Research on the Main Innovative Features of Current International Preprint Platforms (Postprint)

Authors: Zhao Yang, Zhang Zhixiong, Zhang Zhixiong

Date: 2023-03-06T00:00:00+00:00

Abstract

Objective: To analyze the functional innovations and future development directions of current international preprint platforms, providing references and insights for the functional development of domestic preprint platforms in China. Methods: This study systematically reviews the functional innovations of international preprint platforms in areas such as innovative community collaboration, evaluation and assessment, editorial services, and peer review mechanisms, analyzing the innovative service practices introduced by each platform compared to conventional functions. Results: The five major functional innovations of current international preprint platforms are: constructing innovative collaborative community models, integrating multi-party academic exchange and discussion interfaces, serving authors to enhance paper quality, promoting open review of papers, and providing research evaluation functions. Conclusion: By summarizing and analyzing the development trends in functional innovations of current international preprint platforms, this study provides recommendations for the functional development and construction of domestic preprint platforms in China, aiming to further enhance the role of preprints in academic communication within the country.

Full Text

Preamble

Main Innovative Functions of Current International Preprint Platforms

Zhao Yang[1,2] *Zhang Zhixiong*[1,2] [1] National Science Library, Chinese Academy of Sciences, No. 33 Beisihuan West Road, Zhongguancun, Beijing 100190 [2] Department of Information Resources Management, School of Economics and Management, University of Chinese Academy of Sciences, No. 33 Beisihuan West Road, Zhongguancun, Beijing 100190

Abstract

[Purposes] To analyze the functional innovations and future development directions of current international preprint platforms, and to provide reference for the development of China' s preprint platforms. [Methods] This study systematically reviews functional innovations in areas such as innovative community cooperation, assessment and evaluation, editorial services, and review methods, analyzing the innovative service practices introduced by each platform compared to conventional functions. [Findings] The five major functional innovations of current international preprint platforms are: building innovative cooperative community models, integrating multi-party academic communication and discussion channels, serving authors to improve paper quality, promoting open review of papers, and providing scientific evaluation functions. [Conclusions] This paper summarizes and analyzes the development trends in functional innovation among international preprint platforms, and offers recommendations for the functional development and construction of China' s preprint platforms, aiming to further enhance the role of preprints in Chinese academic communication.

Keywords: Preprint platform; Functional innovation; Preprint communication; Innovative practice

Classification Number: G237.6

1. Introduction

Preprint platforms are academic platforms that provide researchers with channels for preprint publishing and dissemination while facilitating the exchange of scholarly manuscripts. Since Paul Ginsparg established the first preprint platform, arXiv, at Los Alamos National Laboratory in 1991, preprint platforms have primarily focused on functions such as manuscript submission, publication, retrieval, and browsing, continuously improving mechanisms for content screening, review methods, and version management to enhance processing efficiency and accelerate manuscript publication and exchange.

In recent years, with the rise of open access and open science, academic communication based on preprints has gained significant attention from the scientific community. Many preprint platforms, beyond their traditional functions of manuscript submission, storage, and publication, have continuously pursued functional innovations and introduced creative features. For example, arXiv has partnered with third-party organizations to create open innovation communities, embedding third-party tools into its article record interface to display relevant content from various perspectives and expand manuscript impact [1]. Guided by the core philosophy of accelerating research dissemination without charging any fees, these platforms aim to play a greater role in preprint-based academic communication. Multiple preprint platforms (such as SSRN and Research Square)

have also collaborated with international academic publishers to explore diverse and distinctive transitional practices between preprint manuscripts and traditional journals [2], challenging conventional academic communication models and providing demonstration models for the development and reform of the publishing industry. The innovative development and practice of preprint platform functions can better serve research, disciplinary, and social communities, promoting academic exchange among researchers. However, few studies have systematically examined the development trends of these emerging functions. Therefore, this paper reviews the continuous emergence of functional innovations in international preprint platforms. Through comparative analysis with these platforms' conventional functions, five major innovative directions merit our attention: collaborating with third-party organizations to build open innovation communities, launching comment panels to support scientific discussion and review, serving authors to improve paper quality, promoting open peer review, and constructing evaluation systems to support research assessment. This paper focuses on analyzing these five functional innovation directions, emphasizing their novelty compared to traditional functions, hoping to provide reference for the future development of China' s preprint platforms while offering insights for the transformation of scientific journals, improvement of publishing mechanisms, and innovation of publishing services.

