

Rethinking scholar research evaluation in the context of open science: An exploratory case study

Cuizhong Guan^{1✉}, Qian Li¹, Ziyi Zhu²

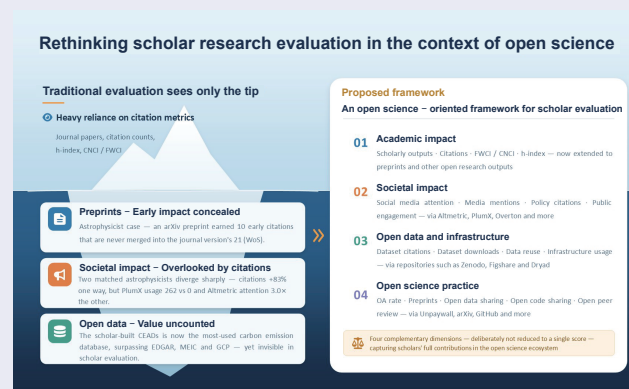
Cite this article: Guan CZ, Li Q, Zhu ZY. *iLibrary* 2026, 1(1): 9690006. <https://doi.org/10.26599/ILIB.2026.9690006>

ABSTRACT: The current research evaluation system is inherently contradictory to the principles of open science. Existing evaluation frameworks remain heavily reliant on citation-based metrics, which often fail to fully recognize the broader value generated by preprints, social dissemination, and open data sharing. This study employs an exploratory case study approach to examine the preprint impact and societal impact of scholarly outputs among astrophysics researchers and to analyze the usage statistics of open data platforms among scholars in the Earth sciences. The findings reveal that traditional research evaluation systems significantly underestimate the early academic impact of preprints, inadequately capture scholars' societal influence, and fail to effectively recognize the value created by open data infrastructure. To address these limitations, this paper proposes an open science-oriented scholar evaluation framework that integrates academic impact, societal impact, open data/infrastructure, and open science practice into a unified assessment system. Through scholarly case studies, this paper provides empirical evidence and offers practical insights for reforming research evaluation systems in the era of open science.

KEYWORDS: open access; open data; preprint; research evaluation; Altmetrics

1 Introduction

Research evaluation is a crucial component of scientific research management and decision-making and plays a vital role in the optimal allocation of national science and education resources, and its research orientation and value leadership merit attention (Cramer and Reinhart, 2026; Lauronen, 2020; Molas-Gallart, 2012). A scientifically sound and rational research evaluation system can stimulate innovation and guide researchers to conduct high-quality scientific research, thereby promoting the advancement of scientific knowledge and sustained progress in science and technology (Groen-Xu et al., 2023). The recognition and impact of academic achievements, as well as researchers' scientific behaviors and directions, are directly influenced by the scientific validity and fairness of evaluation systems. Currently, international scholar evaluation systems include ranking-based systems such as Clarivate's Highly Cited Researchers, Elsevier's Highly Cited Chinese Researchers, and Stanford University's Top 2% Scientists; indicator-based systems such as the h-index; and database profiling systems based on InCites and SciVal, including



highly cited papers, subject rankings, international collaboration, and normalized citation impact indicators (CNCI/FWCI) (Chaignon, 2025; Hirsch, 2005; Ioannidis et al., 2020; Lanchobarrantes and Cantú-Ortiz, 2019). These systems are built upon citation data, possessing advantages in objectivity and quantifiability. However, they also have obvious limitations: single-dimensional academic value measurement, difficulty in reflecting originality and true contributions, and time lag and Matthew effects that disadvantage emerging directions and young scholars.

Open science represents a new paradigm for global scientific research, aiming to promote the free, open, cooperative, and shared nature of scientific knowledge and improve the quality, efficiency, and impact of scientific research (OECD, 2015; UNESCO, 2021). With the continuous advancement of the open science movement, an increasing number of research outputs are being published in open access formats. According to Science Citation Index Expanded (SCIE) data, open access papers accounted for 46% of total global papers in 2024. Open science has transformed the production methods, communication and exchange, and transformation mechanisms of scientific research

¹ Tsinghua University Library, Tsinghua University, Beijing 100084, China. ² School of Journalism and Communication, Tsinghua University, Beijing 100084, China.

✉ Corresponding author. E-mail: guanczh@tsinghua.edu.cn

Received: April 24, 2026; Revised: June 23, 2026; Accepted: August 11, 2026

© The Author(s) 2026. This is an open access article under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0, <http://creativecommons.org/licenses/by/4.0/>).

results. Meanwhile, the deficiencies of traditional research evaluation systems have become increasingly apparent in the open science environment (Nosek et al., 2015).

In the era of open science, discussions on research evaluation reform have increasingly emphasized the principles of responsible evaluation and multidimensional assessment. The San Francisco Declaration on Research Assessment (DORA) calls for eliminating the overreliance on journal impact factors. It stresses that evaluation must be based on the intrinsic value of research itself (American Society for Cell Biology, 2013). The Leiden Manifesto proposes ten principles for the responsible use of research metrics. It opposes the misuse of single indicators such as the impact factor and h-index. It also advocates for transparency in metrics and respect for disciplinary diversity (Hicks et al., 2015). The narrative CVs allow applicants to present a more comprehensive account of their research contributions and career development (Fritch et al., 2021). These initiatives, from various perspectives, advocate for reducing excessive dependence on journal-level papers and citation metrics. They place greater emphasis on the diverse value of research outputs and the comprehensive contributions of scholars.

