

Exploring the effects of journal article features: Implications for automated prediction of scholarly impact

Giovanni Abramo¹, Ciriaco Andrea D'Angelo^{2†}, Leonardo Grilli³

¹ University Telematica Mercatorum, Rome 00186, Italy

² Department of Engineering and Management, University of Rome "Tor Vergata", Rome 00133, Italy

³ Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence 50134, Italy

Citation: Abramo, G., D'Angelo, C. A., & Grilli, L. (2025). Exploring the effects of journal article features: Implications for automated prediction of scholarly impact. *Journal of Data and Information Science*, 10(2), 13–39. <https://doi.org/10.2478/jdis-2025-0010>

DOI: 10.2478/jdis-2025-0010

CSTR: 32295.14.jdis-2025-0010

Received: Sep. 3, 2024; Revised: Dec. 19, 2024; Accepted: Feb. 7, 2025

ABSTRACT

Purpose: Scholars face an unprecedented ever increasing demand for acting as reviewers for journals, recruitment and promotion committees, granting agencies, and research assessment agencies. Consequently, journal editors face an ever increasing scarcity of experts willing to act as reviewers. It is not infrequent that reviews diverge, which forces editors to recur to additional reviewers or make a final decision on their own. The purpose of the proposed bibliometric system is to support of editors' accept/reject decisions in such situations.

Design/methodology/approach: We analyse nearly two million 2017 publications and their scholarly impact, measured by normalized citations. Based on theory and previous literature, we extrapolated the publication traits of text, byline, and bibliographic references expected to be associated with future citations. We then fitted a regression model with the outcome variable as the scholarly impact of the publication and the independent variables as the above non-scientific traits, controlling for fixed effects at the journal level.

Findings: Non-scientific factors explained more than 26% of the paper's impact, with slight variation across

disciplines. On average, OA articles have a 7% greater impact than non-OA articles. A 1% increase in the number of references was associated with an average increase of 0.27% in impact. Higher-impact articles in the reference list, the number of authors and of countries in the byline, the article length, and the average impact of co-authors' past publications all show a positive association with the article's impact. Female authors, authors from English-speaking countries, and the average age of the article's references show instead a negative association.

Research limitations: The selected non-scientific factors are the only observable and measurable ones to us, but we cannot rule out the presence of significant omitted variables. Using citations as a measure of impact has well-known limitations and overlooks other forms of scholarly influence. Additionally, the large dataset constrained us to one year's global publications, preventing us from capturing and accounting for time effects.

Practical implications: This study provides journal editors with a quantitative model that complements peer reviews, particularly when reviewer evaluations diverge. By incorporating non-scientific factors that significantly predict a paper's future impact, editors can make more informed decisions, reduce reliance on additional reviewers, and improve the efficiency and fairness of the manuscript selection process.

Originality/value: To the best of our knowledge, this study is the first one to specifically address the problem of supporting editors in any field in their decisions on submitted manuscripts with a quantitative model. Previous works have generally investigated the relationship between a few of the above publication traits and their impact or the agreement between peer-review and bibliometric evaluations of publications.

Keywords Peer review; Manuscripts evaluation; Non-scientific factors; Scholarly impact; Bibliometrics; Artificial intelligence

1 Introduction

In the current knowledge-based economy, pursuing the effectiveness and efficiency of national research systems is among the top priorities of visionary governments. Performance-based fund allocation systems are implemented to stimulate the improvement of research productivity and to maximize the socioeconomic returns on government research spending. The assessment of researchers' and research institutions' performance requires the measurement of the impact of their research activities. A growing number of bibliometricians and economists have got intrigued by the challenge and a rich literature on the topic has flourished (Abramo, 2018; Bornmann, 2013, 2017; Caputo et al., 2022; Grant et al., 2010; Miettinen et al., 2015; Milat et al., 2015; Penfield et al., 2014; Smit & Hessels, 2021; Wilsdon, 2016).

For a research result to have an impact, it needs to be used (OECD/Eurostat, 2018). Scholarly impact refers to use within the scientific community, while societal impact implies that new knowledge contributes to economic progress, societal well-being, or other public goods. It can be expected that, in general, higher scholarly impact is a harbinger of a higher societal impact. Furthermore, a remarkable heterogeneity of societal impact occurs per discipline, especially for

research in STEM vis-a-vis arts, humanities, and social sciences (Budtz Pedersen et al., 2020).

An aspect of impact assessment in evaluation systems often overlooked or not sufficiently emphasized is that one cannot measure the actual research impact. Instead, at best, one can try to predict it. In fact, for a definitive measurement of impact, one should wait until the scholarly or societal impact of research results is over (Abramo, D'Angelo, & Felici, 2019). Evaluation systems then have to deal with the embedded tradeoff between the level of accuracy of prediction and timeliness in impact measurement. The longer the time elapsed from the "publication" of the research result (i.e., the closer to the end of its impact life-cycle), the more accurate the prediction.

The indicators and methods for predicting the impact of research results depend on the timing of the assessment. Regarding scholarly impact, bibliometric techniques work exceptionally well and promptly as a complement to or substitute for peer review (Abramo, D'Angelo, & Reale, 2019; Aksnes & Taxt, 2004; Reale et al., 2007; Rinia et al., 1998). This is especially true for large-scale evaluations, where the time and cost of peer-review assessment are extremely high. Recently, artificial intelligence (AI) techniques have been applied to predict the future impact of publications; however, studies on this subject are restricted to single research fields (Alohali et al., 2022; Beranová et al., 2022; Himani et al., 2022; Rosenkrantz et al., 2016; Thelwall, 2024).

Bibliometrics builds on the axiom that citing a publication implies using the knowledge embedded in the cited publication (Bloor, 1976; Mulkey, 1976). The higher the number of citations, the higher the scholarly impact of the work.

A critical threshold for any research result to provide an impact, apart from personal usage or trade secrets, is the possibility of making it available to the public. This requires that new knowledge be encoded in written form and then published. Most research results have been published in scientific journals.

Owing to a global rush to publish stimulated by competitive mechanisms being introduced in many countries, the ever-increasing number of submissions has put pressure on editors and reviewers. "In 2022, the article total was 47% higher than in 2016, which has outpaced the limited growth, if any, in the number of practicing scientists" (Hanson et al., 2023). Engaging high-quality scholars in the review process is becoming more and more arduous as demand (including research project evaluations, appointments, tenure, and promotions) remarkably overcomes supply. Selecting which manuscripts to publish from a large number of submissions is then a complicated, time-consuming, and complex process, and critics have argued that the editorial review process is not always free from arbitrium, slowness, and bias (Dickersin et al., 1992; Lee et al., 2013; van Lent et al., 2014). Furthermore, it is not infrequent for reviewers' opinions and recommendations to diverge (Cole et al., 1981; Kirman et al., 2019; Schroter et al., 2022), particularly for interdisciplinary studies (Thelwall, Kousha, Stuart, et al., 2023). This is also true when evaluating published manuscripts, such as those submitted to national research assessment exercises (Bertocchi et al., 2015).

The labor-intensive nature of peer-reviewing has prompted journals to partially automate certain aspects, such as plagiarism checking (Memon, 2020), reviewer selection and assignment (Zhao &

Zhang, 2022), and statistics verification (Baker, 2016). Moreover, efforts have been made in large-scale assessment exercises, particularly in STEMM, to replace peer reviews with bibliometrics (Abramo, 2024; Sivertsen, 2017) or AI (Cárdenas, 2023; Kousha & Thelwall, 2024a). The advent of Large Language Models (LLM), like ChatGPT, might offer valuable alternatives to time-consuming peer review (Liang et al., 2023; Wu et al., 2023).

In theory, an LLM might replace human peer reviewers by assessing the quality of academic articles, especially if provided with guidelines on how to conduct the evaluation. While in the process of writing, Mike Thelwall (2024) undertook the pioneering task of assessing ChatGPT's proficiency in assessing the quality of academic journal articles. Simultaneously, Joost de Winter (2024) explored its efficacy in predicting citation counts, Mendeley readership, and social media engagement.

Ideally, a publication should earn citations because of its scientific content, but in fact, empirical evidence has shown that several non-scientific factors are linked to citation outcomes as well (Kousha & Thelwall, 2024b; Mammola et al., 2022).

We investigated the relative effects of these non-scientific features in predicting the future scholarly impact of journal articles after publication. The findings might serve as guidelines to inform the upcoming assessment of research impact by LLM.

To this end, we analyse the 2017 world articles (nearly 2 million) indexed in the Web of Science (WoS) and count the citations of each up to 2022. The five-year citation window should approximate the articles' relative overall scholarly impact (Abramo et al., 2011). Based on theory and empirical evidence, we extrapolate the traits of each publication that one expects to be related to future citations. We then fit a regression model whereby the outcome variable is the publication yearly impact, and the independent variables are the publication non-scientific features related to our research goal, controlling for fixed journal effects.

Recent advancements in AI-driven citation prediction models provide valuable context for our study. For example, NLP-based models analyze textual features, such as article titles, abstracts, and keywords, to predict citation outcomes (Beranová et al., 2022; Rosenkrantz et al., 2016). Topic modeling and embedding techniques, such as BERT, have been used to capture semantic relationships between articles, providing insights into how thematic relevance affects citation patterns (Alohali et al., 2022; Devlin et al., 2019).

Our approach differs in its focus on non-scientific factors, such as article length and open-access status, which are often overlooked in purely content-driven AI models. By emphasizing these structural characteristics, our study complements AI-driven approaches and offers a broader perspective on the drivers of citation impact. Future research could integrate these approaches, combining content-based NLP features with structural and contextual factors to enhance prediction accuracy while maintaining interpretability.