2. Research Objects and Methods

For the selection of international preprint platforms, this study chose five representative platforms based on criteria including platform scale, operating organization, influence, and subject coverage: arXiv, bioRxiv, Research Square, F1000 Research, and SSRN. Among them, arXiv is the first international preprint platform and remains the most important one globally, operated by Cornell University since 2001, providing preprint deposit, retrieval, publication, and exchange services in eight fields including physics, mathematics, and computer science. bioRxiv is a life sciences preprint platform launched by Cold Spring Harbor Laboratory in November 2013, which has developed rapidly with strong support from the biology community. In 2017, the "rise of biology preprint communication" marked by bioRxiv was selected as one of Science magazine' s top ten scientific breakthroughs [3], representing a major cultural shift in academic communication. F1000Research is a life and biomedical research preprint platform acquired by Taylor & Francis Group, offering innovative open access publishing with rapid publication and open peer review services while supporting data storage and sharing. Research Square is a multidisciplinary preprint platform invested in by Springer Nature, while SSRN is a multidisciplinary preprint platform acquired by Elsevier.

The authors visited the official websites of these five platforms, focusing on the preprint manuscript introduction interfaces. Through comparative analysis with conventional functions (such as online submission, reading, browsing, and retrieval [4]), this study systematically analyzed each platform' s innovative

features and common innovation trends across platforms.

3.1 Building Innovative Cooperation Communities to Strengthen Individual and Organizational Collaboration

To facilitate better exchange, dissemination, and diffusion of preprints, preprint platforms invite community participation and collaboration in developing tools that benefit the scientific community, aiming to contribute to platform development and preprint academic exchange. For example, arXiv actively engages with the community, promoting open innovation cooperation with individuals and organizations by launching a new, formal framework called arXivLabs in September 2020 [1]. This framework supports collaborators in sharing newly developed features on the arXiv platform. These innovative functions are embedded below existing content on arXiv article record pages to discover more knowledge related to manuscripts, with framework content continuously updated as new projects launch.

Through arXivLabs, arXiv has integrated several innovative functions, including manuscript citation functionality, related manuscript discovery, and data/code linking capabilities.

The manuscript citation functionality currently integrates three tools: Bibliographic Explorer, Litmaps, and scite, each presenting citation information related to manuscripts in different forms. Bibliographic Explorer presents user-friendly paper citation navigation trees showing information about papers that cite and are cited by arXiv papers and their versions, with data sourced from Semantic Scholar and Prophy databases. Users can sort and filter by publication year, first author, title, and impact, with the interface shown in [Figure 1: see original paper]. Litmaps visualizes relationships among a paper' s references, allowing users to focus on specific references and explore surrounding citation networks [5], with the analysis interface shown in [Figure 2: see original paper]. scite (smart cite) is an intelligent citation platform [6] that records citation contexts for each paper, showing not only citation counts but also specific citation content in citing documents, intelligently categorized to reveal citation attitudes including supportive, mentioning, and contrasting types. Readers can view scite reports to see how each preprint is discussed or cited, find co-citations, and filter articles by citation type (supportive, contrasting, or mentioning) [7].

The related manuscript discovery function recommends relevant papers to readers for reference, currently supported by CORE Recommender and Connected Papers. CORE Recommender [8] aggregates open access paper resources from arXiv and over 10,000 other open access data providers, recommending relevant open access papers from a global network of research repositories. Connected Papers [9] helps researchers find and explore papers related to their field by analyzing 50,000 papers related to the manuscript from the Semantic Scholar Paper Corpus and ultimately selecting dozens of most closely related papers, visualizing relevant articles to facilitate focus on highly cited papers.

The data/code linking function advances research progress, supported by Papers With Code. Papers With Code [10] provides a simple and convenient way to find relevant code for machine learning-related papers published on arXiv. In addition to official code links mentioned in papers, Papers With Code connects community code released by others, enabling researchers to quickly and conveniently use and reference work. This function also provides links to datasets used in papers and shows how many other papers have used each dataset, offering reference for dataset usage. Furthermore, to enhance data usability, arXiv hosts its integrated collection of 1.7 million preprints on the Kaggle data platform for free user access [11], updated weekly. The dataset includes metadata fields such as paper titles, authors, categories, abstracts, and version information, with full-text information accessible via API calls in batches. This dataset provides rich corpus resources for projects like CORE Recommender and natural language processing technologies.