Beyond discussions at the policy initiative level, some scholars have begun to empirically examine the limitations of traditional research evaluation systems in the context of open science. Kousha et al. (2018) pointed out that traditional academic databases such as Web of Science (WoS) and Scopus inadequately capture early citations of papers, thereby underestimating the impact of preprints and early access papers. Bornmann et al. (2019) argued that research evaluation should not focus solely on academic citations but should also attend to the social dissemination and societal impact of research outputs. Colavizza et al. (2020) contended that the value of open data resources should be recognized in research evaluation. However, current studies mainly focus on single dimensions, such as preprints, societal impact, or open data. Comprehensive empirical research on scholar-level evaluation in open science remains scarce.

Therefore, this paper provides quantitative evidence through case studies of scholars from three dimensions: the obscured value of preprints, the inadequate assessment of scholars' multidimensional impact, and the measurement blind spots of open data resources. By analyzing cases from astrophysics and Earth sciences, we reveal the limitations of traditional scholar evaluation systems in the context of open science. Based on these findings, we propose an open science-oriented evaluation framework. This framework provides references for improving research evaluation in the open science era.

2 Reconsidering the validity of current scholar evaluation in the context of open science

2.1 Value concealment of preprints

To reduce publication delays caused by journal publishing and conference organization and enable rapid visibility of innovative research to the public, scientists in fields such as physics, mathematics, and computer science have chosen to publicly disclose their academic achievements on preprint platforms in advance, with formal publication of these results potentially occurring months or even longer afterward. Since American physicist Paul Ginsparg established arXiv in 1991, preprint platforms targeting different disciplinary fields have continuously emerged, such as bioRxiv in biosciences, medRxiv in medicine, and ChemRxiv in chemistry (Ginsparg, 2021). These platforms accelerate the public dissemination of research results, enabling

them to quickly attract peer attention and discussion. Their free and open characteristics also allow more people to access these results, further enhancing their impact (Fu and Hughey, 2019). Although not formal publications in the traditional sense, they may still be cited by other publications, which increases researchers' academic and social influence.

However, in current databases such as WoS and Scopus, impact data for preprints and other nonformally published versions are still counted separately from formally published versions, with the impact gained by nonformally published versions not being merged into their subsequent formally published versions. Larivière et al. (2014) reported that the publication rate of arXiv preprints is stable at approximately 73%, whereas Abdill and Blekhan (2019) reported that bioRxiv preprints have a publication rate of approximately 67%, indicating that approximately 30% of preprints are never formally published. Fraser et al. (2020) found that in the first few months after submission and storage, preprints that are subsequently not formally published receive no fewer citations than those that are formally published; even without formal publication, these preprints continue to receive citations months later, and citing non-peer-reviewed preprints has become widely accepted in biosciences. According to data from the "Report on the Output Status of Preprints in the Open Science Environment 2024" released by the Institute of Scientific and Technical Information of China, in 2023, the total volume of preprints published on major international platforms was 260,000, with 71,000 citations from international papers in the SCI, SSCI, CPCI, BKCI, and ESCI databases and 263,000 citations from other preprints (Wang et al., 2024).

This study adopts theoretical sampling from astrophysics—a field with a mature preprint ecosystem and long-term, stable data accumulation—as an empirical analysis scenario. Since establishing its section on the arXiv platform in 1992, this field has formed a relatively complete academic exchange and citation system, providing a data foundation for analyzing the relationship between preprints and academic impact. Here, we select professor Zheng Cai from the Department of Astronomy at Tsinghua University as a case study. He has made notable contributions in early galaxy cosmology and actively published scientific research results on preprint platforms. For example, on August 17, 2022, Zheng Cai's group submitted a research achievement titled "The Identification of a Dusty Multiarm Spiral Galaxy at $z = 3.06$ with JWST and ALMA" on the arXiv platform (Wu et al., 2022). This achievement was submitted to the international authoritative journal *Astrophysical Journal Letters* on October 4, 2022 and was published online on December 27, 2022 (Wu et al., 2023). As shown in Fig. 1, by 2024, the *Astrophysical Journal Letters* version of this paper was cited 21 times in the WoS Core Collection, while the arXiv version was cited 10 times in the WoS Core Collection. Citations of the arXiv version occurred mainly in the early period,

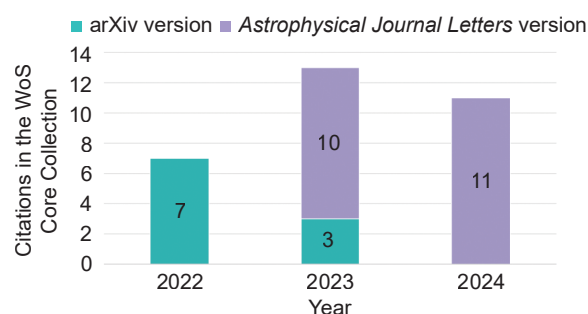


Fig. 1 Citation status of Zheng Cai's arXiv version and journal version in WoS.

with 7 in 2022 and 3 in 2023. After the paper was formally published, 10 papers cited it in 2023, and all citations of the paper in 2024 came from the formally published version.