It is worth emphasizing that automated AI systems make predictions using complex machine learning algorithms, which achieve greater accuracy compared to a regression model. Despite significant progress in "explainable artificial intelligence," machine learning algorithms primarily

focus on prediction rather than explanation. This is why we rely on a regression model to assess the role of each non-scientific factor and estimate its effects, all other things being equal. This understanding is crucial for conscious use of automated prediction systems.

Previous works generally investigated the relationship between a few of the above publication traits and its impact or the agreement between peer-review and bibliometric evaluations of publications (Aksnes & Taxt, 2004; Allen et al., 2009; Bornmann & Leydesdorff, 2013; Fu & Aliferis, 2008).

The conceptual framework of this study is rooted in two complementary perspectives: bibliometric theory and the sociology of science. Bibliometric theory posits that citations serve as proxies for the scholarly use and impact of research output. This principle underpins our choice of citation counts as the dependent variable for assessing impact. However, empirical evidence and theoretical critiques from the sociology of science highlight that citations are influenced not only by the intrinsic quality of a publication but also by various non-scientific factors.

From the sociology of science perspective, we draw on Merton's normative theory Merton (1973), which suggests that citations reflect intellectual influence and credit allocation. However, we also consider constructivist critiques that emphasize the role of social, institutional, and cultural factors in shaping citation practices. These include variables such as gender, institutional affiliation, and open-access status, which may reflect systemic biases or structural disparities in academia.

Guided by this dual framework, our analysis investigates how measurable non-scientific features of publications—categorized as traits of the manuscript, byline, and reference list—affect their scholarly impact. We employed a fixed-effects linear model to isolate the effects of these variables while controlling for journal-specific characteristics, thus addressing the potential influence of publication venue on citation outcomes.

This framework allowed us to examine the relative contributions of diverse factors to citation impact, bridging the theoretical insights of bibliometrics and sociology, with practical implications for research evaluation. By doing so, we aim to provide a more nuanced understanding of the dynamics underlying scholarly recognition and impact.

The proposed investigation can be identified within the Quality & Reproducibility School framework of peer-review improvement (Waltman et al., 2023). Using a quantitative system makes it possible to reduce the subjectivity of the review and (theoretically) improve the accuracy of the assessment. Besides, the methodology can also be attributed to the Efficiency & Incentives School, streamlining the review process, reducing the pressure on the peer review system, and the opportunity costs of reviewing.

While numerous studies have examined individual non-scientific factors influencing citation outcomes, such as open access or the number of authors, the interplay between these factors across disciplines remains underexplored. Furthermore, existing models often focus on a single field or fail to incorporate a comprehensive range of publication characteristics.

Our study addresses these gaps by providing a large-scale, cross-disciplinary analysis of how diverse non-scientific factors—categorized as traits of the manuscript, the byline, and the reference

list—affect citation impact. Using a fixed-effects linear model applied to over 1.4 million articles from 13 disciplines, we identify key differences in how these factors operate across fields.

This work contributes to the literature by i) introducing a holistic framework that integrates multiple non-scientific variables within a single analysis; ii) highlighting cross-disciplinary variations in citation dynamics, thereby providing insights into the contextual factors that shape scholarly recognition; and iii) offering actionable implications for the design of equitable and discipline-sensitive research evaluation systems.

In the following, Section 2 reviews the relevant literature on the topic. Section 3 describes the data and methods. Section 4 presents the results of the analysis. Finally, Section 5 concludes the work with our comments on the main findings and a discussion about their implications.

2 Non-scientific factors affecting citation rates

In perfect markets, high-quality products, all other equals, should sell comparatively more. In reality, markets are not perfect, and all others are not equal. Services incorporated into or associated with a product and its marketing affect buyers' decisions. Non-proprietary knowledge, like that encoded in a written form aimed at final publication in scientific journals, is a public (open access publication) or near-public (non-open access publication) good. It is, therefore, easily transferable within and across sectors and territories and can be used repeatedly without wearing out and diminishing its value.

In the sociology of science, the normative theory based on a milestone work of Robert Merton (1973) affirms that scientists, through the citation of scientific work, recognize credit towards a colleague whose results they have used, meaning that the citations represent an intellectual or cognitive influence on their scientific work. High-quality knowledge (publications), all others equal, should be used more (i.e. earn more citations). Nevertheless, the social constructivist approach (Knorr-Cetina, 1981; Latour & Woolgar, 1979) challenges the validity of the normative theory, arguing that “scientific knowledge is socially constructed through the manipulation of political and financial resources and the use of rhetorical devices” (Knorr-Cetina, 1991), meaning that citations would not necessarily be linked in a consequential manner to the scientific contents of the cited article. In fact, “non-scientific” characteristics of publications may also have a role in determining their citation rates (Tahamtan et al., 2016).

The research questions several evaluative bibliometricians have been fascinated with in recent years are whether: i) it is possible to predict the future scholarly impact of publications objectively; ii) there are factors other than the intrinsic quality of a publication that might affect its impact; and, if yes, iii) to what extent.

There are numerous factors other than quality that affect the citation rates of a manuscript, but not all of them are measurable in large-scale investigations. We distinguish between factors endogenous to the manuscript from exogenous ones. We classified the former into three main categories: i) traits of the manuscript, i.e. manuscript length, language, readability, title, document type, access type, and degree of interdisciplinarity; ii) traits of the byline, i.e. number, gender,

scientific standing of authors, number of organizations, and number of countries; and iii) traits of the reference list, i.e. number of cited works, their average impact, and their age. Among the exogenous factors are worth mentioning: i) the prestige of the hosting journal, although the “direction” of the causal nexus between the quality of a publication and that of the hosting journal is not always clear (Traag, 2021); and the communication initiatives in social media (i.e. blogs, Twitter, Facebook, pre-prints) undertaken by the authors to increase the visibility of the manuscript.

In the following section, we report theoretical and empirical evidence of the link between the above factors and citation outcomes.

The type of access to a publication, open (OA) or non-open (non-OA), affects the ease of knowledge transferability to potential users. One would expect that making access to a publication open, all others equal, would give it a citing advantage over non-OA publications, and in fact, several studies found that OA articles have significantly greater citation counts than non-OA articles (Antelman, 2004; Gargouri, et al., 2010; Wang, Liu, et al., 2015; Yu et al., 2022). However, no significant OA advantage was found in other investigations (Antoniou et al., 2015; Calver & Bradley, 2010; Lansingh & Carter, 2009). A recent review article identified 134 studies on the topic, of which 64 (47.8%) confirmed that OA publications received more citations, 32 (23.9%) found the same only in subsets of their sample, and 1 (0.8%) was inconclusive. In contrast, 37 (27.6%) found no OA advantage (Langham-Putrow et al., 2021). Reasons for contradictory results can be ascribed to differences in the range of disciplines, data sources, methods, date of analysis, and contextual factors. As reasonable as it may seem that OA publications may be more cited than non-OA publications, all other things being equal, one should not be surprised that this may not be the case. OA is an option available to the authors relatively recently. Many prestigious journals have only recently started to publish OA as well, while many newer, generally less prestigious journals only offer authors the OA option. OA articles are likely partly of lower scientific quality than non-OA articles, which would explain why some studies, perhaps not controlling for journals and other factors, do not observe an OA advantage.

The knowledge encoded in publications of different document types (articles, reviews, proceedings papers, letters, books, etc.) has a different average impact, and citations accumulate differently (Wang et al., 2013). All others equal, generally, reviews target a wider audience than articles and, on average, receive more citations (Waltman, 2016). Consequently, the distribution of publications across document types in a journal influences the measurement of the indicators (i.e. IF) that approximate its prestige (Glänzel & Moed, 2002).

Longer papers may embed a greater variety of topics and concepts and, therefore, attract the interest of a larger audience. This feature should lead to a citation advantage over shorter papers, as shown by several investigations (Ball, 2008; Elgendi, 2019; Fox et al., 2016; Xie et al., 2019). In addition, the linguistic properties of a manuscript, its title, and its abstract appear to be associated with citations (Abramo et al., 2016; Didegah & Thelwall, 2013; Heßler & Ziegler, 2022; Rossi & Brand, 2020; Stremersch et al., 2015). Ante (2022) discusses the presence of an incentive for scientists “to (artificially) reduce the readability of their abstracts in order to signal quality and

competence to readers ... and to attract more citations.” The degree of interdisciplinarity of a publication might influence citation outcomes. Empirical evidence on the link between the two shows contrasting results, to be ascribed mainly to the different definitions of interdisciplinarity and the approaches adopted to measure it. Chen, Arsenault, and Larivière (2015) found that in most scientific fields the highly cited publications were those with higher degrees of interdisciplinarity. Yegros-Yegros, Rafols, and D’Este (2015) found instead that very low or very high degrees of interdisciplinarity are associated with a lower number of citations, while some middle degrees with higher citedness.

Abramo, D’Angelo, and Di Costa (2017a) found that highly interdisciplinary research delivers results showing a citation advantage in all disciplines except for Earth sciences, but consider these findings only indicative. Others have revealed that citation outcomes depend greatly on the field (Larivière & Gingras, 2010; Levitt & Thelwall, 2008).