3.2 Launching Comment Windows to Promote Academic Exchange and Discussion

After preprint publication, effective academic exchange and feedback can prompt authors to further improve their research content. To this end, many preprint platforms have developed comment sections on each preprint manuscript interface to collect and publicly display comments from various academic groups and communities. For example, bioRxiv has developed a comment integration dashboard that provides convenient access to multi-source comments [12], encouraging readers to comment on and discuss preprints while supporting feedback to authors through various means. Specifically, the dashboard integrates reader comments, peer reviews, community discussions, and mentions of preprints in social and traditional media, linking to comment sources to enhance manuscript discoverability and make it easier for readers to find and access this information. As shown in [Figure 3: see original paper], the dashboard is guided by a launch bar above the bioRxiv article abstract interface, where different icons represent different preprint discussion or comment sources, each displaying interaction frequencies. Clicking an icon opens a specific comment panel showing detailed entries. The dashboard' s information is provided through continuous integration with third-party platforms and various community initiatives, including Hypothesis, Disqus, Altmetric, and Sciety.

The "Comments" link in the dashboard shows comments submitted through the bioRxiv platform. Readers can comment directly on bioRxiv or through the comment service platform Disqus. Comments primarily discuss scientific content, including feedback and/or specific questions about preprint methods, results, or data. All comments must be posted in English and can only be submitted once. Content that is offensive, irrelevant, or questions the author' s character, ability, or motivation is excluded. All comments undergo moderation to ensure compliance with bioRxiv' s comment policy, typically within 24-48

hours.

The "TRiP Peer Reviews" link displays comments solicited by journals participating in the "Transparent Review in Preprints" (TRiP) pilot program and other peer review service organizations. TRiP [13] is a new pilot program developed by bioRxiv in collaboration with other organizations to help broaden the academic publishing ecosystem. Supported by the web annotation tool Hypothesis, TRiP allows participating organizations to publish dedicated peer review comments on the bioRxiv platform. Authors can opt into TRiP when submitting to participating organizations such as eLife, Review Commons, and Peerage of Science. This enables journals and peer review services to publish their solicited peer review comments alongside submitted versions of preprint manuscripts. Other comment and discussion sources are gradually being developed. The "Community Reviews" link shows comments from community groups on preprint manuscripts. bioRxiv supports initiatives for independent preprint review, though these comments are not endorsed by bioRxiv or preprint authors. bioRxiv is also developing automated methods for screening and evaluating preprints, which will be provided under the "Automated Evaluations" link. To inform readers about discussions of individual preprints in social and traditional media, bioRxiv includes a "Blog/Media" source in the dashboard, covering mentions in blogs, newspapers, and other media. Two additional links display related videos and Twitter discussions about the preprint.

Beyond bioRxiv, other preprint platforms such as Research Square, SSRN, and F1000 Research have also developed corresponding comment sections for academic discussion. Users can also comment on and discuss preprints through third-party platform tools like Hypothesis and Disqus.

3.3 Serving Authors to Improve Paper Quality

To facilitate rapid transition of manuscripts to target journals after publication, preprint platforms provide a series of new knowledge service 增长点 (growth points) around manuscripts to maximize service to authors and improve paper quality. Research Square offers author services to help further improve preprints, enhance paper quality, facilitate formal publication, and promote preprints to expand their impact. Its main author services include editing services, promotion services, and certification services.

Editing services focus on improving paper content quality and include four specific forms [14]: (1) English Editing: domain experts related to the paper' s subject perform language editing, carefully correcting errors in spelling, grammar, wording, language, and punctuation while preserving the author' s original meaning, structure, and logic, and optimizing writing style and flow; (2) Digital Editing [15]: an automated editing service using intelligent technology that improves writing style without affecting content, correcting grammar, wording, and citations in track-changes mode for author discretion, while automatically assessing writing quality and providing language quality scores before and af-

ter editing; (3) Manuscript Formatting: formatting manuscripts to meet target journal requirements, including updating page layout, text formatting, headings, title pages, image placement, and reference formats to comply with journal guidelines, checking reference accuracy, and indicating necessary revisions to meet journal word limits for titles, abstracts, main text, and figure legends; (4) Figure & Table Services: generating figure legends and tables that meet target journal specifications from author manuscripts, including updating file formats, resolution, color space, fonts, scale, line width, and layout for improved readability, and improving table layout, fonts, spacing, borders, and shading to accurately and effectively convey results.