Zheng Cai's case clearly demonstrates the value of preprints in accelerating research dissemination and obtaining early citations. However, in accordance with current statistical standards of the WoS or Scopus databases, citations of preprint versions are listed separately from those of formally published versions and are not included in the citation counts of the formally published versions. This treatment is not technically impossible to integrate but rather reflects differences in the recognition and application of preprint values across different evaluation systems. To some extent, this approach underestimates the true impact achieved during the early dissemination stage and may consequently underestimate scholars' overall academic contributions. Given that preprints significantly influence peer research and that formally published papers typically do not cite both the preprint and formally published versions of a reference simultaneously, incorporating the citation situations of papers on preprint platforms into the statistical system of formally published versions is necessary. This would help in comprehensively and accurately assessing the overall academic impact of papers.

2.2 Assessment deficiencies of scholars' multidimensional impact

With the rapid development of internet technology and the expansion of the open access movement, academic achievements have been widely disseminated and shared, and exchange forms have become more diversified. Social media and online platforms have become crucial channels for researchers to disseminate academic achievements and viewpoints beyond academic journals. As a result, traditional literature citation-based academic value measurement methods can no longer truly and comprehensively reflect the comprehensive societal impact of the academic achievements of institutions or individual researchers. In this context, Altmetrics have emerged as a supplementary tool to traditional citation-centric research evaluation systems (Amath et al., 2017; Zhao et al., 2021). They break through one-sided evaluation methods that rely solely on citations, providing broader and more diverse impact assessments.

Altmetrics include not only traditional citation data but also nontraditional citation indicators such as social media mentions, policy document citations, and online discussion activity, thereby providing a more comprehensive and dynamic picture of research impact for researchers and academic institutions. Currently, major Altmetrics application tools include PlumX (owned by Elsevier)

and Altmetric (owned by Digital Science). PlumX tracks public interactions, sharing, and communication behaviors with academic achievements and conducts measurement analysis from five dimensions: mentions, usage, citations, captures, and social media (Elsevier, 2025). Altmetric focuses on 12 categories of sources, each containing one or more data sources, including public policy documents, online reference managers, social media, blogs, Wikipedia, multimedia, and other online platforms, and reflects the relative impact of each attention source on the basis of weighted composite scores (Altmetric, 2025).

Integrating citation impact indicators and Altmetrics indicators can more comprehensively evaluate the impact of research achievements. To ensure comparability, this study adopts a theoretical sampling approach and selects scholars in astrophysics, a field with a mature preprint culture. Since the 1990s, astrophysics has developed an academic communication system centered on arXiv. Scholars in this field commonly share preprints before formal journal publication. This makes astrophysics well suited for examining the relationship between societal impact and traditional academic impact. Following the principle of the "most similar system design" proposed by Adam Przeworski and Henry Teune, we select Zheng Cai and Shude Mao as case studies. Both work in the same astrophysics subfield and have similar annual publication outputs. Using WoS, PlumX, and Altmetric indicators, this study conducts a multidimensional analysis of the impact of their published work. As shown in Table 1, impact is divided into two major categories: academic impact and societal impact. Academic impact indicators include citations, CNCI, PlumX citations, and PlumX captures; societal impact indicators include PlumX usage, PlumX mentions, PlumX social media, and Altmetric attention. All data were collected up to July 10, 2025.

After reviewing the SCIE papers published by both authors from 2022 to 2024 in the WoS database, we observed that Zheng Cai has 63 papers and Shude Mao has 64 papers. As shown in Fig. 2, Zheng Cai performs outstandingly on academic impact indicators. His citations are 83% higher than those of Shude Mao, CNCI values are 154% higher, and both PlumX citations (32% higher) and captures (68% higher) demonstrate the deep academic penetration of Zheng Cai's research achievements. In contrast, Shude Mao performs better on societal impact indicators. His PlumX usage count is 262 times, compared with 0 for Zheng Cai; PlumX mention counts are 3.1 times as high as those of Zheng Cai; and PlumX social media discussions total 37, compared with 1 for Zheng Cai. Additionally, the Altmetric attention score is 3.0 times as high as that of Zheng Cai. This, to some extent, indicates that Zheng Cai's achievements have increased in impact at the academic level, whereas Shude Mao's achievements have attracted increased attention at the social level.

