice through widening the information divide between privileged and marginalised groups and distorting or erasing the understanding of the experiences of those on the margins.

'Traditional' academic databases, such as Web of Science, have long played a role in making explicit, distributing and evaluating academic information, determining which research is deemed valuable and whose work is funded and disseminated (Fischer et al., 2020). Whilst GenAI has only recently been recognised as an influential stakeholder in the world of publishing (Bozkurt, 2024), all academic search engines have opaque, poorly understood algorithms that rank some writers' work as 'relevant' over the work of others (Jordan & Tsai, 2024). In this study, we are concerned to determine whether the use of 'GenAI' literature search tools amplifies or reduces this effect in the context of literature reviews for the purposes of educational research.

METHODS

In approaching the research design for the study, we faced the common paradox which faces all researchers wishing to understand algorithmic systems. Systems based on opaque algorithms and artificial intelligence are a classic example of a 'black box problem', when systems are so opaque, that they can only be understood through their inputs and outputs, rather than knowing how they function directly (Castelvecchi, 2016; von Eschenbach, 2021). This makes the need for research to understand their implications more pressing but presents a significant challenge for those attempting to conduct such research effectively.

While generative AI is relatively new as a broad social phenomenon, the issue of how to address this challenge and ways to conceptualise and study algorithms has been an issue in internet research prior to this. One of the authors' previous experiences of scholarship in Internet Studies and digital methods provided a link between these fields and the present study. The critical study of algorithms – algorithms being broadly defined, but particularly in response to search engines and social media feeds – positions algorithms as a form of culture, and as such calls for the use of ethnographic methods (Gillespie, 2016; Seaver, 2017). Although the application of ethnographic methods to understand what is obscured by algorithms has previously been more frequently applied to social media, the principles remain relevant and adopting a stance to "follow the medium" and "consider the Internet not so much as an object of study, rather as a source of new methods and languages for understanding contemporary society" (Caliandro, 2018, p.553) remains fitting. As generative AI becomes increasingly embedded in academic culture and practices, there is likely to be further reconceptualising of the definition of algorithms and thinking about ways to approach ethnographic research (e.g. Cellard, 2022). For the present study, we adopted a methodological approach which can be characterised as aligning with 'algorithmic ethnography' (Christin, 2020a) to address these challenges and begin to understand the ways in which algorithms are mediating academic and scholarly culture in this context. Algorithmic ethnography is defined as "the ethnographic study of the computational systems enabling and shaping online interactions", and "entails paying close attention to the role of algorithms in structuring the back and front end of the digital platforms that increasingly mediate digital exchanges" (Christin, 2020a, p.109). The definition of 'algorithm' is notoriously slippery and multiple (Seaver, 2017); in the context of this project, mindful of this, we adopt the position similar to Christin (2020a) that algorithms are technical mediators between individuals and their interaction with content (in this case, scholarly knowledge) online.

In adopting an algorithmic ethnographic stance, it is important to be attentive to the data infrastructures, the role of algorithmic sorting, and metrics deployed by platforms (Christin, 2020a). Algorithmic ethnography is also fitting given the focus and research question guiding the study, which are aligned with a stance of understanding and interpreting the

effects of the technology in practice, rather than from a technical standpoint. For this project, we applied algorithmic ethnography by running a number of searches for a defined sample of topics, across the same range of academic literature search platforms. We then sought to identify trends across the platforms, drawing upon descriptive statistics and our interpretation of the observations and data. As such, through these processes we sought to enact the key strategies set out by Christin (2020b). The strategy of 'algorithmic comparison' was applied through the research design by seeking to compare features across several different platforms, while 'algorithmic triangulation' acknowledges the role that the algorithms themselves play in the study as a source of data.

The first step involved planning and identifying the object of focus for our inquiry, in terms of both the platforms we would include in the sample, and the searches we would run across the platforms. We sought to balance the sample of platforms to include a range of 'traditional' academic literature searches, alongside a range of tools which also provide a way to search the academic literature but state they use generative artificial intelligence to assist in the process. Through collaborative discussion, ten platforms were identified, as shown in Table 1. Except for Scopus AI, to which one of the authors had access via their institution, all the 'GenAI' literature search tools are payment free versions.

Table 1 Platforms included in the sample for searches.

'TRADITIONAL' DATABASES	'GenAI' TOOLS
Academic Search Ultimate	Consensus
ERIC	Elicit
Google Scholar	Research Rabbit
Scopus	Scopus AI
Web of Science	Semantic Scholar

As the authors are all researchers with a range of experiences in educational technology and related fields, we each nominated potential topics to search for. We selected three topics initially, focusing on ones which were relatively short and well-defined terms, to avoid ambiguity. The three terms included 'breakout rooms', 'learning management systems in higher education', and 'microcredentials'. As we engaged with the analytical process, we opted to add a fourth search term – 'citizen science' – to help verify some of the emergent trends.

Data collection took place between October 2024 and April 2025. All data were collected before the decision that ERIC, a US Department of Education funded 'traditional' database, would no longer be updated (Alonso, 2025). For each of the search terms, a search was conducted on each platform, and the first 20 journal article results were logged in a spreadsheet. Most platforms presented results according to 'relevance ranking' or similar by default (Jordan & Tsai, 2024), apart from Scopus which displayed results according to date by default. The ordering of results in Scopus was toggled to ordering 'by relevance', for comparability. By focusing on the order in which articles are presented by different platforms, we put into practice a focus on algorithmic sorting (Christin, 2020a) and applied the strategy of 'algorithmic comparison', a "similarity-and-difference approach to identify the instruments' unique features" (Christin, 2020b, p.897).

Once the articles were logged, the spreadsheet was also populated with metadata and information relating to each one. Given the focus of the study, and to allow potential sources of bias to be viewed from multiple perspectives, the following data was added: (i) publication date; (ii) number of female first authors; (iv) number of countries represented by first authors and the number of these which are 'Global Majority' countries; (v) journal impact factor according to Web of Science (also indicative of whether the journal is indexed by Web of Science); and (vi) number of citations, according to Google Scholar.

In examining these variables, we were aware of the potential for inaccuracies in our data and the potential for our choice of variables to unhelpfully reinforce existing discourses. In examining the number of female first authors we were dependent on the preferred pronouns indicated by the authors in their biographies in online institutional profiles or social media, so there may have been instances where individual authors were misgendered. Whilst we did not restrict analysis to a binary gender framework, we identified no lead authors with non-binary pronouns in our sample.

We categorised the countries represented according to how many are Global Majority countries, broadly defined as those outside of the USA, Europe, Australia and New Zealand. Whilst we acknowledge that there could be much debate over this imperfect, working definition, we also recognize that most publications come from authors affiliated to institutions within the USA, Europe, Australia and New Zealand and that failing to include this factor would have been an omission in research of this kind.

Once collected, the data were analysed using a descriptive statistics approach (Marshall & Jonker, 2010), with an emphasis on displaying the data and understanding its characteristics and distribution, rather than applying statistical tests.

FINDINGS

In this section, the data will be presented according to each of the factors which were examined, in turn. The focus here is upon presenting the data and identifying trends, with a full discussion bringing together the trends more generally in relation to the research question in the subsequent section.

GENDER

Figure 1 shows the number of female first authors within the first 20 search results, for each platform, colour-coded according to the search topic.