Moving to the features of the byline, there has been shown a positive correlation between the number of authors and citation rates (Abramo & D’Angelo, 2015; Didegah & Thelwall, 2013; Fox et al., 2016; Talaat & Gamel, 2023; Wuchty et al., 2007). The findings seem plausible for two reasons: i) each author is a potential broadcaster of the results encoded in the publication (increased audience through multiple author networks) and a potential self-user for future research advancements (self-citations); ii) collaboration increases the pool of expertise and possibly resources for a research project. Recent empirical evidence showed that in STEMM, larger team size is moderately correlated with higher quality publications, while in arts and humanities with an increased audience (Thelwall, Kousha, Abdoli, et al., 2023). In addition, author experience, reputation, and collaboration network positively influence citations (Hurley et al., 2013; Mammola et al., 2022). Specifically, the average quality and impact of previous co-authored articles might correlate with the impact of future work. Likewise, the number of different organizations listed in the byline, including multiple affiliations (Narin & Whitlow, 1990; Sanfilippo et al., 2018), and the number of countries (Glänzel & de Lange, 2002) are associated with citation advantage, as they are “sounding boards” at institutional and national levels. Regarding the traits of the publication reference list, the number of references might reveal an in-depth analysis by the authors of the topics under investigation and be a predictor of good science. Furthermore, “longer reference lists may make papers more visible in online searches, while also attracting tit-for-tat citations, that is, the tendency of cited authors to cite the papers that cited them” (Mammola et al., 2021). Publications with longer reference lists have been found to covariate with higher citation impact (Alimohammadi & Sajjadi, 2009; Fox et al., 2016), even more so if cited works have higher impact (Jiang et al., 2013; Sivadas & Johnson, 2015), up-to-date (Liu et al., 2022; Mammola et al., 2021) publications. Finally, covariance was found with the number of cited fields and their cognitive distance (Wang, Thijs, et al., 2015), and with the share of self-citations in the reference list (Ruan et al., 2020).

Among the exogenous variables that are likely to correlate with the future impact of the manuscript is the prestige of the hosting journal, approximated by such citation-based indicators

as IF (Mammola et al., 2021; Traag, 2021). Of course, a journal's IF depends on citations received by the hosted publications, so the causal link is somehow bi-directional. However, prestigious journals are likely to attract and select for publishing high-quality manuscripts. Finally, it has been shown that articles with exposure to social media have higher citation rates (Özgent, 2022).

3 Data and methods

3.1 The statistical model

We adopt a linear fixed effects model to evaluate the extent of association of several non-scientific features with the scholarly impact of a given publication (e.g. Rabe-Hesketh & Skrondal, 2022). This model accounts for differences between journals, such as prestige or focus area, by including the fixed effects for each journal. Specifically, for an article i published in journal j , the model is

$$Y_{ij} = \beta_0 + \beta_1 X_{1ij} + \dots + \beta_p X_{pij} + \zeta_j + e_{ij} \quad (1)$$

where Y_{ij} is the outcome variable, i.e. the article (normalized) impact, and X_{1ij}, \dots, X_{pij} are the explanatory variables (article's features).

Specifically, the impact is proxied by citations accrued up to 31/12/2022, and the normalized impact (Y) by citations normalized to the average number of citations of all WoS articles classified in the same subject category (SC)^① and indexed in the same year. For ease of reading, in the following, we call Y_{ij} impact.

The number of authors (X_1) of each article is extracted from its byline, as well as the number of institutions^② and countries (X_2) involved. The average impact of the co-authors' publication portfolio (X_3) is measured on their 2012-2016 scientific output. For this purpose, each author is "disambiguated" using the algorithm proposed by Caron and van Eck (2014), and each relative 2012-2016 publication is assigned a field-normalized impact given by the ratio of citations received (31/12/2022) to the average citations received by all cited publications from the same year and SCs. The disambiguation algorithm of Caron and van Eck is also able to identify the gender composition of the byline; in particular, this gender variable (X_4) is specified through a dummy equal to 1 in the presence of at least one woman among the co-authors and 0 otherwise. Finally, regarding the possible "linguistic advantage," again with reference to the byline, we specify an additional dummy (X_5 , "English"), whose value is 1 in the presence of at least one author with affiliation associated with an English-speaking country, 0 otherwise.

The binary character of open access articles (X_6) refers to the categories Green (published, accepted, submitted), Hybrid, Gold.

The number of pages of the article (X_7), and the length of its reference list (X_8) are included. Concerning the latter, we include also the share of references indexed in WoS (X_9), signalling the

^① The WoS classification scheme involves 255 SCs in all, grouped in 13 macro-areas (in the following referred to as areas). The SC of a publication corresponds to that of the journal where it is published. For publications in multidisciplinary journals (multiple SCs) the scaling factor is calculated as the average of the standardized values for each subject category.

^② The strong correlation between the number of authors and institutions leads to the exclusion of the institutions from the model.

extent of recourse to “qualified” literature. Of WoS-indexed references, we have included also their age (X_{10}), their average normalized impact (X_{11}), measured in the same way as Y , and the share of self-citations (X_{12}).

Finally, the degree of interdisciplinarity of an article (X_{13}), or “*diversification ratio (DR)*”, is the share of cited papers falling in SCs other than the dominant one in the article reference list (Abramo et al., 2017b).

Model (1) includes also ζ_j representing the fixed effect for journal j (with the constraint that the sum of the fixed effects across observations is zero), and e_{ij} representing a random error with zero mean and standard deviation σ . The model is fitted using the *xtreg* command of Stata 17 with the *fe* option (StataCorp., 2021).^③

A preliminary analysis based on local polynomial smoothing shows that for most of the variables, the relationships are linear on the logarithmic scale. For this reason, the variables are inserted into the model after a log transformation, except for dummies (X_3) and variables expressed in percentages (X_6 , X_9). To avoid dropping out articles with zero values, for variables having a minimum of zero (Y , X_7), the log transformation is computed after shifting the variable by a small constant equal to 1/20 of the standard deviation.

The log transformation makes these relationships easier to interpret. This means we analyze the percentage changes rather than absolute differences. For example, a 1% increase in the number of authors would result in a proportional change in citation impact based on the estimated coefficient. This approach is particularly useful when variables have wide-ranging values, such as citation counts or reference numbers, which can vary significantly across articles.

The coefficients in our model are elasticity measures, meaning they show the percentage change in citation impact associated with a 1% change in an explanatory variable. For instance, a coefficient of 0.2 for article length (number of pages) indicates that a 1% increase in the number of pages is associated with a 0.2% increase in citation impact. To illustrate, if an article originally has 10 pages and receives 50 citations, increasing the length to 10.1 pages (a 1% increase) would, on average, lead to an additional 0.1 citations (0.2% of 50).

While the literature supports the inclusion of the above variables, we acknowledge that our model does not account for other potentially influential factors, such as institutional reputation and country of origin, which may also shape citation patterns. Although these variables were not included in the present analysis, their omission could limit the predictive power of our model. Institutional reputation, for instance, is known to influence publication impact, particularly in fields where specific institutions have high recognition and prestige (Sanfilippo et al., 2018). Similarly, the country of origin could play a role in shaping visibility and citation rates, particularly in fields with strong regional biases in publication or citation practices.

The fixed-effects linear regression model used in this study is well-suited for isolating the effects

^③ The regression coefficients $\beta_0 \dots \beta_p$ are estimated by running ordinary least squares on the deviations from the journal means (i.e. the outcome variable is transformed into $Y_{ij} - \bar{Y}_j$, where the second term is mean across articles of journal j , and similarly for the independent variables).

of non-scientific factors at the article level while controlling for time-constant journal-specific effects. This approach enables us to focus on within-journal variations, with no need to explicitly control for the confounding influence of journal characteristics, such as prestige or scope.

The analysis could alternatively exploit multilevel models with random effects for journals and subject areas, but this would raise concerns about the validity of the required assumptions about the random effects, in particular, the independence with the explanatory variables (exogeneity). The main advantage of replacing journal fixed effects with random effects is to allow for the inclusion of journal-level characteristics, such as reputation. This would open an interesting research path, but it is out of the scope of our work; therefore, we prefer to rely on a fixed-effects model that is suitable for our research questions while being valid under much weaker assumptions (e.g. Rabe-Hesketh & Skrondal, 2022).

3.2 Dataset

The analysis focuses on WoS publications of the year 2017 having “article” as the document type. To make the analysis more robust and limit the noise introduced by outliers, we excluded records with extreme values for some features, namely, articles with more than 100 authors, more than 100 pages, or more than 200 references. The dataset includes 1,556,053 articles. We further drop out articles with missing values on the covariates (share of women 4.4%, average quality of the authors 4.6%), with an overall reduction of 8.3%, so the final dataset has 1,427,670 articles. For each article, we attribute the SC of the hosting journal and the relevant area. A breakdown of the final dataset per area is presented in Table 1. The last column, “total” is larger than the number of unique articles since some belong to multiple areas (71.6% one area, 23.8% two areas, 4.4% three areas, and 0.1% four or five areas).

Table 1. Number of articles per area of the journal.

	Area	N. of articles
1	Arts and Humanities	17,816
2	Biology	251,921
3	Biomedical Research	164,805
4	Chemistry	194,405
5	Clinical Medicine	316,828
6	Earth and Space Sciences	116,736
7	Economics	40,308
8	Engineering	349,058
9	Law, political and social sciences	65,344
10	Mathematics	57,718
11	Multidisciplinary Sciences	66,985
12	Physics	227,595
13	Psychology	30,896
	Total	1,900,415

Research Papers

We consider most of the measurable variables affecting future citation rates of manuscripts submitted for publication. Table 2 reports the summary statistics of the variables included in the model. The dummy variables are *English* (at least one author belongs to an English-speaking country), *Women* (at least one author is a woman) and *open access* (mode of publication of the article). The outcome variable *impact* and the independent variables *n. authors*, *n. countries*, *average author quality*, *n. pages*, *n. references*, *age cited articles*, and *impact cited articles* have right-skewed distributions where the maximum is many standard deviations from the mean, which suggests a logarithmic transformation for modeling.