Table 1 Metrics for academic impact and societal impact

Category	Metric	Source(s)
Number of publications	Number of journal articles	Science Citation Index Expanded (SCIE)
	Citations	WoS Core Collection
	CNCI	WoS Core Collection
Academic impact	PlumX citations	Citations (CSCD, CrossRef, PubMed, SciELO, Scopus, SSRN, patents, policy)
	PlumX captures	Readers, exports/saves
	PlumX usage	Abstract views (Airiti iRead eBooks, Digital Commons, SciELO, SSRN); downloads (Airiti iRead eBooks, Airiti Library, Digital Commons, Institutional Repositories, Mendeley Data, Pure, RePEC, SSRN); full text views (Airiti iRead eBooks, SciELO); views (Mendeley data)
	PlumX mentions	Blog mentions, news mentions, references (Wikipedia)
Societal impact	PlumX social media	Facebook shares, likes & comments; SSRN ratings
	Altmetric attention	Public policy documents, reference managers, social media, blogs, Wikipedia, multimedia, and other online platforms

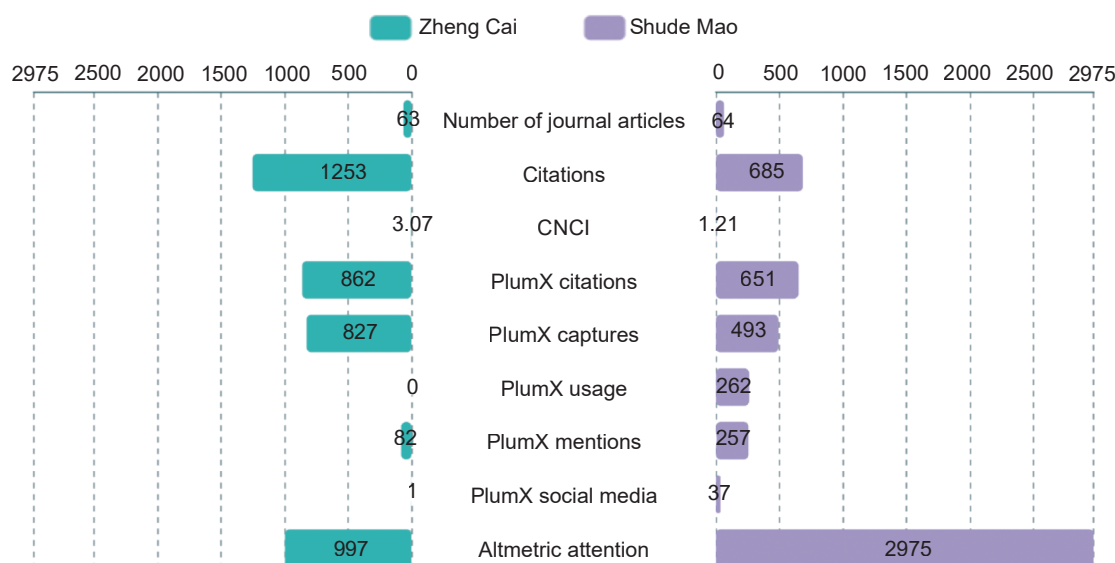


Fig. 2 Comparison of the impact of academic papers published by Zheng Cai and Shude Mao from 2022 to 2024.

Scholars play a key role in scientific research, and the multidimensional assessment of their impact has become a critical issue in the open science era. The cases of Zheng Cai and Shude Mao demonstrate substantial disparities between their academic and societal impact. Relying on any single impact metric to evaluate research outputs would readily introduce estimation bias, thereby undervaluing scholars' genuine contributions in certain dimensions and highlighting the need for multidimensional assessment. Özkent (2022) reported a positive correlation between social media exposure and article citations in the field of communication. Erdt et al. (2016) found that the overall correlation between Altmetrics and citation metrics is weak. Stephen and Stahlschmidt (2024) further noted that among the various measurement dimensions of Altmetrics, Mendeley, usage metrics, ResearchGate, Twitter, and peer ratings have stronger correlations with citations, whereas Facebook, Wikipedia, Google+, blogs, news, and Reddit have relatively weaker correlations with citations. These studies all indicate that Altmetrics provides an indispensable supplementary perspective for assessing scholars' multidimensional impact and further corroborate the necessity of using Altmetrics to complement, rather than replace, citation impact indicators.

2.3 Measurement blind spot of open data resources

Open science encompasses traditional academic outputs, such as journal publications and research data, metadata, open educational resources, software, and source code. In this context, open data typically follow the FAIR principles, allowing anyone to freely access, reuse, preserve, and share data with proper attribution. In the current big data era, massive data resources in fields such as healthcare, industry, and education are being actively mined and applied. Earth sciences are among the most active fields in open data practices. Many studies in this field rely on shared databases for analysis. For example, climate change research widely uses NASA EarthData and NOAA climate databases. Remote sensing studies depend on open satellite data from Landsat and Sentinel. Carbon emission research also relies

heavily on databases such as CEADs and EDGAR. These characteristics make Earth sciences an ideal field for examining how contributions to open data resources can be reflected in research evaluation.