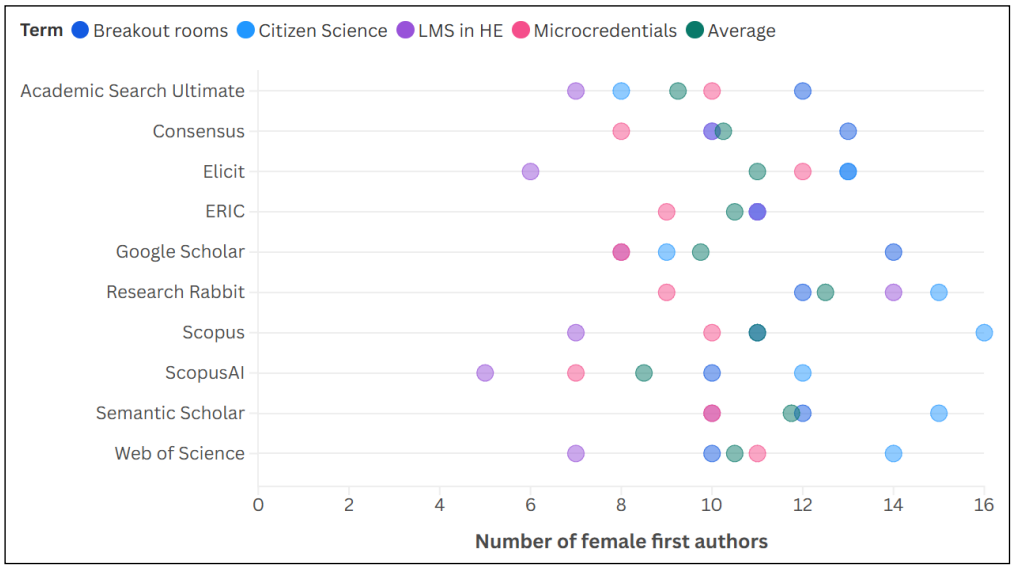

Figure 1 Number of female first authors in the first 20 papers returned from each platform included in the sample. Colour of dots denotes the search term, and green dots show the mean average for each platform. Note that in instances where fewer than five data points are shown, this is due to overlapped data points of the same value – please refer to the colour coding.

Figure 1 suggests that the proportion of female first authors shows variation by platform, but also varies substantially according to each topic, so any platform differences are less clear. Focusing on the mean average per platform (as shown by green dots), there is an overall tendency towards balanced gender representation. Despite this tendency toward balance, there are platforms which represent the extremes of the distribution, with Scopus AI being furthest below the average (higher proportion of males), and Research Rabbit and Semantic Scholar being above average. Among the conventional academic databases, Academic Search Ultimate is the furthest below average, whilst Scopus identifies more female first authors than any other platform studied. There is also considerable variation between the results for different research topics, with the topics of citizen science and breakout rooms generating far more results with female first authors across all platforms and the topics of microcredentials and learning management systems generating far more results with male first authors.

LOCATION

We examined the geographical range of scholars represented by the sampled papers by looking at the countries given by first authors' affiliations. Location is a known bias within academic publishing (Bol et al., 2023), and the visibility of results within academic database searches (Czerniewicz & Wiens, 2013; Czerniewicz et al., 2017). The number of countries represented by first-author affiliations in the first 20 results are shown in Figure 2.

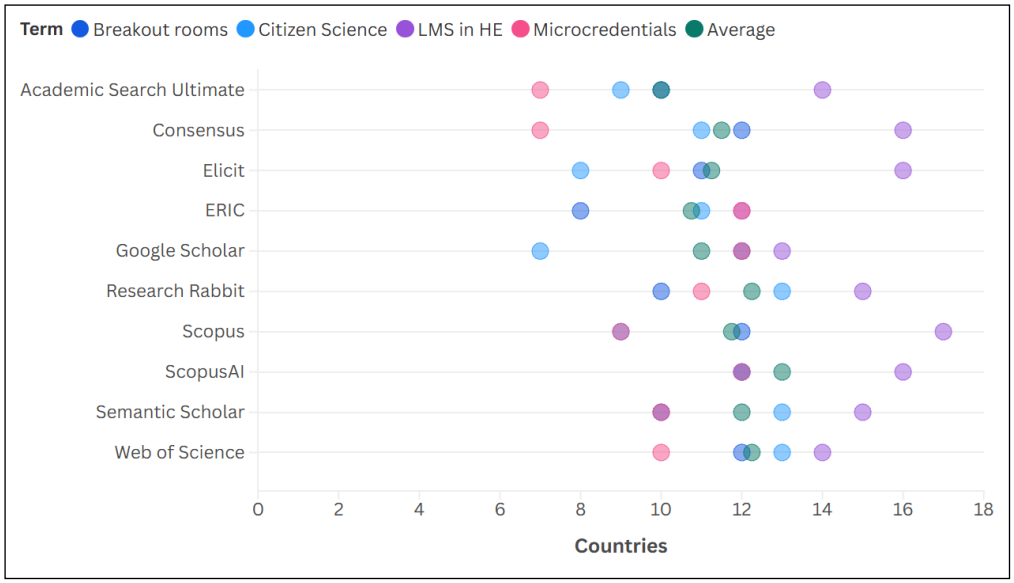

Figure 2 Number of countries represented by first authors in the first 20 papers returned from each platform included in the sample. Colour of dots denotes the search term, and green dots show the mean average for each platform. Note that in instances where fewer than five data points are shown, this is due to overlapped data points of the same value – please refer to the colour coding.

As was the case for gender, Figure 2 shows that the number of countries represented varies by both platform and topic. When considering diversity in terms of the number of countries, it may be notable that the three lowest mean average (less diverse) are 'traditional' literature search databases (Academic Search Ultimate, Google Scholar and ERIC), while two of the three highest are 'GenAI-based' tools (ScopusAI, Research Rabbit). The number of Global Majority countries in the first 20 results are shown in Table 2. Notably, the total for 'GenAI-based' tools is approximately ten percent higher than for 'traditional' literature search databases.

PLATFORM	BREAKOUT ROOMS	MICROCREDENTIALS	LEARNING MANAGEMENT SYSTEMS IN HIGHER EDUCATION	CITIZEN SCIENCE	*TOTAL*
Academic Search Ultimate	4	1	12	2	28
ERIC	4	4	17	3	27
Google Scholar	9	3	15	0	23
Scopus	6	1	13	3	18
Web of Science	5	0	12	1	19
				Total	115
Consensus	9	0	11	1	21
Elicit	6	4	14	0	24
Research Rabbit	9	2	13	1	25
Scopus AI	6	5	13	3	27
Semantic Scholar	9	3	16	1	29
				Total	126

Table 2 Number of papers in the first 20 results per query, per platform, with first authors based in Global Majority countries.

DATE

While publication date is unlikely to be a direct source of bias, it is possible that it may indirectly reflect a role of biased information such as citation counts, which accrue over time, and are known to be a source of bias (Oza, 2023). The distribution of the 20 sampled papers per platform according to date are shown in Figure 3. The distributions vary slightly according to query, but all apart from 'Citizen Science', which typically returns older results, are similar. Consensus, Elicit, ScopusAI and Google Scholar are notably more likely to include older results. The role of citation counts will also be addressed specifically in a subsequent section.

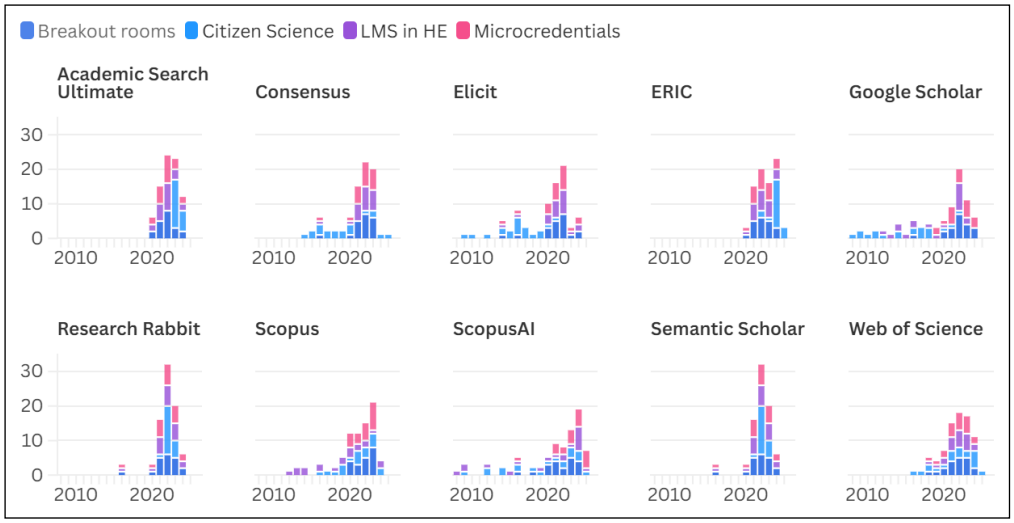

Figure 3 Distribution of publication date (year) for the first 20 papers returned from each platform included in the sample, colour coded according to search topic.