Promotion services help research papers generate impact and communicate research clearly and accurately, including three service types [16]: (1) Video Services: creating accurate summaries of articles, which can be 1-minute video bytes outlining research and its public impact with keywords, short titles, descriptions, and article links; 2-3 minute video abstracts explaining methods, findings, and contributions to the field for professional audiences; or custom videos including GIF animations and promotional videos; (2) Summary Services: capturing key points in news style to help authors communicate research to broad audiences, including research promotion services where professional teams extract key information, create promotional materials, and research highlight services that produce brief written summaries for easier sharing; (3) Infographic Services: visually highlighting the most important research points, including custom infographics describing research in concise, eye-catching, and easily understandable ways, and visual abstracts providing visual summaries of background, methods, and main findings.

Certification services provide badge certification for preprints published on Research Square to guarantee article quality, with badged preprints indicating compliance with established standards for scientific reporting [17]. Badge certification includes two types: (1) Methods Reporting Badge: examining specific material details, research design rationality, data collection and analysis, and software/code for data storage to ensure method transparency and reproducibility; (2) Data Reporting Badge: evaluating the rigor of statistical analysis details and experimental result data.

3.4 Adopting Open Peer Review Models to Promote Paper Open Review

To enable rapid formal publication after preprint submission, preprint platforms actively explore new publishing models by collaborating with international academic publishers in various ways, integrating submission systems with preprints to accelerate manuscript review processes. Platforms like Research Square and SSRN have launched journal preprint zone models [2], advocating that manuscripts submitted to publishers' journals can be first published on preprint platforms and, according to journal requirements, can publicly disclose peer review comments and review processes to varying degrees. Alternatively,

manuscripts published on preprint platforms can be transferred to affiliated journals, accelerating preprint exchange and dissemination in a more open and transparent manner.

F1000 Research fully leverages the advantages of both preprints and journal publishing in its platform, employing an open and transparent, author-driven peer review model for article evaluation and publication [18], thereby promoting open review of papers. Specifically, after manuscript publication, papers are first marked with "AWAITING PEER REVIEW" status next to the title, displayed in an open peer review summary box in both HTML and PDF versions. F1000 Research requires authors to recommend at least five potential experts in their field as reviewers, with formal publication requiring at least two peer review reports on the platform. Once reviewers are confirmed, expert reviewers receive formal invitations to begin the review stage.

Reviewers receive guidelines for each article type to assess scientific soundness, covering four aspects: whether the research is appropriately contextualized within existing literature; whether appropriate methods are employed; whether sufficient information and source data are provided to ensure reproducibility; and whether results support conclusions. For certain article types such as case reports or review articles, reviewers are asked to comment on facts and methods rather than necessarily agreeing with the author' s viewpoints.

Beyond review comments, reviewers must provide explicit decisions on manuscripts, including three approval statuses [19]: (1) Approved: requiring no or only minor modifications, indicating for original research that experimental design (including variable control and methods) is adequate, results are accurately presented, and conclusions are reasonable and data-supported; (2) Approved with Reservations: indicating academic value but requiring minor modifications or specific, significant revisions; (3) Not Approved: indicating poor quality with fundamental flaws that severely undermine results and conclusions. Approval status is displayed with the article, along with reviewer names, affiliations, and detailed reports supporting their decisions. If authors revise articles in response to reviewer comments, all reviewers are invited to provide supplementary reports on the new version. If reviewers initially gave "Approved with Reservations" or "Not Approved" status, F1000 Research encourages re-review to assess whether the work has sufficiently improved to merit better approval status.

Article peer review status updates continuously as reports are published, with review progress clearly marked on each article as part of the citation. Once received, peer review reports are published with the article and display approval status. As more reports are received, peer review status is updated, with all versions, related datasets, and peer review reports deposited in PubMed Central regardless of final review status. Once an article receives two "Approved"statuses or two "Approved with Reservations" statuses plus one "Approved" status, it is indexed in PubMed and other bibliographic databases. If authors struggle to secure reviewers for an extended period, a few articles may be marked as

"PEER REVIEW DISCONTINUED."