The Carbon Emission Accounts and Datasets (CEADs) were created by professor Dabo Guan's team in 2016. It contains multiscale carbon accounting inventories for China and other developing economies, providing free data download services to scholars worldwide. Owing to the continuous increase in greenhouse gas emissions such as carbon dioxide, humanity faces the enormous challenge of global climate warming. According to statistical data from the BP Statistical Review of World Energy 2024, the global average temperature rise is approaching 1.5 °C, and energy-related greenhouse gas emissions exceeded 40 billion tons for the first time in 2023 (BP, 2024). Carbon emissions accounting is important for accurately understanding carbon emission trends, formulating emission reduction policies for various countries, and achieving carbon-neutral goals. Globally, other authoritative carbon emissions databases include the Emissions Database for Global Atmospheric Research (EDGAR), the Multiresolution Emission Inventory for China (MEIC), and the Global Carbon Project (GCP).

These databases have become important data infrastructures for global climate change research, carbon emission accounting, and carbon neutrality policymaking. A large number of academic papers, industry reports, and assessments by international organizations rely on these shared data resources. However, the impact of such data resources remains difficult to measure effectively. As a result, the actual value of open data resources is often underestimated in scholarly evaluations. We used bibliometric methods to compile statistics on the use of various carbon emissions databases in domestic and international academic achievements. The data source for domestic usage statistics was the China National Knowledge Infrastructure (CNKI) Chinese database, while the foreign data source was Elsevier ScienceDirect. The method involved searching for various spelling variants related to carbon emissions databases in full texts, with the search conducted on July 10, 2025. For CEADs, the search formula in the CNKI Chinese database is as follows: Full Text = (中国碳排放数据库 + 中国碳排放核算数据库 + CEADs 数据库 + CEAD 数据库) OR Full Text = (CEADs + CEAD) AND Subject = (碳排放 + 碳减排 + 碳中和)^①. The search formula in Elsevier SD is as follows: ALL = ("China Emission

① For readers unfamiliar with Chinese, the search query used in the China National Knowledge Infrastructure (CNKI) database translates to: Full Text = (China Emission Accounts and Dataset + Carbon Emission Accounts and Dataset + CEADs Dataset + CEAD Dataset) OR Full Text = (CEADs + CEAD) AND Subject = (Carbon Emissions + Carbon Emission Reduction + Carbon Neutrality).

Accounts and Dataset” OR “Carbon Emission Accounts and Dataset” OR “China Emission Accounts & Dataset” OR “CEADs database” OR “CEADs dataset”).

The statistical results are shown in Fig. 3, which indicates that since its establishment in 2016, the influence of CEADs has been increasing annually. It has now become the most frequently used carbon emissions database, surpassing the other three databases, EDGAR, MEIC, and GCP, especially in the usage of domestic academic achievements. We believe that when professor Dabo Guan’s personal impact is evaluated, the influence of CEADs should be included in the statistics. Scholars’ contributions to open data platforms and open source software platforms should be included in the personal evaluation system. When scholars’ personal achievements are being evaluated, we should consider their impact and contributions to open data resource construction.

3 Reflections on research evaluation in the context of open science

In the open science era, the reform of research evaluation systems is necessary and urgent and requires joint efforts from researchers, research institutions, government departments, and the general public. Traditional evaluation systems have deficiencies and limitations in measuring the breadth and depth of research achievements, scholars’ innovation capabilities, and social contributions, especially in terms of preprint impact, scholars’ societal impact, and contributions to open data resource construction, which have not been completely reflected in current research evaluation. Therefore, research evaluation systems need to keep pace with the times and make corresponding updates and adjustments in this open science era.

3.1 Recognizing the value of preprints and open data resources

The rise of preprint platforms such as arXiv, bioRxiv, and ChinaXiv has greatly shortened the dissemination cycle of research achievements, enabling research to become known and discussed by peers and the public more rapidly. These nonformally published versions accelerate knowledge dissemination and promote early feedback and cooperation in research work, playing a crucial role in driving research progress. Therefore, incorporating preprint impact into evaluation systems is important for comprehensively and fairly evaluating research activities. For this purpose, when counting paper citation frequencies, the citation situations of not only formally published versions in journals but also nonformally published versions, such as preprints, must be considered. Integrating citation data from preprint platforms and formally published versions can more accurately reflect the overall impact of papers and scholars’ academic contributions.

Open science has promoted open access and sharing of scientific data, enabling more researchers to use these data for in-depth research. Contributions from open data platforms, such as the Global Initiative on Sharing All Influenza Data (GISAID) and Carbon Emission Accounts and Datasets (CEADs), provide important data support for scientific research to promote interdisciplinary cooperation and knowledge innovation. Therefore, when the academic achievements of scholars are being evaluated, the contributions on open data platforms should be considered. This can be comprehensively measured by counting the publication records of datasets, documents, and other resources and their usage situations, such as downloads, citations, and other indicators. These indicators can reflect the contributions and impact of open data resources, offering new perspectives for evaluating scholars’ innovation capabilities and social contributions.