WEB OF SCIENCE INDEXING

For each of the papers included in the sample, we recorded whether it is in a journal which is indexed by Web of Science. This is a notable characteristic, given the focus of our inquiry, for two reasons. First, there are known biases towards whether journals are indexed by Web of Science according to the country or region they are based in (Mills et al., 2023). Second, only Web of Science-indexed journals are given Journal Impact Factor (JIF) metrics, both being products of the parent company Clarivate. While we consider journal impact factors separately, our examination of whether or not a paper was Web of Science indexed provided a way of including and considering papers which are not indexed and do not have a JIF.

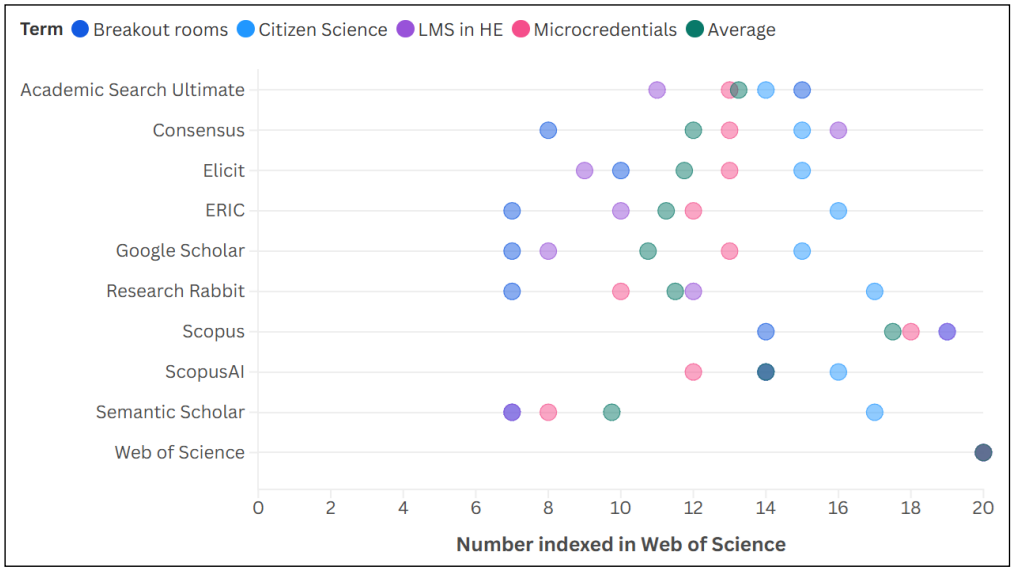

Figure 4 Number of papers in the first 20 papers returned from each platform included in the sample which are published in journals indexed by Web of Science. Colour of dots denotes the search term, and green dots show the mean average for each platform. Note that in instances where fewer than five data points are shown, this is due to overlapped data points of the same value – please refer to the colour coding.

The number of papers published in Web of Science-indexed journals in the first 20 results per query and platform is shown in Figure 4. Here, a clear divide appears between major, publisher-linked platforms and others (both 'GenAI' tools and 'traditional' literature search databases). Web of Science (by definition), Scopus, Scopus AI and Academic Seach Ultimate are all relatively high in terms of this measure. The other platforms are notably lower, which may suggest the inclusion of potentially more diverse sources.

CITATION COUNTS

The number of citations a paper has accrued is a highly influential metric within academic publishing and evaluation of academic performance (Kulczycki, 2023). However, the choice of whose work to cite – and whose work not to cite – is an act which often serves to reinforce power imbalances in academia (Ray et al., 2024). As a widely available, numeric form of data, citation counts may be relatively easy to incorporate in computational ways of ranking search

results. For the sampled papers, Google Scholar was used as a source of citation count data. While citation counts may be provided by the platforms themselves, Google Scholar was used as the source of citation data as it includes a wider range of sources compared to Web of Science, for example, which gives citation counts but only for indexed sources (e.g. Gerasimov et al., 2024). The range of citation counts for the sampled papers per platform, according to each topic, are shown in Figures 5 to 8.

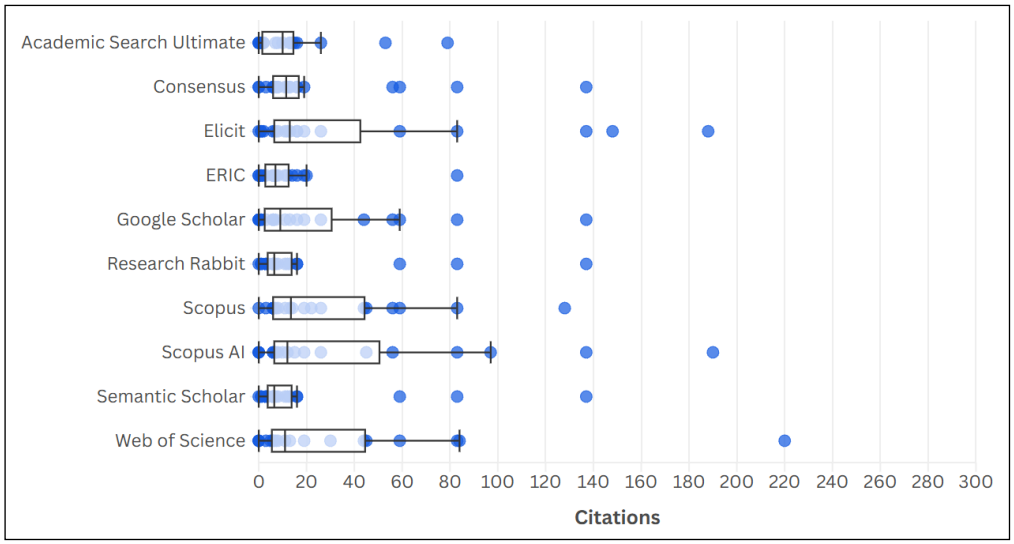

Figure 5 Citation counts according to Google Scholar for the first 20 papers returned from each platform included in the sample on the topic of breakout rooms. Three outliers – articles with citations >300 – are not shown (one each from Academic Search Ultimate, Elicit, and Scopus AI).

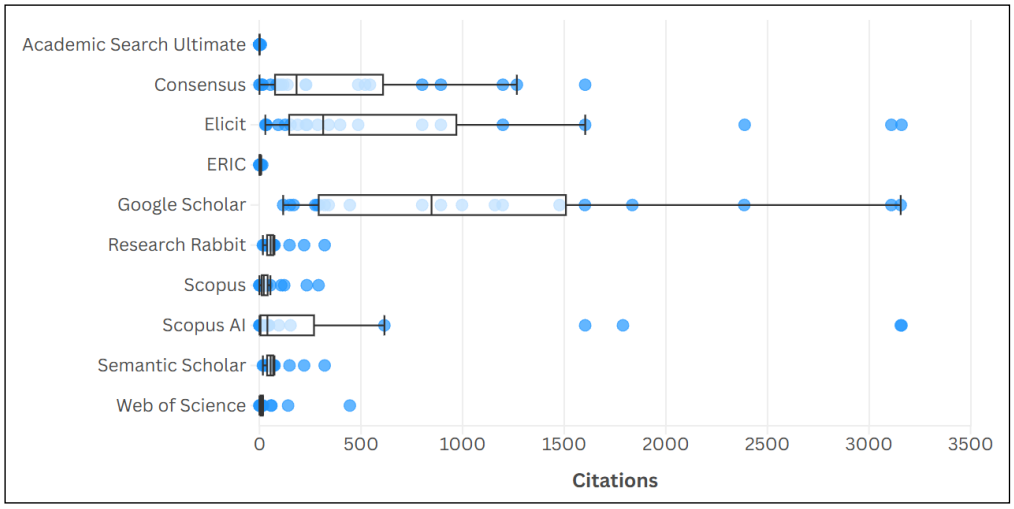

Figure 6 Citation counts according to Google Scholar for the first 20 papers returned from each platform included in the sample on the topic of citizen science.