Two independent variables, the percentage of self-citations in the reference list (X_{12}) and the level of interdisciplinarity in the list (X_{13}), are not reported in Table 2 because they contribute negligibly to the model for predicting the outcome.

Table 2. Summary statistics of the variables of the dataset (1,427,670 articles).

	Variable	Mean	Std dev.	Min	Max
Y	impact	0.871	1.992	0	1,039.638
X ₁	n. authors	5.476	3.959	1	100
X ₂	n. countries	1.417	0.894	1	44
X ₃	avg authors quality	1.015	1.053	0.001	204.831
X ₄	Women	0.605	0.489	0	1
X ₅	English	0.395	0.489	0	1
X ₆	open access	0.453	0.498	0	1
X ₇	n. pages	11.000	6.259	1	100
X ₈	n. references	41.317	22.570	1	200
X ₉	% references in WoS	76.874	20.630	0.515	100
X ₁₀	cited articles age	9.431	3.716	0	37
X ₁₁	cited articles impact	11.800	23.462	0.010	1,995.620

4 Results

The fixed effects linear model outlined in Section 3 was fitted first on all articles, and then separately on the articles of each area. The results of the former approach are shown in Table 3, while those of the latter are shown in Table 4, separately by area. The model includes all the explanatory variables listed in Table 2, which are suggested by theoretical considerations and empirical findings.

As for the overall model (Table 3), given the large dataset size, all coefficients are statistically significant at 1/1000. The R squared is 0.26, which means that the considered non-scientific features of an article can predict about one-fourth of the variation of its normalized impact.

Given that the outcome variable is log-transformed, the point estimates of the regression coefficients are to be interpreted in terms of the relative change. For example, the 0.070 estimated coefficient for dummy variable *open access* means that an open access article has on average, 7%

more normalized impact than a non-OA article. This means that, all else being equal, an article published with open access would receive 107 citations for every 100 citations a comparable non-open-access article would receive.

Table 3. Estimates of linear models for the log impact of articles with fixed effects for the journals overall.

	Coef.	Std Err.	t	P>t	[95% Conf. Interval]	
X ₁ (log) n. authors	0.085	0.001	66.1	<0.01	0.082	0.087
X ₂ (log) n. countries	0.041	0.002	24.4	<0.01	0.037	0.044
X ₃ (log) avg authors quality	0.224	0.001	242.3	<0.01	0.222	0.226
X ₄ Women	-0.024	0.001	-17.1	<0.01	-0.026	-0.021
X ₅ English	-0.022	0.002	-15.0	<0.01	-0.025	-0.020
X ₆ open access	0.070	0.002	38.5	<0.01	0.066	0.073
X ₇ (log) n. pages	0.200	0.002	86.8	<0.01	0.195	0.204
X ₈ (log) n. references	0.268	0.002	165.8	<0.01	0.265	0.271
X ₉ % references in WoS	0.002	0.000	39.2	<0.01	0.002	0.002
X ₁₀ (log) cited articles age	-0.374	0.002	-231.6	<0.01	-0.377	-0.370
X ₁₁ (log) cited articles impact	0.061	0.001	97.5	<0.01	0.060	0.063
Intercept	-1.436	0.007	-201.9	<0.01	-1.450	-1.422

Number of articles: 1,427,670; Number of journals: 13,135; R-squared: 0.26; Intra-journal correlation: 0.32.

At the same time, the 0.002 estimated coefficient for the share of references indexed in WoS means that an increase of one percentage point of such share is associated with an average increase of 0.2% in the impact. The other explanatory variables are log-transformed; thus, the regression coefficient is an elasticity measure. For example, an increase of 1% in the number of references is associated with an average increase of 0.268% in the impact. The share of references indexed in WoS has instead a negligible association (+0.002).

We register an impact advantage for publications citing higher-impact articles (+0.061). In addition, the number of authors in the byline (+0.085) and the international character of the co-authorship (number of countries) (+0.041) show a positive association with the article's impact; the same holds true for the article length (number of pages) (+0.200). As expected, we register a very high positive association with impact (+0.224) of the average impact of co-authors' past publications.

A few variables show a negative association with impact:

- The average age of the article's references (-0.374) means that when an article's scientific content relies on old literature, it is less likely to represent a reference for future scientific advances.
- English (-0.022), implying that articles co-authored by scholars from English-speaking countries, all others being equal, have a lower impact.
- Women (-0.024), attesting to a gender disadvantage for articles authored by females.

Table 4. Estimates of linear models for the log impact of the articles with fixed effects for the journals (standard errors in parenthesis)†, by area.

Area#	1	2	3	4	5	6	7	8	9	10	11	12	13
X ₁ (log) n. authors	0.079 (0.018)	0.082 (0.003)	0.072 (0.004)	0.044 (0.003)	0.088 (0.003)	0.105 (0.004)	0.122 (0.009)	0.071 (0.003)	0.069 (0.007)	0.087 (0.009)	0.111 (0.005)	0.084 (0.003)	0.057 (0.009)
X ₂ (log) n. countries	0.109 (0.026)	0.033 (0.004)	0.044 (0.005)	0.009† (0.005)	0.045 (0.003)	0.055 (0.005)	0.075 (0.010)	0.039 (0.004)	0.056 (0.010)	0.112 (0.010)	-0.002† (0.007)	0.034 (0.004)	0.053 (0.012)
X ₃ (log) avg authors quality	0.179 (0.008)	0.235 (0.002)	0.177 (0.003)	0.282 (0.003)	0.157 (0.002)	0.225 (0.003)	0.220 (0.005)	0.242 (0.002)	0.212 (0.004)	0.299 (0.005)	0.233 (0.004)	0.259 (0.002)	0.243 (0.007)
X ₄ Women	0.024† (0.017)	-0.034 (0.003)	-0.016 (0.004)	-0.023 (0.003)	-0.017 (0.003)	-0.053 (0.004)	-0.013† (0.008)	-0.028 (0.003)	0.011† (0.008)	-0.044 (0.008)	-0.007† (0.006)	-0.023 (0.003)	-0.011† (0.011)
X ₅ English	0.020† (0.017)	-0.023 (0.003)	-0.019 (0.004)	-0.044 (0.004)	-0.005† (0.003)	-0.054 (0.005)	-0.038 (0.009)	-0.033 (0.003)	0.017† (0.008)	-0.052 (0.008)	-0.002† (0.006)	-0.014 (0.004)	-0.008† (0.010)
X ₆ open access	0.164 (0.018)	0.070 (0.004)	0.101 (0.006)	0.020 (0.004)	0.114 (0.004)	0.068 (0.006)	0.087 (0.009)	0.037 (0.004)	0.139 (0.008)	0.038 (0.010)	0.087 (0.018)	0.033 (0.004)	0.079 (0.010)
X ₇ (log) n. pages	0.230 (0.026)	0.169 (0.005)	0.230 (0.007)	0.217 (0.006)	0.270 (0.005)	0.205 (0.008)	0.134 (0.015)	0.193 (0.005)	0.218 (0.013)	0.090 (0.010)	0.086 (0.011)	0.221 (0.005)	0.187 (0.018)
X ₈ (log) n. references	0.269 (0.017)	0.277 (0.004)	0.239 (0.005)	0.249 (0.004)	0.243 (0.003)	0.267 (0.006)	0.313 (0.010)	0.268 (0.003)	0.275 (0.008)	0.316 (0.008)	0.287 (0.008)	0.262 (0.004)	0.291 (0.012)
X ₉ % references in WoS	0.007 (0.001)	0.002 (0.000)	0.001 (0.000)	0.000† (0.000)	0.002 (0.000)	0.002 (0.000)	0.002 (0.000)	0.002 (0.000)	0.004 (0.000)	0.002 (0.000)	-0.000† (0.000)	0.000 (0.000)	0.001 (0.000)
X ₁₀ (log) cited articles age	-0.193 (0.013)	-0.408 (0.004)	-0.383 (0.005)	-0.456 (0.004)	-0.319 (0.004)	-0.414 (0.006)	-0.490 (0.011)	-0.386 (0.003)	-0.322 (0.008)	-0.357 (0.009)	-0.429 (0.008)	-0.395 (0.004)	-0.463 (0.015)
X ₁₁ (log) cited articles impact	0.093 (0.008)	0.043 (0.001)	0.070 (0.002)	0.029 (0.001)	0.075 (0.001)	0.090 (0.003)	0.127 (0.005)	0.076 (0.001)	0.096 (0.004)	0.103 (0.004)	0.051 (0.003)	0.045 (0.002)	0.093 (0.005)
Intercept	-2.312 (0.071)	-1.264 (0.017)	-1.377 (0.021)	-0.952 (0.021)	-1.676 (0.015)	-1.409 (0.025)	-1.308 (0.047)	-1.326 (0.015)	-1.821 (0.036)	-1.303 (0.037)	-1.160 (0.038)	-1.269 (0.018)	-1.266 (0.056)
Number of articles	17,816	251,921	164,805	194,405	316,828	116,736	40,308	349,058	65,344	57,718	66,985	227,595	30,896
Number of journals	1,618	2,187	1,187	600	2,954	931	919	2,104	1,800	746	202	864	525
R-squared	0.26	0.27	0.26	0.31	0.24	0.27	0.28	0.28	0.22	0.26	0.24	0.30	0.22
Intra-journal correlation	0.32	0.27	0.29	0.34	0.28	0.24	0.26	0.28	0.25	0.18	0.39	0.26	0.26

1, Arts and Humanities; 2, Biology; 3, Biomedical Research; 4, Chemistry; 5, Clinical Medicine; 6, Earth and Space Sciences; 7, Economics; 8, Engineering; 9, Law, political and social sciences; 10, Mathematics; 11, Multidisciplinary Sciences; 12, Physics; 13, Psychology.