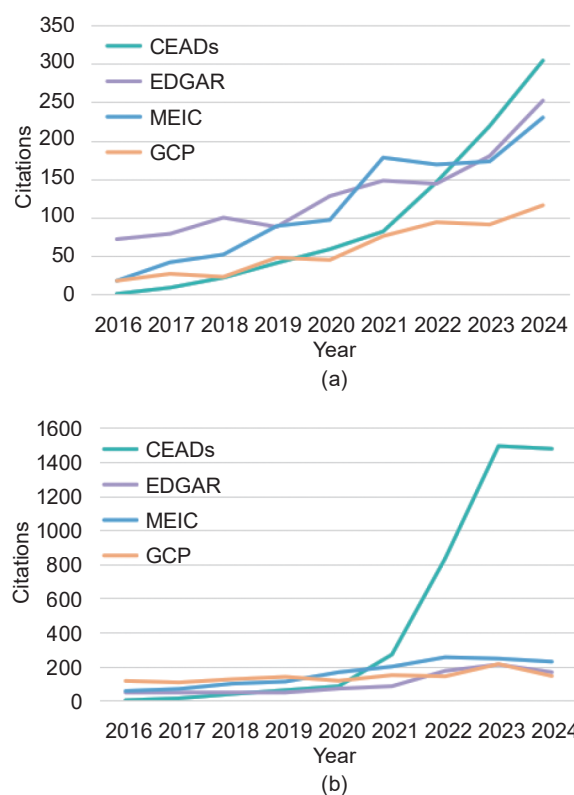


Fig. 3 Citation status of global carbon emissions databases. (a) Citation status in the CNKI Chinese citation database; (b) citation status in the Elsevier ScienceDirect database.

3.2 Taking Altmetrics seriously but cautiously

In the open science environment, the dissemination and exchange of academic achievements are no longer confined to traditional academic journals and conferences. Social networking sites, blogs, online forums, and other social media platforms have also become major channels for the dissemination of academic achievements. Behaviors such as mentions, shares, and discussions on these platforms reflect not only social attention to academic achievements but also scholars’ influence. Altmetrics combine traditional citation indicators with measures of societal impact. This provides a more comprehensive view of the multidimensional value and overall impact of scholarly outputs. At the same time, the limitations of Altmetrics should not be overlooked. These include platform bias, language bias, manipulation risks, disciplinary differences, and the distinction between online attention and genuine societal impact.

First, Altmetric data show clear platform and language biases. Most Altmetric data come from English-language online platforms such as Twitter/X, Facebook, Mendeley, and blogs, while coverage of non-English sources remains limited. This places non-English-speaking scholarly communities at a disadvantage in research evaluation. Second, Altmetric indicators can be artificially inflated through media exposure, institutional promotion, or even fake accounts, making them vulnerable to manipulation. Third, public attention varies significantly across disciplines. Fields such as medicine, public health, and environmental science often receive greater social media attention, whereas mathematics, engineering, and some areas of the humanities and social sciences receive much less attention. Finally, online attention does not necessarily reflect genuine societal impact. High social media attention may simply reflect the novelty, controversy, or public interest of a study. It does not

necessarily indicate substantive contributions to technological innovation or social development. Therefore, Altmetric indicators are better suited for measuring the online visibility of research outputs and the level of public engagement. They should not be treated as direct measures of societal impact.

3.3 An open science-oriented framework for scholar research evaluation

The case studies show clear limitations in existing scholarly research evaluation systems in the era of open science. Based on the findings of this study and related research on responsible evaluation and open science, we propose an open science-oriented framework for scholar research evaluation, as shown in Table 2. The framework includes four dimensions: academic impact, societal impact, open data and infrastructure, and open science practice.

Academic impact: Academic impact remains a core dimension of scholarly research evaluation. It primarily reflects a scholar's research contributions and academic recognition. It is typically measured using indicators such as scholarly outputs, citations, FWCI/CNCI, and the h-index. However, in the era of open science, the scope of academic impact is expanding. Evaluation should not be limited to formally published journal articles. The impact of preprints and other open research outputs should also be considered to provide a more comprehensive assessment of scholars' academic contributions.

Societal impact: Societal impact reflects how widely scholars' research outputs spread and how they influence society. With the growth of open access and internet technologies, research outputs can reach policymakers, industry, the media, and the public more quickly. Indicators such as social media attention, media mentions, policy citations, and public engagement help capture the role of research in knowledge sharing, public discussion, and policymaking.

Open data and infrastructure: In data-intensive disciplines, datasets, databases, software, code repositories, and computing platforms have become important forms of research output. Many research activities rely heavily on these shared resources. However, their value is often not adequately captured in traditional evaluation systems. Indicators such as dataset citations, downloads, reuse frequency, and infrastructure usage can help provide a more comprehensive assessment of scholars' contributions to the development of shared research resources.

Open science practice: This dimension focuses on the openness of scholars' research processes. It includes practices such as OA rate, preprint, open data sharing, open code sharing, and open peer review. Unlike the previous three dimensions, which focus on research outputs and their impacts, this dimension emphasizes the research process itself. It reflects scholars' efforts to improve research transparency, reproducibility, accessibility, and collaboration. It also shows how actively they engage with the

principles of open science.

Notably, these four dimensions should not be reduced to a single composite score. Instead, they should be viewed as complementary dimensions of scholars' research contributions. Compared with traditional evaluation systems, this framework moves beyond a publication-centered approach and captures broader contributions within the open science ecosystem. It distinguishes academic impact from societal impact while also recognizing contributions to data resources and open science practice. This provides research institutions with a more comprehensive and actionable framework for reforming scholar evaluation in the era of open science.