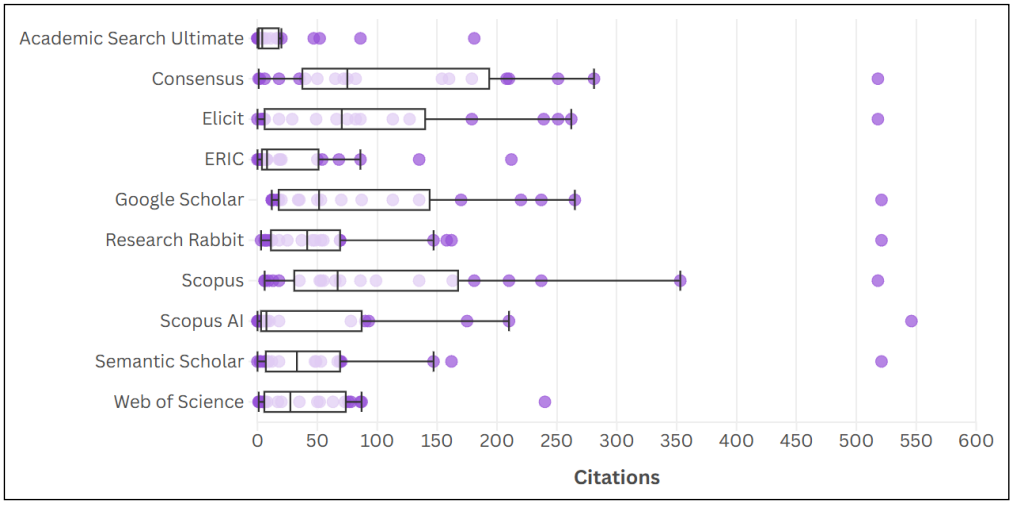

Figure 7 Citation counts according to Google Scholar for the first 20 papers returned from each platform included in the sample on the topic of learning management systems in higher education. Three outliers – articles with citations >600 – are not shown (one from Consensus, and two from Scopus AI).

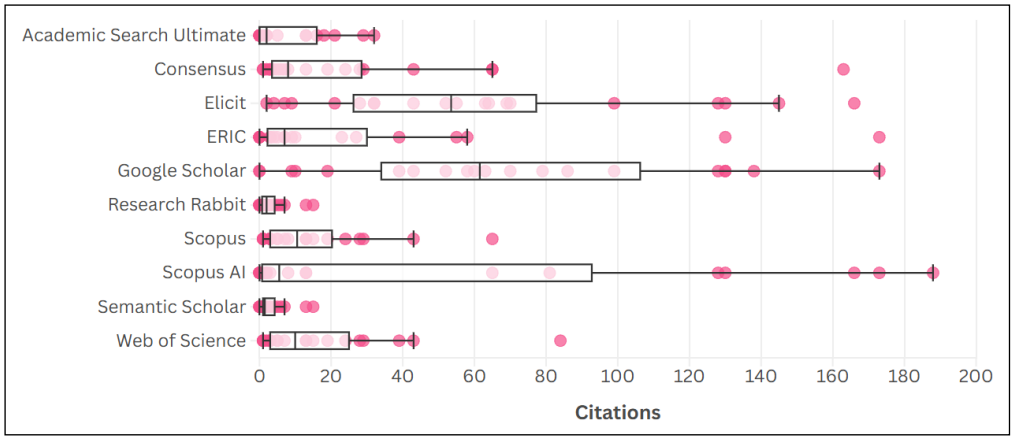

Figure 8 Citation counts according to Google Scholar for the first 20 papers returned from each platform included in the sample on the topic of microcredentials.

Figures 5 to 8 show that the range of citation counts varies considerably according to each topic. As such, it is useful to consider the relative position of platforms. Table 3 shows the platforms per topic, ranked by median, from the largest median citation value (rank 1) to the smallest (rank 10). Within Table 3, there is not a clear distinction between the 'GenAI' tools and 'traditional' literature search databases; the sample of both types shows some which appear to be more influenced by citations, and others which are not. Looking at the 'total' ranking column, a lower aggregate figure here indicates that a platform is more often highly ranked, reflecting a more highly cited sample of papers. Within 'traditional' literature search databases, Google Scholar and Scopus stand out as being more often highly ranked, while this is true of Elicit and Consensus within the 'GenAI' literature search tools.

PLATFORM	BREAKOUT ROOMS	MICROCREDENTIALS	LEARNING MANAGEMENT SYSTEMS IN HIGHER EDUCATION	CITIZEN SCIENCE	*TOTAL*
Academic Search Ultimate	5	8.5	10	10	33.5
ERIC	8	6	8	9	31
Google Scholar	7	1	4	1	13
Scopus	1	3	3	7	14
Web of Science	6	4	7	8	25
Consensus	4	5	1	3	13
Elicit	2	2	2	2	8
Research Rabbit	9.5	8.5	5	4.5	27.5
Scopus AI	3	7	9	6	25
Semantic Scholar	9.5	10	6	4.5	30

Table 3 Ranking by median number of citations for the sampled papers per query.

JOURNAL IMPACT FACTOR

We also examined journal impact factor (JIF), as, alongside citations, this is a key metric within academic publishing and would potentially readily lend itself to computational methods of defining relevance. JIF is linked to the earlier discussions about the binary variable of whether or not a journal is indexed by Web of Science, as not all journals are assigned a JIF by Clarivate.

Boxplots showing the range of JIFs – for the papers within the sample which have them – are shown in Figures 9 to 12.

Similar to the citation data, the range of JIF within the sample varies considerably according to the search topic. Likewise, we consider the relative positions of the median JIF per platform, per query, as shown in Table 4. With the exception of Scopus AI, which tends to use high-JIF sources, the 'GenAI' tools may tend to use lower-JIF sources. However, in the 'traditional' literature search databases group, it is also notable that the two least JIF-linked platforms are Scopus and Web of Science.

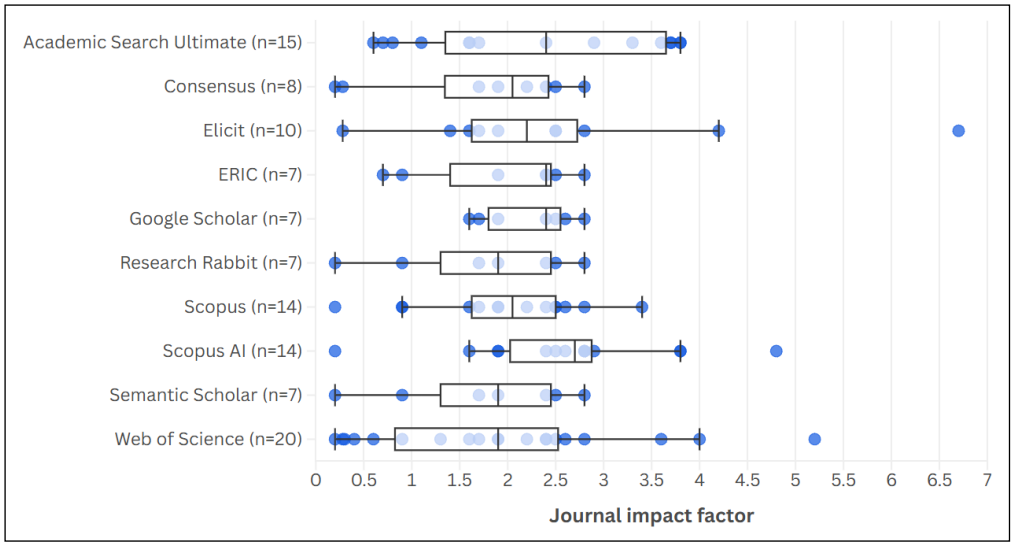

Figure 9 Journal impact factors, for papers published in journals which have them (note differing 'n'), within the first 20 papers returned from each platform included in the sample on the topic of breakout rooms.