† All coefficients are statistically significant at 1%, except those marked with †.

While these findings provide important insights into citation dynamics, they also point to potential systemic biases in scholarly recognition and citation practices. The gender gap in citations may reflect broader disparities in academic networks, visibility, and resource access rather than differences in the intrinsic quality of research. Previous studies have shown that female scholars often have smaller international collaboration networks, reduced mobility, and limited access to high-prestige publishing venues, all of which may contribute to lower citation rates. Such biases can perpetuate inequities in research evaluations, particularly when citation-based metrics are used in isolation to assess scholarly performance.

Coming to estimates obtained at the area level (Table 4), compared to others in the Arts and Humanities, the impact of articles is much less sensitive to the age of cited articles. This is an area where the coverage of WoS and the intensity of publications are low compared to other areas (Archambault et al., 2006; Hicks, 1999). Furthermore, the pace of discoveries and technological advances is notoriously slower than in STEM. The impact of articles in the Arts and Humanities is much more sensitive than in other areas to the open-access features of publications. Interpretation of these results is complex. A possible explanation is that the countries' contribution to this area has a larger geographical scope than that of STEM. Free access to literature in less affluent countries might partly explain this marked difference. Mathematics is another area with peculiar features. Within STEM, it shows the lowest intensity of publications and collaboration work (D'Angelo & Abramo, 2015), the lowest citation counts for journals, papers, and authors (Adler et al., 2008), and the lowest citation accrual speed (Wang, 2013). Compared to other areas, the associations to the impact of the number of countries in the byline and the reference list length are markedly higher, and vice versa, that of the manuscript length is appreciably lower. Evidently, in mathematics, numbers count more than words; alongside cooperative work involving different countries and works based on more sources contribute more than other features to the manuscript's impact. The average quality of authors' past scientific publications is confirmed to be a meaningful and relevant factor with no exception; the positive association ranges from a minimum in Clinical Medicine (0.157) to a maximum in Mathematics (0.299). On the contrary, "linguistic advantage" shows several exceptions: in particular, in Arts and Humanities and Law, political and social sciences, the coefficient is positive, although not significant. One possible interpretation is that being an English mother tongue does not help in quantitative discipline studies, unlike others. The same holds for gender influence. Anyway, in the other eight areas, the negative association with the impact of female co-authors remains statistically significant. The interpretation here is not straightforward. While there are certain areas of the globe where gender homophily and citation sexism are more pronounced (Wang et al., 2021), they should represent only a slight share of overall world citations. A more plausible explanation is that female scholars, in general, hold lower social capital than males (Rhoten & Pfirman, 2007), as shown by their smaller international collaboration networks (Abramo et al., 2013; Larivière et al., 2011; Uhly et al., 2015), which is probably due to their lower inclination to travel abroad for conferences and research, with some exceptions in specific fields (Zimmer et al., 2006). Their comparatively lower social relationships

at the international level might contribute to lower citation rates, with all others being equal. It is worth noting that there is a slight variation of the R squared across areas (min 0.22 in Psychology, max 0.31 in Chemistry); thus, the predictive ability of the considered explanatory variables is similar.

Lastly, the intra-journal correlation, namely the residual correlation between the impact of two articles published in the same journal, is 0.32 in the whole dataset, indicating the presence of a systematic association between the hosting source and the future impact of publications. We registered relevant variations across areas, with a minimum of 0.18 in Mathematics and a maximum of 0.39 in Multidisciplinary Sciences. The larger correlation in Multidisciplinary Sciences can be partly explained by the higher heterogeneity of its journals, which include such prestigious outlets as *Nature*, *Science*, *PNAS*, but also 33 journals (equal to 16% of the area total) with nihil impact factor. Therefore, systematic differences in the impact of articles published in the same source are more limited in this area with respect to others.

The practical use of the proposed model is to assess the contribution of non-scientific factors to the impact of journal articles after publication to support the automated prediction of future impact. We have evaluated the model's prediction ability through 5-fold cross-validation: the mean absolute prediction error of the impact on the log scale is 0.639, corresponding to 1.895 in the original scale. This is slightly less than the standard deviation of the impact (1.992). In summary, even if the predictive ability is moderate, the model still provides valuable information that can be combined with other sources.

Our analysis highlights pronounced variations in the influence of non-scientific factors on citation outcomes across academic disciplines, as detailed in Table 4. These differences underscore the complexity of modeling citation behavior in a way that is broadly applicable. For example, we observed that open access is significantly more influential in the Arts and Humanities compared to STEMM fields. This could be attributed to the broader geographical and institutional access disparities in these disciplines, where open access enhances visibility and usability, particularly in less-resourced regions.

Conversely, Mathematics exhibits lower sensitivity to the number of authors, reflecting its traditionally individualistic or small-team research culture. This stands in contrast to disciplines like Biology or Engineering, where larger collaborative efforts are more common and strongly associated with increased citation rates. Similarly, the reliance on the number of references is notably more pronounced in Mathematics, likely due to the field's emphasis on theoretical depth and comprehensive bibliographic foundations.

These variations illustrate the necessity of tailoring citation models to the unique characteristics of each field. While our overall model captures key trends, applying it uniformly across all disciplines without adjustment could lead to misinterpretations or less effective predictions. Future research should focus on refining discipline-specific models to address these nuances.

As AI has increasingly become a key tool in research evaluation, integrating insights from explainable AI (XAI) is critical for ensuring transparency and accountability. While our study

employs a regression-based approach to identify and quantify the influence of non-scientific factors on citation impact, XAI techniques could extend this analysis to AI-driven models. For instance, feature importance methods, such as SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-Agnostic Explanations), can elucidate the role of specific predictors in machine learning algorithms, helping to uncover the relationships between variables like open access, number of authors, and article impact (Lundberg & Lee, 2017; Ribeiro et al., 2016).

Such techniques could make AI-driven citation predictions more interpretable, allowing researchers and evaluators to understand how non-scientific factors influence outcomes. This is particularly important when these predictions inform high-stakes decisions, such as funding allocations or career advancements. Incorporating XAI could also help identify and mitigate biases in AI systems, ensuring fairer and more equitable evaluations across disciplines and demographic groups.

5 Conclusions

Knowledge production has socioeconomic value once it is used to further advance knowledge (scholarly impact) and/or improve practices, goods, and services by incorporating new knowledge into products or processes (social impact). In universities and public research institutions, formal and informal mechanisms are at work to incentivize scholars to produce new knowledge and to devote efforts to encode it in written form to make it available to all potential users. The written communication channel far more utilized by scholars for knowledge diffusion is that of scientific journals. Journals' editors recur to peer-reviewers' recommendations to select which manuscripts to publish among those submitted to the journal.

The opportunity cost linked to peer reviewing is notably significant, particularly for esteemed scientists, who must allocate time from their research responsibilities to conduct reviews. In 2020, reviewers worldwide dedicated over 100 million hours solely to peer reviews for journals, equating to more than 15 thousand years (Aczel et al., 2021). The estimated monetary value of reviewers' time investment in the United States exceeded 1.5 billion USD during the same period. These figures likely underestimate the genuine societal cost, as they only address a portion of global journals and solely consider the monetary value of time, overlooking the opportunity costs. Scholars are currently grappling with the escalating demand for reviews from journal editors. Additional reviews for hiring committees, tenure review boards, granting agencies, and research evaluation agencies impose rapidly diminishing marginal benefits and escalating marginal costs on scientists.

Efforts have been ongoing for years to substitute costly peer reviews with bibliometrics in STEMM (Abramo, 2024; Sivertsen, 2017) and, more recently, with AI (Cárdenas, 2023; Kousha & Thelwall, 2024a). The advent of Large Language Models (LLMs), such as ChatGPT, could potentially offer viable alternatives to the labor-intensive peer-review process. Initial attempts are underway to test ChatGPT's ability to assess the quality of journal articles (Thelwall, 2024). The subsequent phase might involve evaluating its capability to predict the scholarly impact of research

works directly, rather than solely assessing their quality.

In theory, a publication should accrue citations based on its scientific content. However, empirical evidence has demonstrated that numerous non-scientific factors are associated with citation outcomes (Kousha & Thelwall, 2024b; Mammola et al., 2022). This should not be surprising, as it is acknowledged that factors beyond the quality of a product or service influence consumers' choices. Similarly, non-scientific factors might influence scholars' utilization of the existing literature. If maximizing research impact is the ultimate goal of policymakers and society (akin to profit maximization for for-profit organizations), then such factors should also play a role in determining the long-term scholarly impact of research output.

The fundamental question driving our work is whether it is feasible to investigate the relative effects of these non-scientific features in predicting the future scholarly impact of journal articles after publication.

The explanatory variables in our inferential model, which aim to predict the future impact of publications, account for approximately 26% of the impact, with slight variations across disciplinary areas. The remaining 74% is presumably influenced by the actual quality of the manuscript, the marketing efforts of the authors, and other non-scientific factors not encompassed in the statistical model we have employed.

Our quantitative analysis provides valuable insights into the influence of non-scientific factors on citation impact. However, it is essential to recognize the limitations of our approach and propose pathways for future improvements.

Notably, our analysis relies on data from Web of Science (WoS), which, despite its robustness, has known limitations. This may underrepresent certain disciplines, languages, or regional research outputs, potentially biasing our findings. Furthermore, i) potential biases may arise due to variations in citation practices, resulting in unequal recognition of certain publications; ii) citation impact requires time, and as such, results are sensitive to the chosen time citation window; iii) various publication types may impact the citation outcome; iv) unjustified self-citation and cross-citation practices can artificially inflate the perceived impact of publications; v) negative citations may occur, especially in the early years of publication; vi) citations do not always certify real use and may not be representative of all use; and vii) results are sensitive to the classification schemes adopted for publications.