4 Conclusions

This study re-examines the limitations of traditional scholarly research evaluation systems in the era of open science through three exploratory case studies. The results show that traditional evaluation systems, which mainly rely on publication and citation metrics, are becoming less able to capture the new forms of scholarly contributions emerging in the open science ecosystem.

First, the astrophysics case study shows that preprints speed up the dissemination of research outputs. They can generate substantial academic impact even before formal publication. However, this early impact is usually not counted together with postpublication impact. Second, the comparison of two astrophysicists suggests that academic impact and societal impact do not always align. Altmetric indicators therefore offer an important additional perspective for evaluating scholars' societal impact. Third, the CEADs case shows that open data have become an important form of research output in data-intensive disciplines. However, their value is still difficult to recognize in paper-centered evaluation systems.

Based on these findings, we propose an open science-oriented framework for scholar research evaluation. This framework moves beyond traditional publication- and citation-based metrics and captures broader contributions in the context of open science. It consists of four core dimensions: academic impact, societal impact, open data and infrastructure, and open science practice. Compared with traditional evaluation systems, it offers a more comprehensive approach to assessing scholars' contributions in the open science era. More importantly, it helps connect open science principles with ongoing reforms in scholarly research evaluation.

This study has several limitations that should be acknowledged. The case analyses focus on representative scholars from two disciplines, namely, astrophysics and Earth sciences. This may limit the generalizability of the findings to other fields. Future research could expand the sample range to include open science practice in a broader range of disciplines, such as social sciences, humanities, and engineering. This would improve the generalizability and explanatory power of the findings.

Table 2 An open science-oriented framework for scholar research evaluation

Category	Metric	Data source
Academic impact	Scholarly outputs, citation, FWCI/CNCI, h-index	Bibliographic databases (e.g., WoS, Scopus, CNKI)
Societal impact	Social media attention, media mentions, policy citations, public engagement	Altmetric aggregators (e.g., Altmetric, PlumX); policy databases (e.g., Overton, WHO, OECD); social media platforms (e.g., Twitter/X, Facebook, WeChat); news databases (e.g., LexisNexis, GDELT, CNKI Newspaper database)
Open data and infrastructure	Dataset citations, dataset downloads, data reuse, infrastructure usage	Data repositories (e.g., Zenodo, Figshare, Dryad)
Open science practice	OA rate, preprints, open data sharing, open code sharing, open peer review	OA databases (e.g., Unpaywall, DOAJ); preprint servers (e.g., arXiv, bioRxiv, medRxiv); code repositories (e.g., GitHub, GitLab)