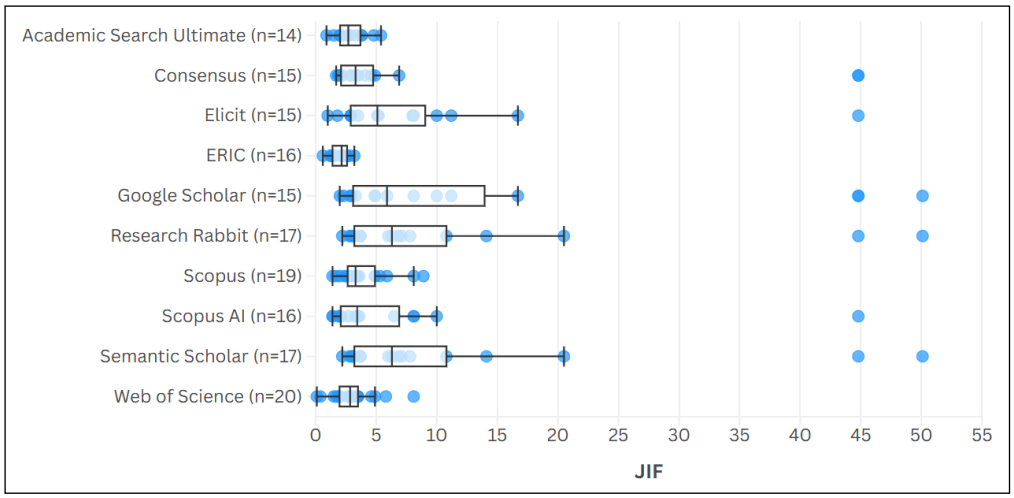

Figure 10 Journal impact factors, for papers published in journals which have them (note differing 'n'), within the first 20 papers returned from each platform included in the sample on the topic of citizen science.

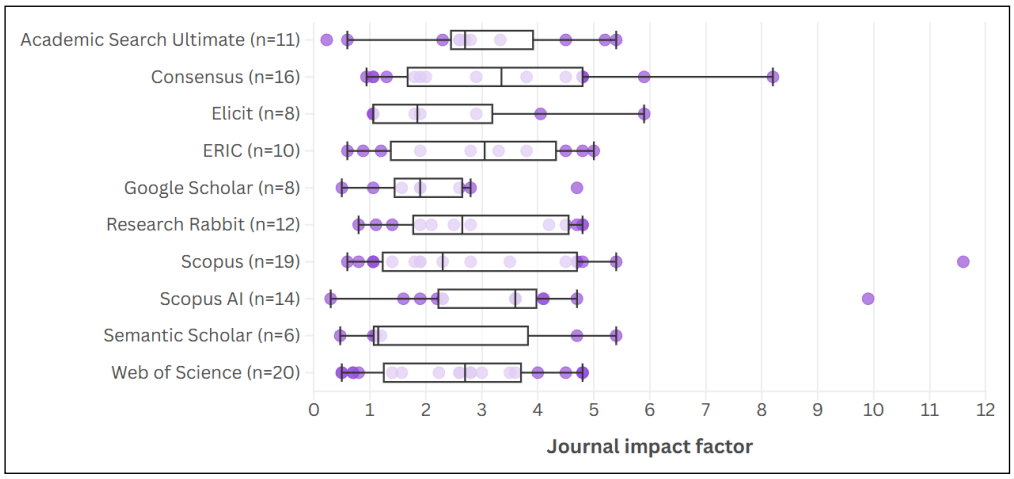

Figure 11 Journal impact factors, for papers published in journals which have them (note differing 'n'), within the first 20 papers returned from each platform included in the sample on the topic of learning management systems in higher education.

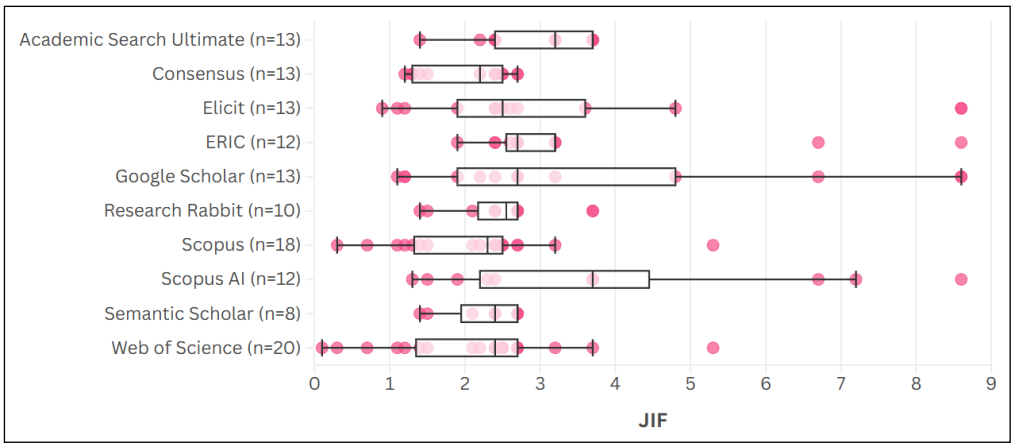

Figure 12 Journal impact factors, for papers published in journals which have them (note differing 'n'), within the first 20 papers returned from each platform included in the sample on the topic of microcredentials.

PLATFORM	BREAKOUT ROOMS	MICROCREDENTIALS	LEARNING MANAGEMENT SYSTEMS IN HIGHER EDUCATION	CITIZEN SCIENCE	*TOTAL*
Academic Search Ultimate	2	2	4.5	9	17.5
ERIC	2	3.5	3	10	18.5
Google Scholar	2	3.5	8	3	16.4
Scopus	5.5	9	7	7	28.5
Web of Science	9	7.5	4.5	8	29
Consensus	5.5	10	2	6	23.5
Elicit	4	6	9	4	23
Research Rabbit	9	5	6	1.5	21.5
Scopus AI	1	1	1	5	8
Semantic Scholar	9	7.5	10	1.5	28

Table 4 Ranking by median JIF for the sampled papers per query.

DISCUSSION

We set out to explore whether 'GenAI' literature search tools have any influential social biases around gender and geography in academic research. Using algorithmic ethnography, we analysed 800 search results retrieved through five 'GenAI' literature search tools and five 'traditional' literature search databases and following we provide a critical discussion of the findings and offer some practical implications for researchers.

'GENAI' LITERATURE SEARCH TOOLS: SOME PATTERNS OVER CLEAR BIAS

To return to our overall question of how 'GenAI' literature search tools represent the diversity of scholarly work around gender and geography in comparison to 'traditional' literature search databases, we found little evidence to support the initial hypothesis that 'GenAI' literature search tools might be inherently biased towards research conducted by male researchers in Global Minority countries. In terms of gender, there is no overall discernible trend. The overall tendency towards a balanced gender representation contrasts with the extent of gender biases in academic publishing overall, which tend to favour male authors and reflect a range of sources of bias in the sector (Larivière et al., 2013). However, given that our focus is on education, within the social sciences, female academics may be more likely to be represented than in some other disciplines (Goyanes et al., 2025), which may account for the overall balance here. There is, however, considerable variation between platforms. Among the 'GenAI' tools, Research Rabbit and Semantic Scholar identified the highest number of articles with female first authors, whilst Scopus AI identified the lowest number. Among the 'traditional' literature search databases, Academic Search Ultimate is the furthest below average, whilst Scopus identifies more female first authors than any other platform studied. There is also considerable variation between the results for different research topics.

In terms of geographic diversity, the 'GenAI-based' searches were more likely to represent a higher number of countries than the 'traditional' database searches, and the proportion of papers identified from authors within institutions in Global Majority countries was approximately ten percent higher for the 'GenAI-based' searches. We have not identified any factors that might account for this difference. Whilst the 'GenAI' platforms, particularly Consensus, Elicit, Scopus AI and Scopus AI were notably more likely to include older results, a finding which may have indicated an algorithmic bias toward sources with higher numbers of citations, there is no clear distinction between the different types of platforms in terms of citation counts. Overall, there is no evidence within our study that 'GenAI' literature search tools display inherent algorithmic epistemic injustice (Byrnes & Spear, 2023) when compared with the 'traditional' literature search databases. Rather, the picture is more nuanced, with individual platforms in both groups displaying different biases when responding to different queries. Whilst the increased diversity associated with the 'GenAI' platforms may seem positive, this difference may be linked to systemic publishing inequities or rooted in hidden patterns of large-scale data extraction and exploitation.