3.5 Building Evaluation Systems to Support Research Assessment

To further expand preprint impact, preprint platforms have introduced various indicator systems to evaluate papers from different perspectives. For example, SSRN combines relevant indicators to compile statistics and rankings for papers, authors, and organizations, forming an evaluation system that provides multi-dimensional value judgments for preprint manuscripts to support research assessment. Specifically, SSRN uses three methods to measure paper impact: paper statistics, rankings, and PlumX Metrics [20]. Paper statistics can be viewed on each article' s abstract page, including views, downloads, and other article-level metrics such as citation frequency. Rankings sort papers, authors, and organizations on the platform according to relevant criteria. Paper rankings primarily calculate download counts and rank papers across different disciplinary categories, including the ten most downloaded papers since publication (all-time Top 10) and the top ten papers by download count in various SSRN e-journals over the past 60 days (recent Top 10). Author rankings are primarily based on affiliation and download counts, with authors requiring at least one publicly available full-text paper on SSRN to qualify, while schools and organizations must meet SSRN-specific criteria. PlumX Metrics primarily provide insights into how people interact with research outputs (articles, conference proceedings, book chapters, etc.) in online environments, collecting and aggregating appropriate research metrics for all types of scholarly output across five independent categories [21]: Citations, Usage, Captures, Mentions, and Social Media interactions.

Beyond SSRN' s impact indicators, other preprint platforms provide their own metrics on manuscript interfaces. bioRxiv' s Metrics interface records paper usage, allowing users to click and view views, downloads, and Altmetrics counts for the current or past six months. Research Square' s Engagement function shows current manuscript views and downloads on the platform, as well as Altmetrics and Dimensions counts. F1000 Research' s Metrics data records current manuscript views and downloads on the platform and in PubMed Central (PMC), with PMC usage updated monthly. Metrics also record citation counts in Scopus and PubMed databases, as well as Altmetrics and Dimensions counts.

4. Summary and Implications

Facing the continuously innovating preprint platforms, this paper identifies several development directions for functional innovation: (1) Integrating into communities for open innovation collaboration by associating and cooperating with external organizations to expand preprint influence throughout the academic communication ecosystem, promoting broad dissemination and exchange of early research findings, and providing research inspiration for users and read-

ers in relevant fields. (2) Introducing various scientific discussion and comment channels to support open review of papers, effectively promoting academic exchange and publication of preprints. (3) Helping authors further improve paper content through preprint platform value-added services, creating high-quality papers and advancing the publication process. (4) Combining effectively with traditional citation metrics to provide multi-dimensional evaluation indicators that expand preprint impact. Displaying more diverse content indicator data helps understand research papers' potential impact on science and society, enabling readers to quickly discover high-quality articles and authors to witness the impact of their work. (5) Diversified academic output pathways centered on papers, including data and materials, are gradually opening. Beyond research papers, preprint platforms also provide publishing channels for other research outputs. For example, F1000 Research also publishes academic posters, slides, and documents covering basic science, translational research, and clinical research in life sciences and medicine. Research Square can also publish video content and laboratory manuals (Protocol Exchange) [22], which include detailed instructions on experimental procedures, computational models, equipment operation, and other processes reported in manuscript methods, providing objective and convincing experimental evidence. SSRN integrates with Mendeley Data to allow researchers to upload their research data together with papers during or after submission [23], with datasets displayed on the author' s article page.

The various distinctive features of current international preprint platforms offer strong implications for the future development of China's preprint platforms and warrant widespread attention. China' s preprint platforms can learn from these successful cases of functional innovation by effectively cooperating with research communities and academic groups, such as developing comment columns for preprints, inviting peer experts to publicly provide review comments, and linking to media interaction platforms like Weibo, Zhihu, and Douban for academic discussion. Platforms can also introduce various statistical indicators to count preprint exposure across platforms, further expanding preprint influence, and cooperate with third parties to provide editing and promotion services that meet authors' actual needs. Good service functions will provide users with better experiences, further expanding the influence of preprints and platforms themselves, attracting more potential users while broadening the beneficiary community.