Although our dataset spans 13 academic fields, disciplinary differences in citation practices may not be fully captured by the current model. For example, disciplines like the Arts and Humanities often exhibit lower citation intensities and longer citation half-lives compared to STEM fields. Future studies could adopt discipline-specific models to capture broader dimensions of research impact.

Our model does not include certain variables, such as institutional reputation, funding acknowledgments, social media promotion, and country of origin, which could further refine our understanding of citation dynamics. Future research should explore the inclusion of these variables to capture a more complete picture of the factors that influence scholarly recognition. Additionally,

regional differences in publication practices and access to research could further explain some of the unexplained variance in the citation impact.

Our fixed-effects linear model assumes linear relationships between the logarithm of the citation impact and the explanatory variables, with continuous variables log-transformed. This specification, which has been chosen on the basis of fit, represents a type of non-linear relationship. However, real-world relationships may involve more complex non-linearities or interactions between factors that the current model cannot fully capture. Advanced machine learning techniques, such as random forests or neural networks, could help uncover these complex patterns, although this may come at the cost of interpretability.

A further limitation of the fixed-effects model is that, while effective for isolating article-level effects, it cannot include journal-level explanatory variables. This prevents the analysis of the effects of journal characteristics, which is, however, beyond the scope of the current work. This limitation could be overcome by moving from fixed to random effects, which requires additional assumptions. In addition, the disciplines could also be modelled with random effects; however, this approach would only account for the variability of the intercepts across disciplines, whereas it is of interest to study how the regression coefficients change with the discipline, namely, to evaluate interactions between article features, such as number of authors or open access, and the discipline. To this end, we fitted the model separately for each discipline and compared the estimates.

Citations accrue over time, and their patterns may vary depending on the article's age. Our analysis uses a fixed citation window, which may not account for long-term trends or differences in citation dynamics across disciplines. Longitudinal studies or dynamic modeling approaches could address this limitation.

The observed differences in the influence of non-scientific factors across disciplines suggest that universal citation models may overlook important field-specific behaviors. For instance, the heightened impact of open access in Arts and Humanities and the minimal influence of manuscript length in Mathematics point to underlying differences in dissemination practices and research cultures. These findings imply that any automated systems for citation prediction or impact assessment, such as those leveraging AI, must incorporate field-specific adjustments to enhance accuracy and relevance.

By identifying these variations, our study lays the groundwork for developing tailored approaches to research evaluation. Policymakers and evaluators can use these insights to design assessment systems that are equitable and reflective of disciplinary norms. Future studies should investigate the socio-cultural and structural factors driving these trends to further refine the prediction models.

The findings of our study could potentially inform forthcoming assessments of research impact by ChatGPT. Providing associations between non-scientific factors of journal articles and future citations to inform assessments of research impact by ChatGPT could have several implications.

Improved prediction accuracy stands out as the first benefit. By integrating non-scientific

factors into the assessment process, ChatGPT could potentially enhance its ability to forecast the future impact of research articles. Factors such as journal prestige, author reputation, number of authors, and article accessibility may yield valuable insights into citation patterns.

Moreover, incorporating a diverse range of factors into impact assessments could mitigate bias in the evaluation process. ChatGPT could provide a more comprehensive view of research impact by considering non-scientific factors alongside traditional citation metrics, reducing reliance on metrics still susceptible to limitations and assumptions.

Enhanced decision-making emerges as another advantage. Researchers and stakeholders could make more informed decisions about which articles to prioritize or promote based on their predicted future impact. This could lead to more efficient resource allocation and increased visibility for high-impact research.

Additionally, encouraging engagement could foster positive outcomes. Researchers may be incentivized to engage in activities that boost the visibility and impact of their work, such as promoting articles on social media or collaborating with high-profile authors. This could facilitate the broader dissemination of research findings and broader societal impact.

However, alongside these benefits, it is crucial to consider potential drawbacks. There is a risk of misuse, with the possibility that the association between non-scientific factors and future citations could be misinterpreted or exploited. It is essential to communicate the limitations of these associations and emphasize that they are just one aspect of a comprehensive impact assessment framework.

Ethical considerations also come into play. Using non-scientific factors to assess research impact may raise concerns about privacy or biases related to author demographics. It is important to ensure that the use of such factors are transparent, fair, and aligned with ethical standards.

Furthermore, our findings underline the need for careful consideration of systemic biases in citation practices. Metrics that include variables such as gender or institutional affiliation must be used responsibly to avoid perpetuating existing inequities. Future research should explore methods to adjust for these biases, such as developing equity-aware citation models or integrating alternative metrics that better capture the diverse contributions of scholars from underrepresented groups. We can work towards more inclusive and fair systems for evaluating research impact by addressing these challenges.

Undoubtedly, the future of research assessment will witness a gradual replacement of costly peer review with automatic systems such as bibliometrics and LLM (Schulz et al., 2022), which will increasingly integrate. The challenge will be to make these automatic evaluation systems transparent to both those being evaluated and decision-makers, who should also be educated on interpreting tool results. Currently, LLMs appear as both miraculous and opaque to the user.

Integrating LLM tools, such as ChatGPT, into the peer review process holds significant promise, particularly in reducing time and labor costs and improving scalability. However, adopting LLMs in peer reviews must address several critical risks. LLM systems are trained on existing data, which may reflect historical biases in publication practices, citation patterns, and peer review

decisions. For example, the underrepresentation of certain demographics, disciplines, or methodologies in training datasets could lead to biased evaluations. If left unchecked, such systems risk reinforcing systemic inequities rather than mitigating them. LLMs function as “black boxes,” making their decision-making processes difficult to interpret. This lack of transparency could erode trust among researchers, particularly when applied to high-stakes evaluations of complex or interdisciplinary work, where human judgment often involves nuanced understanding and contextual interpretation.

Although LLM systems have the potential to revolutionize peer review, their deployment must be approached cautiously and ethically. By addressing risks such as bias and opacity through robust auditing, explainable models, and human oversight, LLMs can complement human reviewers and enhance the efficiency and fairness of the review process. Future work should prioritize developing hybrid systems that integrate human expertise with AI’s scalability and speed, ensuring that evaluations remain rigorous and equitable across all disciplines.

Future developments may consider integrating XAI techniques with holistic modeling frameworks to make AI-driven citation prediction systems accurate and transparent. Combining insights from bibliometric analyses with AI methods, such as NLP and topic modeling, could provide a comprehensive understanding of citation dynamics. Moreover, applying XAI methods to explain the influence of non-scientific factors in complex machine learning models would bridge the gap between predictive power and interpretability, ensuring these tools are both effective and ethically sound.

Author contributions

Giovanni Abramo (Email: giovanni.abramo@uniroma2.it, ORCID: 0000-0003-0731-3635): Conceptualization (Equal), Investigation (Equal), Supervision (Lead), Validation (Equal), Writing - original draft (Equal), Writing - review & editing (Lead).

Ciriaco Andrea D’Angelo (Email: dangelo@dii.uniroma2.it, ORCID: 0000-0002-6977-6611): Conceptualization (Equal), Data curation (Lead), Investigation (Equal), Methodology (Equal), Visualization (Equal), Writing - original draft (Equal).

Leonardo Grilli (Email: leonardo.grilli@unifi.it, ORCID 0000-0002-3886-7705): Data curation (Equal), Investigation (Equal), Methodology (Lead), Visualization (Equal), Writing - original draft (Equal).

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abramo, G. (2018). Revisiting the scientometric conceptualization of impact and its measurement. *Journal of Informetrics*, 12(3), 590-597.

- Abramo, G. (2024). The forced battle between peer-review and scientometric research assessment: Why the CoARA initiative is unsound. *Research Evaluation*, rvac021, DOI: 10.1093/reseval/rvac021.
- Abramo, G., Cicero, T., & D'Angelo, C.A. (2011). Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics*, 5(4), 659-667.
- Abramo, G., & D'Angelo, C.A. (2015). The relationship between the number of authors of a publication, its citations and the impact factor of the publishing journal: Evidence from Italy. *Journal of Informetrics*, 9(4), 746-761.
- Abramo, G., D'Angelo, C.A., & Di Costa, F. (2016). The effect of a country's name in the title of a publication on its visibility and citability. *Scientometrics*, 109(3), 1895-1909.
- Abramo, G., D'Angelo, C.A., & Di Costa, F. (2017a). Do interdisciplinary research teams deliver higher gains to science? *Scientometrics*, 111(1), 317-336.
- Abramo, G., D'Angelo, C.A., & Di Costa, F. (2017b). Specialization versus diversification in research activities: the extent, intensity and relatedness of field diversification by individual scientists. *Scientometrics*, 112(3), 1403-1418.
- Abramo, G., D'Angelo, C.A., & Felici, G. (2019). Predicting long-term publication impact through a combination of early citations and journal impact factor. *Journal of Informetrics*, 13(1), 32-49.
- Abramo, G., D'Angelo, C.A., & Murgia, G. (2013). Gender differences in research collaboration. *Journal of Informetrics*, 7(4), 811-822. DOI: 10.1016/j.joi.2013.07.002
- Abramo, G., D'Angelo, C.A., & Reale, E. (2019). Peer review vs bibliometrics: Which method better predicts the scholarly impact of publications? *Scientometrics*, 121(1), 537-554.
- Aczel, B., Szaszi, B., & Holcombe, A.O. (2021). A billion-dollar donation: Estimating the cost of researchers' time spent on peer review'. *Research Integrity and Peer Review*, 6, 1-8.
- Adler, R., Ewing, J., & Taylor, P. (2008). *Citation statistics*. International Mathematical Union, in cooperation with the International Council of Industrial and Applied Mathematics and the Institute of Mathematical Statistics. <https://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf>
- Aksnes, D.W., & Taxt, R.E. (2004). Peer reviews and bibliometric indicators: A comparative study at Norwegian University. *Research Evaluation*, 13 (1), 33-41.
- Alimohammadi, D., & Sajjadi, M. (2009). Correlation between references and citations. *Webology*, 6(2), a71.
- Allen, L., Jones, C., Dolby, K., Lynn, D., & Walport, M. (2009). Looking for landmarks: The role of expert review and bibliometric analysis in evaluating scientific publication outputs. *PLoS ONE*, 4(6).
- Alohali, Y.A., Fayed, M.S, Mesallam, T., Abdelsamad, Y., Almuhawes, F., & Hagr, A. (2022). A machine learning model to predict citation counts of scientific papers in otology field. *BioMed Research International*. DOI: 10.1155/2022/2239152
- Ante, L. (2022). The relationship between readability and scientific impact: Evidence from emerging technology discourses. *Journal of Informetrics*, 16(1), 101252. DOI: 10.1016/j.joi.2022.101252
- Antelman, K. (2004). Do open-access articles have a greater research impact?. *College & Research Libraries*, 65(5), 372-382.
- Antoniou, G.A., Antoniou, S.A., Georgakarakos, E.I., Sfyroeras, G.S., & Georgiadis, G.S. (2015). Bibliometric analysis of factors predicting increased citations in the vascular and endovascular literature. *Annals of Vascular Surgery*, 29(2), 286-92.
- Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329-342.
- Baker, M. (2016). Stat-checking software stirs up psychology. *Nature*, 540(7631), 151-152.
- Ball, P. (2008). A longer paper gathers more citations. *Nature*, 455(7211), 274.