Examination of the number of papers published in Web of Science-indexed journals show a clear divide appears between major, publisher-linked platforms and others, in both 'GenAI' and 'traditional' groups. Whilst Web of Science (by definition), Scopus, Scopus AI and Academic Seach Ultimate are all relatively high in terms of this measure, other platforms are notably lower, which may suggest the inclusion of potentially more diverse sources. Except for Scopus AI, the 'GenAI' tools may tend to use lower-JIF sources. When compiling our data, we noted our own perceptions of the platforms we used for the study, including our tendency to describe the results identified by the 'traditional' database searches and Scopus AI as 'higher quality papers'. These perceptions may well be influenced by our ideas around the value of the major, publisher-linked platforms and journal impact factors. As such, these perceptions also reflect our own contribution to testimonial injustice (Fricker, 2008), when the author's credibility is diminished by the readers' prejudice. Our perceptions of what constitutes a 'high quality' publication are also reflective of the hermeneutical injustice (Fricker, 2008) and 'third order exclusion' (Dotson, 2014) apparent within the institutions within which we study and work and within in the wider world of academic writing and publishing.

ADDITIONAL OBSERVATIONS ON ACCESSIBILITY AND COST

The research process revealed significant social justice concerns regarding accessibility and cost barriers associated with 'GenAI' literature search tools. A primary issue involves geographical restrictions, as certain 'GenAI' tools remain unavailable in specific countries, exacerbating global research inequalities (Castillo-Segura et al., 2023). This challenge became particularly evident when one co-author based in China encountered accessibility limitations, requiring a virtual private network to access some tools due to national internet regulations.

Furthermore, many advanced 'GenAI' tools operate through costly subscription models, creating financial barriers for underfunded researchers and institutions. For instance, Scopus AI remains exclusively available through institutional subscriptions, with only one co-author having access, thereby restricting researchers from less-resourced institutions. Similarly, while Elicit and Consensus offer enhanced features, their monthly upgrade fees (currently $12 and $8.99, respectively) present substantial cost barriers. Although Semantic Scholar and Research Rabbit provide free access, preliminary observations suggest that these tools deliver comparatively lower functionality than non-free alternatives, such as Scopus AI. While existing literature has not thoroughly examined cost-related barriers for these specific tools, scholars universally acknowledge that financial obstacles to research technologies disproportionately affect researchers in low- and middle-income countries and those lacking institutional support, thereby widening global research disparities (Cox & Abbott, 2021; Gabriel, 2024; Khisro & Fenlon, 2025). These accessibility and affordability challenges ultimately limit equitable participation in GenAI-enhanced academic research, particularly for marginalised scholar communities.

TOWARDS A BETTER INTEGRATION: SOME PRACTICAL IMPLICATIONS

Based on the findings, researchers require an appreciation of the variability of platforms, both 'GenAI' literature search tools and 'traditional' literature search databases, and the extent of their individual potential to enhance or limit epistemic diversity. Just as previous research uncovers the potential for some literature to be prioritised over others if only a single 'traditional' database is used (Heck et al., 2024; Wanyama et al., 2022), the same may be true for 'GenAI' review tools. Researchers might use multiple search tools strategically to achieve a higher diversity in results and a higher coverage. Combining geographically inclusive 'GenAI' tools like Semantic Scholar with discipline-specific databases, such as ERIC, could mitigate single-tool biases. Combining 'GenAI' tools with 'traditional' literature search databases might also enhance the geographic diversity of results. It is important to be aware that the diversity of results obtained from individual search tools is likely to vary for different search topics. Researchers need skills to critically evaluate literature search results using a social justice lens to enhance the rigor and trustworthiness of their research.

Higher education institutions must go beyond simple GenAI adoption policies to take account of algorithmic considerations when adopting and providing access to literature search platforms, both 'GenAI' literature search tools and 'traditional' literature search databases. Based on diversity of results according to search topic, they might consider specific tools for certain disciplines. Additionally, given that literature searching requires not only tool literacy

but also the skills to critically evaluate search results, institutions should devise professional development programmes for staff and students to learn how to use the platforms and to learn about their algorithmic advantages and limitations from social justice perspectives. This is an essential part of preparing students for a future that not only requires technical skill but the development of a critical perspective regarding the use of GenAI (Bozkurt et al., 2024; Gupta et al., 2024). As institutions rush to integrate GenAI tools for research, we warn here against conflating diversity metrics with true equity. Surfacing Global Majority scholarship matters little if those works remain algorithmically buried beneath high-JIF publication, so institutions should esteem diverse knowledge generation and epistemologies and encourage scholars to rethink their own contribution to epistemic justice in terms of whose work they read and cite. We affirm Bozkurt's call for the cultivation of the capacity for epistemic resilience, so that humans learn how to 'ask the questions that machines cannot' (Bozkurt, 2025).

CONCLUSION

We conclude that whilst there is there is no overall discernible difference in how the 'GenAI' literature search tools represent the diversity of scholarly work in terms of gender compared with 'traditional' literature search databases, there is a notable, unexpected positive difference in terms of geographical diversity, where the proportion of papers identified from authors within institutions in Global Majority countries was approximately ten percent higher for the 'GenAI'-based tools.

We acknowledge some limitations that need to be considered when interpreting the results of this study. The sample was limited to 10 platforms and four search terms within a single discipline. Except for Scopus AI, the 'GenAI' literature search tools used were free versions. The use of premium versions of these tools might generate different results. In addition, the binary categorisation of so called 'GenAI' tools versus 'traditional' literature search databases may oversimplify the continuum of algorithmic sophistication, as even conventional databases employ AI-like ranking systems, e.g. Google Scholar ranking. Also, the snapshot in time and the ways in which tools are in continuous development or decline might not indicate how algorithmic updates change the results over time.

It is not just authors' geography and gender that determine whose work is most valued in academia. Future research could consider additional factors that might be associated with epistemic injustice in searching for literature, including university ranking. The location of the publisher and those with editorial oversight of a journal article also determine a paper's influence (Czerniewicz et al., 2017). Further research should also include the open access status of papers in the dataset, not from a perspective of bias as such, but to examine whether any evidence could be found to support the possibility that open access publications are more open to exploitation by GenAI (Decker, 2025; Pollock & Michael, 2024). Furthermore, a valuable area for future research would be to consider whether there is evidence emerging about if and how 'GenAI' tools are in turn shaping scholarly norms.

DATA ACCESSIBILITY STATEMENT

The datasets generated and analysed during the current study are available in the Lancaster University repository, https://doi.org/10.17635/lancaster/researchdata/732.

SUSTAINABLE DEVELOPMENT GOALS (SDGs)

This study is linked to the following SDG(s): Quality education (SDG 4), Gender equality (SDG 5), Reduced inequalities (SDG 10).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS (CRediT)

Kathy Chandler: Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. Katy Jordan: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft. Ishaq Al-Naabi: Conceptualization, Data curation, Investigation, Methodology, Project administration, Writing – original draft. Panagiota Tzanni: Conceptualization, Data curation, Investigation, Methodology, Resources, Writing – original draft. Leone Gately: Conceptualization, Data curation, Writing – review & editing. All authors have read and agreed to the published version of the manuscript.