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

- Abdill, R. J.; Blekhman, R. 2019. Tracking the popularity and outcomes of all bioRxiv preprints. *eLife* Vol. 8, Article No. e45133.
- Altmetric. 2025. The donut and Altmetric attention score. Available at <https://www.altmetric.com/about-us/our-data/donut-and-altmetric-attention-score/>
- Amath, A.; Ambacher, K.; Leddy, J. J.; Wood, T. J.; Ramnanan, C. J. 2017. Comparing alternative and traditional dissemination metrics in medical education. *Medical Education* Vol. 51, No. 9, 935–941.
- American Society for Cell Biology. 2013. San Francisco declaration on research assessment. Available at <https://sfidora.org/read/>
- Bornmann, L.; Haunschild, R.; Adams, J. 2019. Do Altmetrics assess societal impact in a comparable way to case studies? An empirical test of the convergent validity of Altmetrics based on data from the UK research excellence framework (REF). *Journal of Informetrics* Vol. 13, No. 1, 325–340.
- BP. 2024. BP statistical review of world energy 2024. Available at <https://www.energyinst.org/statistical-review>
- Chaignon, L. 2025. Highly cited researchers: Anatomy of a list. *Quantitative Science Studies* Vol. 6, 305–327.
- Colavizza, G.; Hrynaskiewicz, I.; Staden, I.; Whitaker, K.; McGillivray, B. 2020. The citation advantage of linking publications to research data. *PLoS One* Vol. 15, No. 4, Article No. e0230416.
- Cramer, M.; Reinhart, M. 2026. Governing science through evaluation: A global heuristic of research evaluation regimes. *Minerva* Vol. 64, No. 2, 1–26.
- Elsevier. 2025. PlumX metrics. Available at <https://www.elsevier.com/insights/metrics/plumx>
- Erdt, M.; Nagarajan, A.; Sin, S. J.; Theng, Y. L. 2016. Altmetrics: An analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics* Vol. 109, No. 2, 1117–1166.
- Fraser, N.; Momeni, F.; Mayr, P.; Peters, I. 2020. The relationship between bioRxiv preprints, citations and Altmetrics. *Quantitative Science Studies* Vol. 1, No. 2, 618–638.
- Fritch, R.; Hatch, A.; Hazlett, H.; Vinkenburg, C. 2021. Using narrative CVs: Process optimization and bias mitigation. Zenodo. Available at <https://doi.org/10.5281/zenodo.5799414>
- Fu, D. Y.; Hughey, J. J. 2019. Releasing a preprint is associated with more attention and citations for the peer-reviewed article. *eLife* Vol. 8, Article No. e52646.
- Ginsparg, P. 2021. Lessons from arXiv's 30 years of information sharing. *Nature Reviews Physics* Vol. 3, No. 9, 602–603.
- Groen-Xu, M.; Börs, G.; Teixeira, P. A.; Voigt, T.; Knapp, B. 2023. Short-term incentives of research evaluations: Evidence from the UK research excellence framework. *Research Policy* Vol. 52, No. 6, Article No. 104729.
- Hicks, D.; Wouters, P.; Waltman, L.; de Rijcke, S.; Rafols, I. 2015. Bibliometrics: The Leiden manifesto for research metrics. *Nature* Vol. 520, No. 7548, 429–431.
- Hirsch, J. E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* Vol. 102, No. 46, 16569–16572.
- Ioannidis, J. P. A.; Boyack, K. W.; Baas, J. 2020. Updated science-wide author databases of standardized citation indicators. *PLoS Biology* Vol. 18, No. 10, Article No. e3000918.
- Kousha, K.; Thelwall, M.; Abdoli, M. 2018. Can Microsoft Academic assess the early citation impact of in-press articles? A multi-discipline exploratory analysis. *Journal of Informetrics* Vol. 12, No. 1, 287–298.
- Lancho-Barrantes, B. S.; Cantú-Ortiz, F. J. 2019. Science in Mexico: A bibliometric analysis. *Scientometrics* Vol. 118, No. 2, 499–517.
- Larivière, V.; Sugimoto, C. R.; Macaluso, B.; Milojević, S.; Cronin, B.; Thelwall, M. 2014. arXiv e-prints and the journal of record: An analysis of roles and relationships. *Journal of the Association for Information Science and Technology* Vol. 65, No. 6, 1157–1169.
- Lauronen, J. P. 2020. The dilemmas and uncertainties in assessing the societal impact of research. *Science and Public Policy* Vol. 47, No. 2, 207–218.
- Molas-Gallart, J. 2012. Research governance and the role of evaluation: A comparative study. *American Journal of Evaluation* Vol. 33, No. 4, 583–598.
- Nosek, B. A.; Alter, G.; Banks, G. C.; Borsboom, D.; Bowman, S. D.; Breckler, S. J.; Buck, S.; Chambers, C. D.; Chin, G.; Christensen, G.; et al. 2015. Promoting an open research culture. *Science* Vol. 348, No. 6242, 1422–1425.
- OECD. 2015. Making open science a reality. OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing. Available at <https://doi.org/10.1787/5jrs2f963zsl-en>
- Özgent, Y. 2022. Social media usage to share information in communication journals: An analysis of social media activity and article citations. *PLoS One* Vol. 17, No. 2, Article No. e0263725.
- Stephen, D.; Stahlschmidt, S. 2024. Contrasting cross-correlation: Meta-analyses of the associations between citations and 13 Altmetrics, incorporating moderating variables. *Scientometrics* Vol. 129, No. 10, 6049–6063.
- UNESCO. 2021. UNESCO sets ambitious international standards for open science. Available at <https://www.unesco.org/en/articles/unesco-sets-ambitious-international-standards-open-science>
- Wang, H.Y.; Yang, D.Q.; Li, J.; Yang, S. 2024. Construction and analysis of international preprint academic/social influence framework in 2023. *Chinese Journal of Scientific and Technical Periodicals* Vol. 35, No. 9, 1290–1302.
- Wu, Y.; Cai, Z.; Sun, F.; Bian, F.; Lin, X.; Li, Z.; Li, M.; Bauer, F. E.; Egami, E.; Fan, X. 2022. The identification of a dusty multiarm spiral galaxy at $z = 3.06$ with JWST and ALMA. *arXiv preprint arXiv: 2208.08473*.
- Wu, Y.; Cai, Z.; Sun, F.; Bian, F.; Lin, X.; Li, Z.; Li, M.; Bauer, F. E.; Egami, E.; Fan, X. 2023. The identification of a dusty multiarm spiral galaxy at $z = 3.06$ with JWST and ALMA. *The Astrophysical Journal Letters* Vol. 942, No. 1, Article No. L1.
- Zhao, W.; Gao, M.; Xia, X.; Zhang, Y. 2021. Analysis and research on Altmetrics index of domestic scholars in the field of international library and information science. In: Proceedings of the IEEE/ACIS 20th International Fall Conference on Computer and Information Science, 49–55.



Cuizhong Guan is a librarian in the Development Research Department of Tsinghua University Library. Her research interests focus on intelligence analysis and intellectual property information services.



Qian Li is a librarian of the Digital Scholarship and Communication Service at Tsinghua University Library. His research primarily focuses on open access, academic communication, and open data.



Ziyi Zhu is a Ph.D. candidate at the School of Journalism and Communication, Tsinghua University. Her research interests focus on science communication and knowledge production.