5. Conclusion

Through systematic analysis of the five major functional innovation directions of these preprint platforms, it is evident that platform functional innovations have expanded the leading role of preprints in academic communication, enabling preprint platforms to better integrate into the academic environment. This paper provides example analyses of corresponding preprint platforms for each innovative function, summarizes current development directions of func-

tional innovation, and offers preliminary recommendations for the functional development and construction of China' s preprint platforms. These innovative initiatives also provide valuable insights for various aspects of academic publishing in China, offering references for introducing preprint models to academic journals, accelerating publication processes, and providing value-added services.

References

[1] arXivLabs: a space for community innovation [EB/OL]. [2022-08-26]. https://blog.arxiv.org/2020/09/28/arxivlabs-a-space-for-community-innovation/

[2] 杨硕. 预印本库与传统期刊合作的典型模式研究 [J]. 科技与出版,2020(07):130-136.

[3] Science. 2017 BREAKTHROUGH OF THE YEAR[EB/OL]. [2022-08-26]. https://vis.sciencemag.org/breakthrough2017/finalists/#cosmic-convergence

[4] 康宏宇, 钱庆, 李姣等. 国内外预印本系统比较研究与展望 [J]. 医学信息学杂志,2020,41(11):26-31+40.

[5] Litmaps[EB/OL]. [2022-08-26]. https://www.litmaps.com/

[6] scite_[EB/OL]. [2022-08-26]. https://scite.ai

[7] New arXivLabs collaboration provides citations in context from scite[EB/OL]. [2022-08-26]. https://blog.arxiv.org/2021/08/26/arxivlabs-citations-in-context-scite/

[8] Knoth P, Anastasiou L, Charalampous A, et al. Towards effective research recommender systems for repositories[J]. arXiv preprint arXiv:1705.00578, 2017.

[9] Connected Papers [EB/OL]. [2022-08-26]. https://www.connectedpapers.com/

[10] New arXivLabs feature provides instant access to code [EB/OL]. [2022-08-26]. https://blog.arxiv.org/2020/10/08/new-arxivlabs-feature-provides-instant-access-to-code/

[11] arXiv Dataset [EB/OL]. [2022-08-26]. https://www.kaggle.com/Cornell-University/arxiv

[12] bioRxiv. An easy access dashboard now provides links to scientific discussion and evaluation of bioRxiv preprints [EB/OL]. [2022-08-26]. https://connect.biorxiv.org/news/2021/05/14/dashboard

[13] Cold Spring Harbor Laboratory. Transparent review in preprints [EB/OL]. [2022-08-26]. https://www.prnewswire.com/news-releases/transparent-review-in-preprints-300927942.h

[14] Research Square. Professional Editing—From the lab to a publication-ready manuscript [2022-08-26]. https://www.researchsquare.com/researchers/editing

[15] Research Square. Digital Editing—High-quality technical editing in under 10 minutes [EB/OL]. [2022-08-26]. https://www.researchsquare.com/researchers/digital-editing

[16] Research Square. Research Promotion [EB/OL]. [2022-08-26]. https://www.researchsquare.com/researcher

[17] Research Square. Badges [EB/OL]. [2022-08-26]. https://www.researchsquare.com/researchers/badges

[18] F1000Research's peer review model [EB/OL]. [2022-08-26]. https://f1000research.com/for-referees/guidelines#rar

[19] F1000 Research. How it works [EB/OL]. [2022-08-26]. https://f1000research.com/about

[20] How can researchers gauge a paper' s impact? [EB/OL]. [2022-08-26]. https://service.elsevier.com/app/answers/detail/a_{id}/34355/supporthub/ssrn/

[21] About PlumX Metrics [EB/OL]. [2022-08-26]. https://plumanalytics.com/learn/about-metrics/

[22] Research Square. Protocols [EB/OL]. [2022-08-26]. https://www.researchsquare.com/browse?journal=prot exchange

[23] How does SSRN and Mendeley Data integration work? [EB/OL]. [2022-08-26]. https://service.elsevier.com/app/answers/detail/a_{id}/34386/supporthub/ssrn/

Author Contribution Statement:

Zhao Yang: Designed research plan, implemented research process, wrote and revised paper;
Zhang Zhixiong: Proposed research direction and framework, revised paper, finalized manuscript.

Note: Figure translations are in progress. See original paper for figures.