AUTHOR AFFILIATIONS

Kathy M. Chandler orcid.org/0000-0002-9396-5578
Lancaster University, UK

Katy Jordan orcid.org/0000-0003-0910-0078
Lancaster University, UK

Ishaq Al-Naabi orcid.org/0000-0001-8829-2922
Lancaster University, UK

Panagiota Tzanni orcid.org/0000-0003-1584-9923
Lancaster University, UK

Leone Gately orcid.org/0009-0003-5105-8165
Lancaster University, UK

REFERENCES

Alonso, J. (2025, April 23). *Collection of free education research may lose funding*. Retrieved July 18, 2025, from Inside Higher Ed https://www.insidehighered.com/news/quick-takes/2025/04/23/collection-free-education-research-may-lose-funding

Anis, S., & **French, J. A.** (2023). Efficient, explicatory, and equitable: Why qualitative researchers should embrace AI, but cautiously. *Business & Society*, *62*(6), 1139–1144. https://doi.org/10.1177/00076503231163286

Bendels, M. H. K., Müller, R., Brueggmann, D., & **Groneberg, D. A.** (2018). Gender disparities in high-quality research revealed by Nature Index journals. *PLOS ONE*, *13*(1), e0189136. https://doi.org/10.1371/journal.pone.0189136

Bjelobaba, S., Waddington, L., Perkins, M., Foltýnek, T., & **Weber-Wulff, D.** (2024). Research integrity and GenAI: A systematic analysis of ethical challenges across research phases. *arXiv*. https://doi.org/10.48550/arxiv.2412.10134

Bol, J. A., Sheffel, A., Zia, N., & **Meghani, A.** (2023). How to address the geographical bias in academic publishing. *BMJ Global Health*, *8*(12). https://doi.org/10.1136/bmjgh-2023-013111

Bozkurt, A. (2024). GenAI et al.: Cocreation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis*, *16*(1), 1–10. https://doi.org/10.55982/openpraxis.16.1.654

Bozkurt, A. (2025). Algorithmically manufactured minds: Generative and agentic AI in a time of post-truth, reconfiguration of student agency and death of critical pedagogy. *Open Praxis*, *17*(2), 206–210. https://doi.org/10.55982/openpraxis.17.2.792

Bozkurt, A., Xiao, J., Farrow, R., Bai, J. Y. H., Nerantzi, C., Moore, S., Dron, J., Stracke, C. M., Singh, L., Crompton, H., Koutropoulos, A., Terentev, E., Pazurek, A., Nichols, M., Sidorkin, A. M., Costello, E., Watson, S., Mulligan, D., Honeychurch, S., ... Asino, T. I. (2024). The manifesto for teaching and learning in a time of generative AI: A critical collective stance to better navigate the future. *Open Praxis*, *16*(4), 487–513. https://doi.org/10.55982/openpraxis.16.4.777

Byrnes, J., & **Spear, A.** (2023). Epistemic injustice and algorithmic epistemic injustice in healthcare. *International Conference on Computer Ethics*. Illinois Institute of Technology in Chicago, IL: International Conference on Computer Ethics, pp. 1–4. May 16–18, 2023, Chicago, IL, United States. https://journals.library.iit.edu/index.php/Soremoindex.php/CEPE2023/article/view/238

Caliandro, A. (2018). Digital methods for ethnography: Analytical concepts for ethnographers exploring social media environments. *Journal of Contemporary Ethnography*, *47*(5), 551–578. https://doi.org/10.1177/0891241617702960

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature News*, *538*(7623), 20. https://doi.org/10.1038/538020a

Castillo-Segura, P., Alario-Hoyos, C., Kloos, C. D., & **Fernandez Panadero, C.** (2023). Leveraging the potential of generative AI to accelerate systematic literature reviews: An example in the area of educational technology. IEEE IFEES World Engineering Education Forum and Global Engineering Deans Council: Convergence for a Better World: A Call to Action, *WEEF-GEDC 2023- Proceedings*. October 23–27, 2023. Monterrey, Mexico. https://doi.org/10.1109/WEEF-GEDC59520.2023.10344098

Cellard, L. (2022). Algorithms as figures: Towards a post-digital ethnography of algorithmic contexts. *New Media & Society*, *24*(4), 982–1000. https://doi.org/10.1177/14614448221079032

Christin, A. (2020a). Algorithmic ethnography, during and after COVID-19. *Communication and the Public*, *5*(3–4), 108–111. https://doi.org/10.1177/2057047320959850

Christin, A. (2020b). The ethnographer and the algorithm: Beyond the black box. *Theory and Society*, *49*(5), 897–918. https://doi.org/10.1007/s11186-020-09411-3

Consensus. (n.d.). *Less noise. More knowledge.* Retrieved July 18, 2025, from https://consensus.app/home/about-us/

Cox, A., & **Abbott, P.** (2021). Librarians' perceptions of the challenges for researchers in Rwanda and the potential of open scholarship. *Libri*, *71*(2), 93–107. https://doi.org/10.1515/libri-2020-0036

Czerniewicz, L., Goodier, S., & **Morrell, R.** (2017). Southern knowledge online? Climate change research discoverability and communication practices. *Information, Communication & Society*, *20*(3), 386–405. https://doi.org/10.1080/1369118X.2016.1168473

Czerniewicz, L., & **Wiens, K.** (2013). The online visibility of South African knowledge: Searching for poverty alleviation. *The African Journal of Information and Communication (Online)*, *13*. https://doi.org/10.23962/10539/19274

Decker, S. (2025, April 15). *Guest post – The open access – AI conundrum: Does free to read mean free to train?* The Scholarly Kitchen. Retrieved May 10, 2025, from https://scholarlykitchen.sspnet.org/2025/04/15/guest-post-the-open-access-ai-conundrum-does-free-to-read-mean-free-to-train/

Dhingra, H., Jayashanker, P., Moghe, S., & **Strubell, E.** (2023). Queer people are people first: Deconstructing sexual identity stereotypes in large language models. arXiv. https://doi.org/10.48550/arXiv.2307.00101

Dotson, K. (2014). Conceptualizing epistemic oppression. *Social Epistemology*, *28*(2), 115–138. https://doi.org/10.1080/02691728.2013.782585

Elicit Help Center. (n.d.). *Information and advice from the Elicit team*. Retrieved July 18, 2025, from https://support.elicit.com/en/articles/552705

Elsevier. (2025). *Scopus AI: Explore. Focus. Advance*. Retrieved July 18, 2025. from https://www.elsevier.com/en-gb/products/scopus/scopus-ai

Fassinger, R., & **Morrow, S. L.** (2013). Toward best practices in quantitative, qualitative, and mixed-method research: A social justice perspective. *Journal for Social Action in Counseling & Psychology*, *5*(2), 69–83. https://doi.org/10.33043/jsacp.5.2.69-83

Fischer, G., Lundin, J., & **Lindberg, J. O.** (2020). Rethinking and reinventing learning, education and collaboration in the digital age—from creating technologies to transforming cultures. *International Journal of Information and Learning Technology*, *37*(5), 241–252. https://doi.org/10.1108/IJILT-04-2020-0051

Fricker, M. (2008). Forum: Miranda Fricker's Epistemic injustice: Power and the ethics of knowing. *Theoria, An International Journal for Theory, History and Foundations of Science*, *23*(1), 69–71. https://doi.org/10.1387/theoria.7

Gabriel, S. (2024). Generative AI and educational (in)equity. *Proceedings of the International Conference on AI Research, ICAIR 2024*, 4(1), 133–142. December 5–6, 2024. Lisbon, Portugal. https://doi.org/10.34190/icair.4.1.3153

Gerasimov, I., KC, B., Mehrabian, A., Acker, J., & **McGuire, M. P.** (2024). Comparison of datasets citation coverage in Google Scholar, Web of Science, Scopus, Crossref, and DataCite. *Scientometrics*, *129*(7), 3681–3704. https://doi.org/10.1007/s11192-024-05073-5

Gillespie, T. (2016). #Trendingistrending: When algorithms become culture. In R. Seyfert & J. Roberge (Eds.), *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies* (pp. 52–75). Routledge.

Goyanes, M., Demeter, M., Bajić, N. S., & **de Zúñiga, H. G.** (2025). Gender disparities in first authorship: Examining the Matilda effect across communication, political science, and sociology. *Scientometrics*, *130*(5), 2947–2961. https://doi.org/10.1007/s11192-025-05303-4

Gupta, A., Atef, Y., Mills, A., & **Bali, M.** (2024). Assistant, parrot, or colonizing loudspeaker? ChatGPT metaphors for developing critical AI literacies. *Open Praxis*, *16*(1), 37–53. https://doi.org/10.55982/openpraxis.16.1.631

Heck, T., Keller, C., & **Rittberger, M.** (2024). Coverage and similarity of bibliographic databases to find most relevant literature for systematic reviews in education. *International Journal on Digital Libraries*, *25*(2), 365–376. https://doi.org/10.1007/s00799-023-00364-3

Hengel, E. (2022). Publishing while female: Are women held to higher standards? Evidence from peer review. *The Economic Journal*, *132*(648), 2951–2991. https://doi.org/10.1093/ej/ueac032

Hernandez, A. (2023). Epistemic oppression and affective exclusion: A pragmatist approach. *Passion: Journal of the European Philosophical Society for the Study of Emotions*, *1*(2), 154–168. https://doi.org/10.59123/passion.v1i2.13804

Hill Collins, P. (2000). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment* (2nd Ed.). Routledge.