- Beranová, L., Joachimiak, M. P., Kliegr, T., Rabby, G., & Sklenák, V. (2022). Why was this cited? Explainable machine learning applied to COVID-19 research literature. *Scientometrics*, 127(5), 2313-2349.
- Bertocchi, G., Gambardella, A., Jappelli, T., Nappi, C. A., & Peracchi, F. (2015). Bibliometric evaluation vs informed peer review: Evidence from Italy. *Research Policy*, 44(2), 451-466.
- Bloor, D. (1976). *Knowledge and Social Imagery*. London: Routledge, Kegan and Paul.
- Bornmann, L. (2013). What is societal impact of research and how can it be assessed? A literature survey. *Journal of the American Society of Information Science and Technology*, 64(2), 217-233.
- Bornmann, L. (2017). Measuring impact in research evaluations: A thorough discussion of methods for, effects of and problems with impact measurements. *Higher Education*, 73(5), 775-787.
- Bornmann, L., & Leydesdorff, L. (2013). The validation of (advanced) bibliometric indicators through peer assessments: A comparative study using data from InCites and F1000. *Journal of Informetrics*, 7(2), 286-291.
- Budtz Pedersen, D., Grønvaad, J. F., & Hvidtfeldt, R. (2020). Methods for mapping the impact of social sciences and humanities - A literature review. *Research Evaluation*, 29, 4-21.
- Calver, M.C., & Bradley, J.S. (2010). Patterns of citations of open access and non-open access conservation biology journal papers and book chapters. *Conservation Biology*, 24(3), 872-80.
- Caputo, A., Manesh, M.F., Farrukh, M., Farzipoor Saen, R., & Randolph-Seng, B. (2022). Editorial: Over a half-century of management decision: a bibliometric overview. *Management Decision*, 60(8), 2129-2147.
- Cárdenas, J. (2023). Inteligencia artificial, investigación y revisión por pares: escenarios futuros y estrategias de acción [Artificial intelligence, research, and peer review: Future scenarios and action strategies]. *Revista Española De Sociología*, 32(4), a184. DOI: 10.22325/fes/res.2023.184
- Caron, E., & van Eck, N. J. (2014). Large scale author name disambiguation using rule-based scoring and clustering. In E. Noyons (Ed.), *19th International Conference on Science and Technology Indicators. "Context counts: Pathways to master big data and little data"* (pp. 79-86). Leiden: CWTS-Leiden University.
- Chen, S., Arsenaault, C., & Larivière, V. (2015). Are top-cited papers more interdisciplinary? *Journal of Informetrics*, 9(4), 1034-1046.
- Cole, S., Cole, J.R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science*, 214/4523, 881-886.
- D'Angelo, C.A., & Abramo, G. (2015). Publication rates in 192 research fields. In A. Salah, Y. Tonta, A.A.A. Salah, C. Sugimoto (Eds), *Proceedings of the 15th International Society of Scientometrics and Informetrics Conference - (ISSI 2015)* (pp. 909-919). Istanbul: Bogazici University Printhouse.
- de Winter, J. (2024). Can ChatGPT be used to predict citation counts, readership, and social media interaction? An exploration among 2222 scientific abstracts. *Scientometrics*, 129, 2469-2487.
- Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Volume 1 (Long and Short Papers), 4171-4186.
- Dickersin, K., Min, Y., & Meinert, C.L. (1992). Factors influencing publication of research results: Follow-up of applications submitted to two institutional review boards. *JAMA*, 267(3), 374-378.
- Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, 7(4), 861-873.
- Elgendi, M. (2019). Characteristics of a highly cited article: A machine learning perspective. *IEEE Access*, 7, 87977-87986.
- Fox, C. W., Paine, C. T., & Sauterey, B. (2016). Citations increase with manuscript length, author number, and references cited in ecology journals. *Ecology and Evolution*, 6(21), 7717-7726.

- Fu, L. D., & Aliferis, C. (2008). Models for predicting and explaining citation count of biomedical articles. In *AMIA Annual symposium proceedings* (Vol. 2008, p. 222). American Medical Informatics Association.
- Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-selected or mandated, open access increases citation impact for higher quality research. *PLoS ONE*, 5(10), e13636.
- Glänzel, W., & de Lange, C. (2002). A distributional approach to multinationality measures of international scientific collaboration. *Scientometrics*, 54, 75-89.
- Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171-193
- Grant, J., Brutscher, P. B., Kirk, S. E., Butler, L., & Wooding, S. (2010). Capturing Research Impacts: A Review of International Practice. Documented Briefing. *Rand Corporation*. www.rand.org/pubs/documented_briefings/DB578.html
- Hanson, M.A., Gómez Barreiro, P., Crosetto, P., & Brockington, D. (2023). The strain on scientific publishing. *arXiv*. DOI: 10.48550/arXiv.2309.15884.
- Heßler, N., & Ziegler, A. (2022). Evidence-based recommendations for increasing the citation frequency of original articles. *Scientometrics*, 127, 3367-3381.
- Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44, 193-215.
- Himani, S., Kumar, M. H., Enduri, M. K., Begum, S. S., Rageswari, G., & Anamalamudi, S. (2022). A comparative study on machine learning based prediction of citations of articles. *Proceedings of the 6th International Conference on Trends in Electronics and Informatics (2022)*, 1819-1824. DOI: 10.1109/ICOEI53556.2022.9777184.
- Hurley, L. A., Ogier, A. L., & Torvik, V. I. (2013). Deconstructing the collaborative impact: Article and author characteristics that influence citation count. *Proceedings of the American Society for Information Science and Technology*, 50(1), 1-10.
- Jiang, J., He, D., & Ni, C. (2013). The correlations between article citation and references' impact measures: What can we learn? *Proceedings of the American society for information science and technology*, 50(1), 1-4. DOI: 10.1002/meet.14505001162
- Kirman, C.R., Simon, T.W., & Hays, S.M. (2019). Science peer review for the 21st century: Assessing scientific consensus for decision-making while managing conflict of interests, reviewer and process bias. *Regulatory Toxicology and Pharmacology*, 103, 73-85.
- Knorr-Cetina, K. D. (1981). *The Manufacture of knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford, UK: Pergamon Press.
- Knorr-Cetina, K. D. (1991). Merton sociology of science: the first and the last sociology of science. *Contemporary Sociology*, 20(4), 522-526.
- Kousha, K., & Thelwall, M. (2024a). Artificial intelligence to support publishing and peer review: A summary and review. *Learned Publishing*, 37(1), 4-12.
- Kousha, K., & Thelwall, M. (2024b). Factors associating with or predicting more cited or higher quality journal articles: An Annual Review of Information Science and Technology (ARIST) paper. *Journal of the Association for Information Science and Technology*, 75(3), 15-44.
- Langham-Putrow, A., Bakker, C., & Riegelman, A. (2021). Is the open access citation advantage real? A systematic review of the citation of open access and subscription-based articles. *PLoS ONE*, 16(6): e0253129. DOI: 10.1371/journal.pone.0253129
- Lansingh, V.C., & Carter, M.J. (2009). Does open access in ophthalmology affect how articles are subsequently cited in research?. *Ophthalmology*, 116(8), 1425-31.

- Larivière, V., & Gingras, Y. (2010). On the relationship between interdisciplinary and scientific impact. *Journal of the American Society for Information Science and Technology*, 61(1), 126-131.
- Larivière, V., Vignola-Gagné, E., Villeneuve, C., Gélinas, P., & Gingras, Y. (2011). Sex differences in research funding, productivity and impact: An analysis of Quebec university professors. *Scientometrics*, 87(3), 483-498. DOI: 10.1007/s11192-011-0369-y
- Latour, B., & Woolgar, S. (1979). *Laboratory Life: The Social Construction of Scientific Facts*. London:Sage.
- Lee, C.J., Sugimoto, C.R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1), 2-17.
- Levitt, J. M., & Thelwall, M. (2008). Is multidisciplinary research more highly cited? A macro-level study. *Journal of the American Society for Information Science and Technology*, 59(12), 1973-1984.
- Liang, W., Zhang, Y., Cao, H., Wang, B., Ding, D., Yang, X., & Zou, J. (2023). Can large language models provide useful feedback on research papers? A large-scale empirical analysis. *arXiv*. <https://arxiv.org/abs/2310.01783>
- Liu, J., Chen, H., Liu, Z., Bu, Y., & Gu, W. (2022). Non-linearity between referencing behavior and citation impact: A large-scale, discipline-level analysis. *Journal of Informetrics*, 16(3), 101318.
- Lundberg, S.M., & Lee, S.I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774.
- Mammola, S., Fontaneto, D., Martínez, A., & Chichorro, F. (2021). Impact of the reference list features on the number of citations. *Scientometrics*, 126(1), 785-799.
- Mammola, S., Piano, E., Doretto, A., Caprio, E., & Chamberlain, D. (2022). Measuring the influence of non-scientific features on citations. *Scientometrics*, 127(7), 4123-4137.
- Memon, A. R. (2020). Similarity and plagiarism in scholarly journal submissions: bringing clarity to the concept for authors, reviewers and editors. *Journal of Korean medical science*, 35(27), e217.
- Merton, R. K. (1973). Priorities in scientific discovery. In R. K. Merton (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 286-324). Chicago: University of Chicago Press.
- Miettinen, R., Tuunainen, J., & Esko, T. (2015). Epistemological, artefactual and interactional-institutional foundations of social impact of academic research. *Minerva*, 53, 257-77.
- Milat, A.J., Bauman, A.E., & Redman, S. (2015). A narrative review of research impact assessment models and methods. *Health Research Policy and Systems*, 13, 18.
- Mulkay, M. (1976). Norms and ideology in science. *Social Science Information*, 15(4-5), 637-656.
- Narin, F., & Whitlow, E.S. (1990). *Measurement of scientific cooperation and co-authorship in CEC-related areas of science* (Vol. 1). Publications Office of the European Union.
- OECD/Eurostat (2018). *Oslo manual 2018: Guidelines for collecting, reporting and using data on innovation (4th ed.)*. The measurement of scientific, technological and innovation activities. Luxembourg: OECD Publishing. DOI: 10.1787/9789264304604-en
- Özkent, Y. (2022). Social media usage to share information in communication journals: An analysis of social media activity and article citations. *PLoS ONE*, 17(2), e0263725. DOI: 10.1371/journal.pone.0263725.
- Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). Assessment, evaluations, and definitions of research impact: A review. *Research Evaluation*, 23(1), 21-32.
- Rabe-Hesketh, S., & Skrondal, A. (2022). *Multilevel and longitudinal modeling using stata (4th ed.)*. College Station, TX: Stata Press.
- Reale, E., Barbara, A., & Costantini, A. (2007). Peer review for the evaluation of academic research: Lessons from the Italian experience. *Research Evaluation*, 16(3), 216-228.
- Rhoten, D., & Pfirman, S. (2007). Women in interdisciplinary science: Exploring preferences and consequences. *Research Policy*, 36(1), 56-75. DOI: 10.1016/j.respol.2006.08.001

- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144.
- Rinia, E.J., van Leeuwen, Th.N., van Vuren, H.G., & van Raan, A.F.J. (1998). Comparative analysis of a set of bibliometric indicators and central peer-review criteria, evaluation of condensed matter physics in the Netherlands. *Research Policy*, 27(1), 95-107.
- Rosenkrantz, A. B., Doshi, A. M., Ginocchio, L. A., & Aphinyanaphongs, Y. (2016). Use of a machine-learning method for predicting highly cited articles within general radiology journals. *Academic Radiology*, 23(12), 1573-1581.
- Rossi, M. J., & Brand, J. C. (2020). Journal article titles impact their citation rates. *Arthroscopy*, 36, 2025-2029.
- Ruan, X., Zhu, Y., Li, J., & Cheng, Y. (2020). Predicting the citation counts of individual papers via a BP neural network. *Journal of Informetrics*, 14(3), 101039.
- Sanfilippo, P., Hewitt, A. W., & Mackey, D. A. (2018). Plurality in multidisciplinary research: multiple institutional affiliations are associated with increased citations. *PeerJ*, 6, e5664. DOI: 10.7717/peerj.5664
- Schroter, S., Weber, W. E. J., Loder, E., Wilkinson, J., & Kirkham, J. J. (2022). Evaluation of editors' abilities to predict the citation potential of research manuscripts submitted to the BMJ: A cohort study. *British Medical Journal*, 379. DOI: 10.1136/bmj-2022-073880.
- Schulz, R., Barnett, A., Bernard, R., Brown, N. J., Byrne, J. A., Eckmann, P., ... & Weissgerber, T. L. (2022). Is the future of peer review automated?. *BMC Research Notes*, 15(1), 203. DOI: 10.1186/s13104-022-06080-6
- Sivadas, E., & Johnson, M.S. (2015). Relationships between article references and subsequent citations of marketing journal articles. In *Revolution in marketing: Market driving changes: Proceedings of the 2006 Academy of Marketing Science (AMS) Annual Conference* (pp. 199-205). Cham: Springer International Publishing.
- Sivertsen, G. (2017). Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective. *Palgrave Communications*, 3(1), 1-6.
- Smit, J. P., & Hessels, L. K. (2021). The production of scientific and societal value in research evaluation: A review of societal impact assessment methods. *Research Evaluation*, 30(3), 323-335.
- StataCorp. (2021). Stata: Release 17 [Statistical software]. College Station, TX: StataCorp LLC.
- Stremersch, S., Camacho, N., Vanneste, S., & Verniers, I. (2015). Unraveling scientific impact: Citation types in marketing journals. *International Journal of Research in Marketing*, 32(1), 64-77.
- Tahamtan, I., Afshar, A.S., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, 107(3), 1195-1225.
- Talaat, F.M., & Gamel, S.A. (2023). Predicting the impact of no. of authors on no. of citations of research publications based on neural networks. *Journal of Ambient Intelligence and Humanized Computing*, 14, 8499-8508. DOI: 10.1007/s12652-022-03882-1
- Thelwall, M. (2024). Can ChatGPT evaluate research quality? *Journal of Data and Information Science*, 9(2), 1-21. DOI: 10.2478/jdis-2024-0013
- Thelwall, M., Kousha, K., Abdoli, M., Stuart, E., Makita, M., Wilson, P., & Levitt, J. (2023). Why are co-authored academic articles more cited: Higher quality or larger audience? *Journal of the Association for Information Science and Technology*, 74(7), 791-810. DOI: 10.1002/asi.24755
- Thelwall, M., Kousha, K., Stuart, E., Makita, M., Abdoli, M., Wilson, P., & Levitt, J.M. (2023). Does the perceived quality of interdisciplinary research vary between fields? *Journal of Documentation*, 79(6), 1514-1531. DOI: 10.1108/JD-01-2023-0012
- Traag, V.A. (2021). Inferring the causal effect of journals on citations. *Quantitative Science Studies*, 2(2), 496-504.

- Uhly, K. M., Visser, L. M., & Zippel, K. S. (2015). Gendered patterns in international research collaborations in academia. *Studies in Higher Education*, 42(4), 760-782. DOI: 10.1080/03075079.2015.1072151
- van Lent, M., Overbeke, J., & Out, H.J. (2014). Role of editorial and peer review processes in publication bias: analysis of drug trials submitted to eight medical journals. *PLoS ONE*, 9(8), e104846.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365-391. DOI: 10.1016/j.joi.2016.02.007
- Waltman, L., Kaltenbrunner, W., Pinfield, S., & Woods, H.B. (2023). How to improve scientific peer review: Four schools of thought. *Learned Publishing*, 36(3), 334-347.
- Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, 94(3), 851-872. DOI: 10.1007/s11192-012-0775-9
- Wang, D., Song, C., & Barabási, A. (2013). Quantifying long-term scientific impact. *Science*, 342(6154), 127-132. DOI: 10.1126/science.1237825
- Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PLoS ONE*, 10(5), e0127298.
- Wang, X., Dworkin, J.D., Zhou, D., Stiso, J., Falk, E.B., Bassett, D.S., & Lydon-Staley, D.M. (2021). Gendered citation practices in the field of communication. *Annals of the International Communication Association*, 45(2), 134-153.
- Wang, X., Liu, C., Mao, W., & Fang, Z. (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics*, 103(2), 555-564. DOI: 10.1007/s11192-015-1547-0
- Wilsdon, J. (2016). *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management*. London: Sage Publications, Ltd.
- Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122-1136.
- Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036-1039.
- Xie, J., Gong, K., Cheng, Y., & Ke, Q. (2019). The correlation between paper length and citations: A meta-analysis. *Scientometrics*, 118(3), 763-786.
- Yegros-Yegros, A., Rafols, I., & D'Este, P. (2015). Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. *PLoS ONE*, 10(8). DOI: 10.1371/journal.pone.0135095
- Yu, X., Meng, Z., Qin, D., Shen, C., & Hua, F. (2022). The long-term influence of open access on the scientific and social impact of dental journal articles: An updated analysis. *Journal of Dentistry*, 119, 104067. DOI: 10.1016/j.jdent.2022.104067.
- Zhao, X., & Zhang, Y. (2022). Reviewer assignment algorithms for peer review automation: A survey. *Information Processing & Management*, 59(5), 103028.
- Zimmer, A., Krimmer, H., & Stallmann, F. (2006). Winners among losers: Zur feminisierung der Deutschen universitäten [Winners among losers: On the feminization of German universities]. *Beiträge zur Hochschulforschung*, 28(4), 30-56.