Jordan, K., & **Tsai, S. P.** (2024). Keywords, citations and 'algorithm magic': Exploring assumptions about ranking in academic literature searches online. *Learning, Media and Technology*, 1–15. https://doi.org/10.1080/17439884.2024.2392108

Kay, J., Kasirzadeh, A., & **Mohamed, S.** (2025). Epistemic injustice in generative AI. In S. Das, B. P. Green, K. Varshney, M. Ganapini, & A. Renda (Eds.), *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, *7*, pp. 684–697. October 21–23, 2024. San Jose, California, USA. Available at: https://dl.acm.org/doi/10.5555/3716662.3716722

Khisro, J., & **Fenlon, K.** (2025). Equity in public access to scientific research results: Insights from federal agency responses to the Nelson Memorandum Policy. *Proceedings of the 58th Hawaii International Conference on System Sciences*, pp. 2144–2153. January 7–10, 2025. Big Island, Hawaii. https://doi.org/10.24251/hicss.2025.263

Kulczycki, E. (2023). *The evaluation game: How publication metrics shape scholarly communication*. Cambridge University Press. https://doi.org/10.1017/9781009351218

Kumar, S., & **Gunn, A.** (2024). Doctoral students' reflections on generative artificial intelligence (GenAI) use in the literature review process. *Innovations in Education and Teaching International*, *62*(4), 1395–1408. https://doi.org/10.1080/14703297.2024.2427049

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & **Sugimoto, C. R.** (2013). Bibliometrics: Global gender disparities in science. *Nature*, *504*(7479), 211–213. https://doi.org/10.1038/504211a

Li, M., Enkhtur, A., Yamamoto, B. A., Cheng, F., & **Chen, L.** (2025). Potential societal biases of ChatGPT in higher education: A scoping review. *Open Praxis*, *17*(1), 79–94. https://doi.org/10.55982/openpraxis.17.1.750

Marshall, G., & **Jonker, L.** (2010). An introduction to descriptive statistics: A review and practical guide. *Radiography*, *16*(4), e1–e7. https://doi.org/10.1016/j.radi.2010.01.001

Mei, K. X., Fereidooni, S., & **Caliskan, A.** (2023). Bias against 93 stigmatized groups in masked language models and downstream sentiment classification tasks. *Proceedings of the 2023 ACM Conference on Fairness Accountability and Transparency*, pp. 1699–1710. June 12–15, 2023. Chicago, USA. https://doi.org/10.1145/3593013.3594109

Mills, D., Kingori, P., Branford, A., Chatio, S., Robinson, N., & **Tindana, P.** (2023). *Who counts? Ghanaian academic publishing and global science*. African Minds. https://doi.org/10.47622/9781928502647

Mozelius, P., & **Humble, N.** (2024). On the use of Generative AI for literature reviews: An exploration of tools and techniques. *Proceedings of the European Conference on Research Methodology for Business and Management Studies*, pp. 161–168. July 4–5, 2024. Porto, Portugal. https://doi.org/10.34190/ecrm.23.1.2528

Onken, J., Chang, L., & **Kanwal, F.** (2021). Unconscious bias in peer review. *Clinical Gastroenterology and Hepatology*, *19*(3), 419–420. https://doi.org/10.1016/j.cgh.2020.12.001

Oza, A. (2023, December 22). *Citations show gender bias—And the reasons are surprising*. Nature. https://doi.org/10.1038/d41586-023-03474-9

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & **Granka, L.** (2007). In google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, *12*(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Pollock, D., & **Michael, A.** (2024, December 10). *News and Views: How much content can AI legally exploit?* Delta Think. Retrieved May 10, 2025, from https://www.deltathink.com/news-and-views-how-much-content-can-ai-legally-exploit

Ray, K. S., Zurn, P., Dworkin, J. D., Bassett, D. S., & **Resnik, D. B.** (2024). Citation bias, diversity, and ethics. *Accountability in Research*, *31*(2), 158–172. https://doi.org/10.1080/08989621.2022.2111257

ResearchRabbit. (2025). *Reimagine research*. Retrieved July 18, 2025, from https://www.researchrabbit.ai

Rycroft-Smith, L., & **Macey, D.** (2025). All sizzle, no steak: AI tools are not able to act as credible knowledge brokers by summarising evidence in mathematics education. In T. Fujita (Ed.), *Proceedings of the British Society for Research into Learning Mathematics*, pp. 1–6. March 1, 2025. Online. https://bsrlm.org.uk/wp-content/uploads/2025/05/BSRLM-CP-45-1-10.pdf

Schoeb, D., Suarez-Ibarrola, R., Hein, S., Dressler, F. F., Adams, F., Schlager, D., & **Miernik, A.** (2020). Use of artificial intelligence for medical literature search: Randomized controlled trial using the hackathon format. *Interactive Journal of Medical Research*, *9*(1), e16606. https://doi.org/10.2196/16606

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, *4*(2), 205395171773810. https://doi.org/10.1177/2053951717738104

Selwyn, N. (2022). The future of AI and education: Some cautionary notes, *European Journal of Education*, *57*(4), 620–631. https://doi.org/10.1111/ejed.12532

Semantic Scholar. (n.d.). Our product. Retrieved July 18, 2025. from https://www.semanticscholar.org/product

Spillias, S., Tuohy, P., Andreotta, M., Annand-Jones, R., Boschetti, F., Cvitanovic, C., Duggan, J., Fulton, E. A., Karcher, D. B., Paris, C., Shellock, R., & **Trebilco, R.** (2024). Human-AI collaboration to identify literature for evidence synthesis. *Cell Reports Sustainability*, *1*(7), 100132. https://doi.org/10.1016/j.crsus.2024.100132

Tlili, A., Bond, M., Bozkurt, A., Arar, K., Chiu, T. K. F., & **Rospigliosi, P. 'Asher'.** (2025). Academic integrity in the generative AI (GenAI) era: A collective editorial response. *Interactive Learning Environments*, *33*(3), 1819–1822. https://doi.org/10.1080/10494820.2025.2471198

von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology*, *34*(4), 1607–1622. https://doi.org/10.1007/s13347-021-00477-0

Wanyama, S. B., McQuaid, R. W., & **Kittler, M.** (2022). Where you search determines what you find: The effects of bibliographic databases on systematic reviews. *International Journal of Social Research Methodology*, *25*(3), 409–422. https://doi.org/10.1080/13645579.2021.1892378

Wellmon, C., & **Piper, A.** (2017). Publication, power, and patronage: On inequality and academic publishing. *Critical Inquiry*. https://doi.org/10.6084/M9.FIGSHARE.4558072.V1

White, D. (2025, March 11). *Agency vs Efficiency (The AI learning gambit)*. David White: Digital and Education. Retrieved March 13, 2025, from https://daveowhite.com/aigambit/

Whitfield, S., & **Hofmann, M. A.** (2023). Elicit: AI literature review research assistant. *Public Services Quarterly*, *19*(3), 201–207. https://doi.org/10.1080/15228959.2023.2224125

Yan, L., Greiff, S., Teuber, Z., & **Gašević, D.** (2024). Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behaviour*, *8*(10), 1839–1850. https://doi.org/10.1038/s41562-024-02004-5

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & **Gašević, D.** (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, *55*(1), 90–112. https://doi.org/10.1111/bjet.13370

TO CITE THIS ARTICLE:
Chandler, K. M., Jordan, K., Al-Naabi, I., Tzanni, P., & Gately, L. (2026). 'GenAI' Literature Search Tools and Scholarly Diversity: An Algorithmic Ethnographical Analysis. *Open Praxis*, *18*(2), pp. 291–309. DOI: https://doi.org/10.55982/openpraxis.18.2.970

Submitted: 31 July 2025
Accepted: 05 March 2026
Published: 02 June 2026

Open Praxis is a peer-reviewed open access journal published by International Council for Open and Distance Education.

INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION