Volume 2, Issue 1, 2025

Intuitions & Insights

An Interdisciplinary Research Journal

ISSN: 3048-6793

A flourishing world of preprint repositories: An outline

Hirak Roychaudhuri

Librarian, Khatra Adibasi Mahavidyalaya, West Bengal, India

Email: roychaudhuri.hirak81@gmail.com

Abstract: **This article delves into the realm of preprint repositories, elucidating their subject areas, benefits, challenges, and limitations. It highlights their role in accelerating research dissemination, enhancing visibility, and promoting open access while addressing issues such as quality control and misinformation. Preprint repositories serve as online platforms for researchers to share their findings prior to peer review or formal publication. This article provides a detailed overview of their history, functionality, advantages, and ongoing developments, offering insights for researchers and stakeholders in academic publishing.**

Keywords: Digital Object Identifier, Open Science Framework, Repositories, Social Science Research Network.

*Corresponding author: **Hirak Roychaudhuri**

Received: 01.12.2024; **Accepted**: 31.01.2025; **Published**: 10.02.2025

1. Introduction

Preprints are manuscripts shared publicly before undergoing peer review or formal publication. They serve as a means for researchers to receive early feedback, accelerate dissemination, and improve access to their work. While not all preprints ultimately lead to publication in renowned journals, their primary purpose lies in fostering academic dialogue and community engagement.

Preprint servers, also known as preprint archives, assign persistent identifiers such as Digital Object Identifiers (DOIs), enabling seamless citation. Though DOIs for preprints differ from those assigned to

published versions, preprints remain a part of the scholarly record. Notably, while preprints can be withdrawn, essential metadata, such as author information and withdrawal reasons, remain visible.

2. Literature Review

This article draws upon resources such as Maastricht University's *A Practical Guide to Preprints*, various preprint databases (e.g., arXiv, bioRxiv, SSRN, medRxiv), and journal articles(e.g:Science Direct etc.). These sources provide foundational insights into the functionality and evolution of preprint repositories.

Advantages and Disadvantages of Preprints:

Advantages for Researchers:

1. **Rapid Publication:** Preprints are uploaded quickly, bypassing lengthy review processes.
2. **Open Licenses:** Documents are accessible under open licenses.
3. **Enhanced Visibility:** Public access to preprints continues to grow.
4. **Recognition:** Researchers gain visibility and recognition for their work.
5. **Timely Feedback:** Quick feedback strengthens confidence and quality.
6. **Citations:** Articles with DOIs are easier to cite.
7. **Cost-Effectiveness:** Publishing preprints incurs minimal costs.

3. Advantages for Both Researchers and the Public:

1. **Open Access:** Free access fosters equitable knowledge dissemination.
2. **Community Engagement:** Preprints invite public comments and collaboration. (Hettne, K. M.2021)

Disadvantages:

1. **For Researchers:** Some journals may reject manuscripts previously shared as preprints.
2. **For Both Researchers and Publishers:**
    - Risk of pseudoscience dissemination.
    - Difficulty in assessing research quality and originality.(Hettne, K. M.2021)

Historical Overview and Recent Developments:

The concept of preprints dates back to the early 1990s, with the launch of arXiv, primarily for physics. Subsequently, disciplines such as mathematics, computer science, and biology embraced preprint repositories. Key milestones include:

- **1994:** Establishment of the Social Sciences Research Network (SSRN).
- **2007:** Introduction of Nature Proceedings for biomedical sciences (discontinued in 2012).
- **2013:** Launch of bioRxiv for life sciences.
- **2016:** The Open Science Framework (OSF) initiated preprint servers, initially targeting psychology.
- **2019:** Launch of medRxiv for medical disciplines, gaining prominence during the COVID-19 pandemic.

Preprint platforms such as Research Square, Preprints.org, and Authorea allow early-stage linkage to specific journals. Other platforms, like F1000 Research, offer features akin to preprints but operate under distinct editorial parameters.(Hettne, K. M.2021).

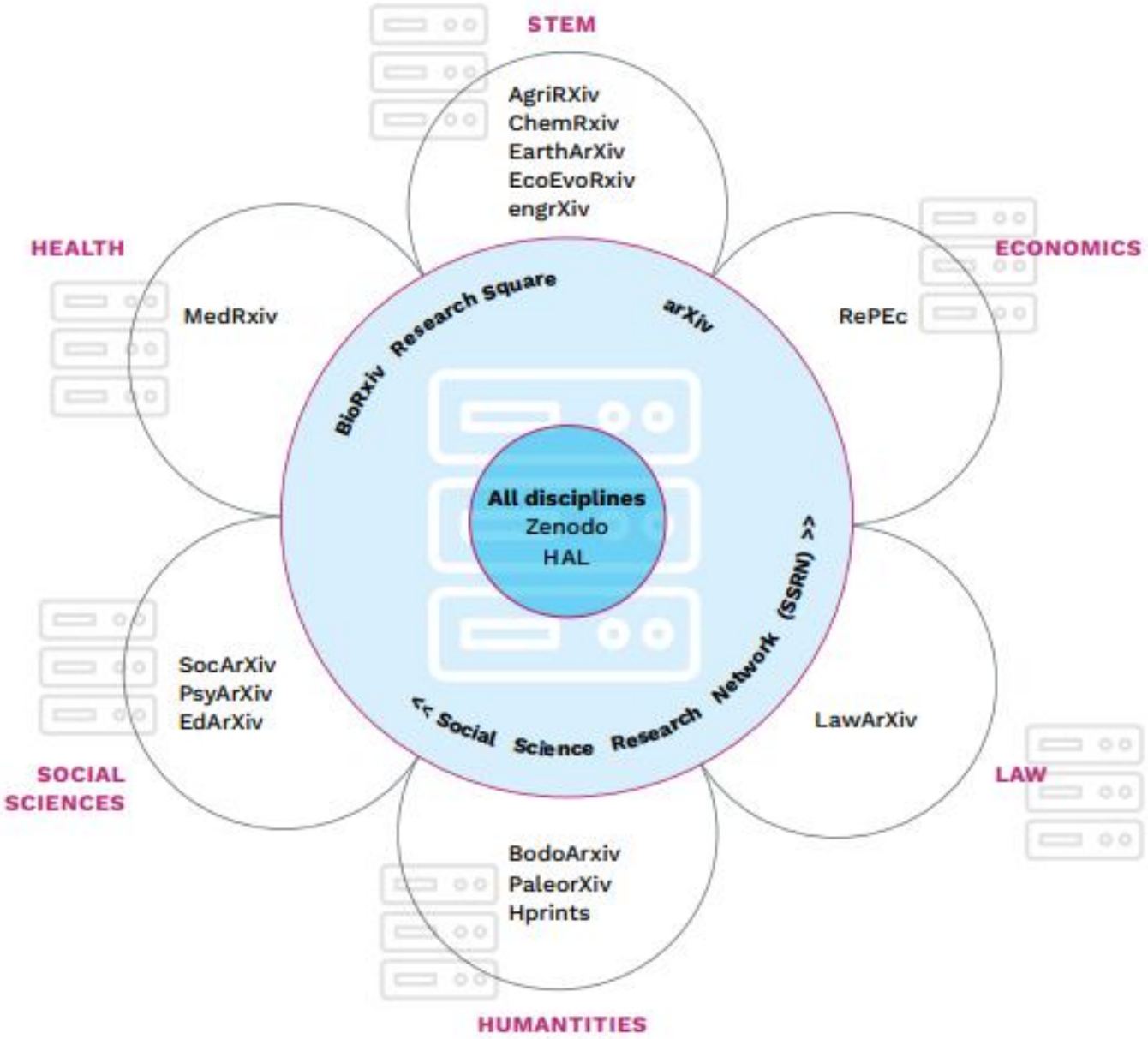
Factors in Choosing and Using Preprint Servers:

When selecting a preprint server, researchers should consider audience reach, interdisciplinary compatibility, and long-term visibility. Essential features include:

1. **Persistent Identifiers:** Such as DOIs.
2. **Longevity:** Guarantee of online visibility.
3. **Version Updates:** Support for revisions.
4. **Quality Checks:** Screening for plagiarism, metadata accuracy, and ethical concerns.
5. **Open Licenses:** Encouraging broad accessibility.
6. **Indexation:** Inclusion in search engines like Google Scholar.

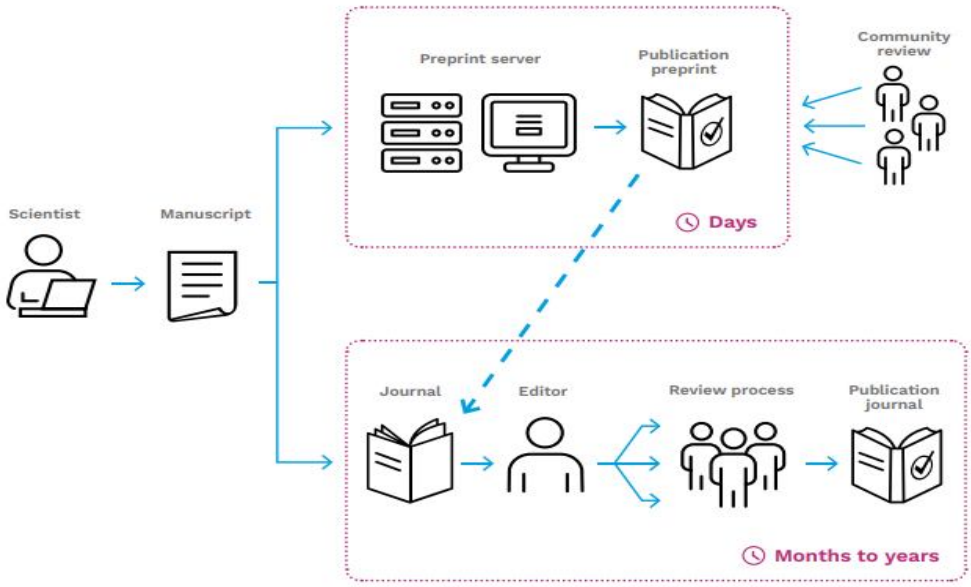
Posting, Updating, and Withdrawing Preprints:

1. **Posting:** Verify journal policies, obtain co-author consent, and share supplementary data.
2. **Updating:** Most servers support versioning without overriding earlier versions.
3. **Withdrawing:** Allowed for legal or ethical reasons, ensuring metadata transparency.



How to link a preprint to the published journal article:

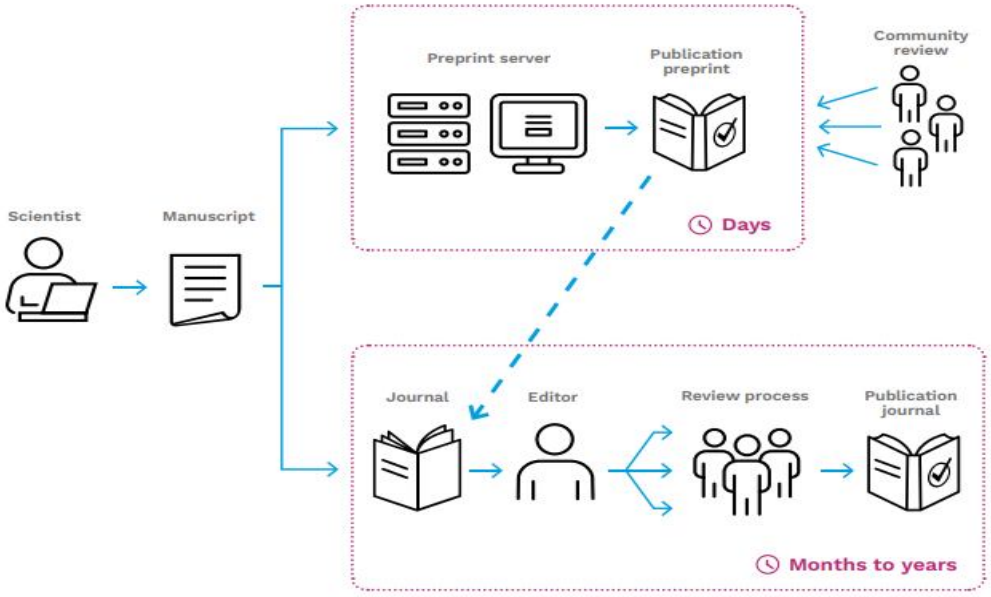
Some preprint servers create links to the published journal article and add them to the metadata of the preprint. If such a link does not automatically show on one's preprint page, one can check with the preprint server if he/she needs to do it himself/herself or if they can do it for him/her. Several journals add links from the published article back to the preprint. If the latter is not done, one should add a footnote with the link to the preprint to the manuscript that one will submit to the journal. It increases the transparency of one's work. (Hettne, K. M.2021).



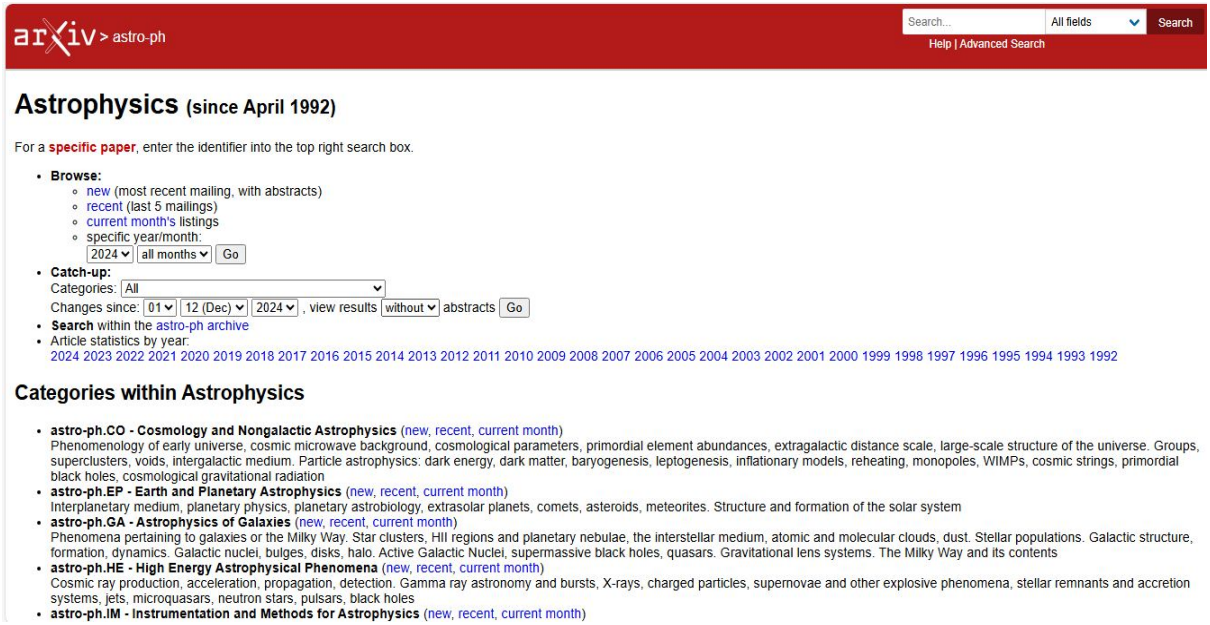
4. Some specific points to recognize a preprint

a) The first page of the preprint generally declares that one dealing with a paper that has not been "published" officially. One may find words like 'preprint', 'concept paper', 'working paper' or 'author manuscript' (not to be messed with author's accepted manuscript; which is a peer reviewed but not yet "published" version of a paper).

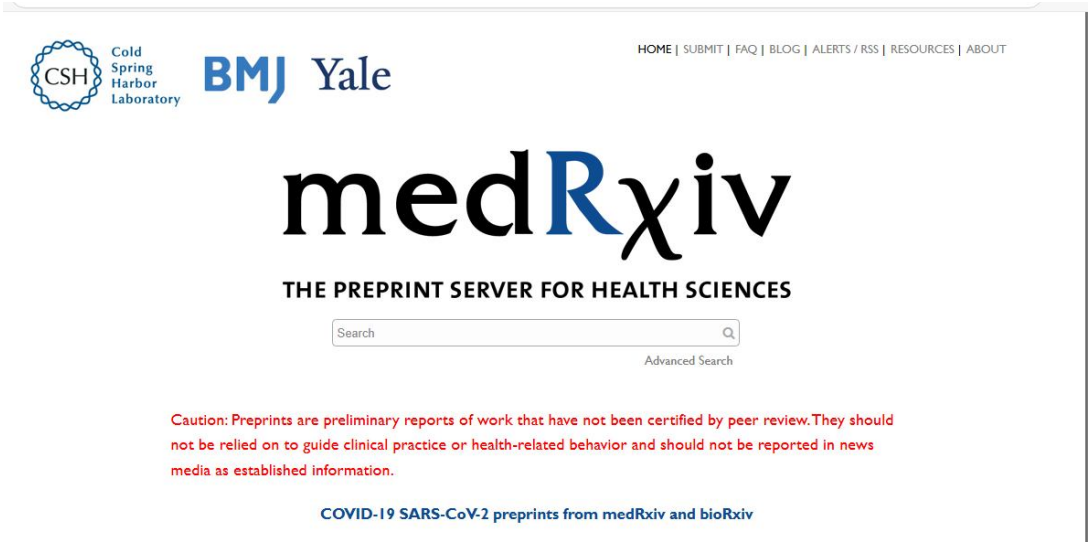
b) A preprint is generally posted on a preprint server . It can be stated that some preprint servers also host the author's received manuscript and that some generic platforms like Zenodo and Figshare can host any versions of papers.((Hettne, K. M.2021)



Some images of Preprint repositories:



ArXiv



(https://arxiv.org/)

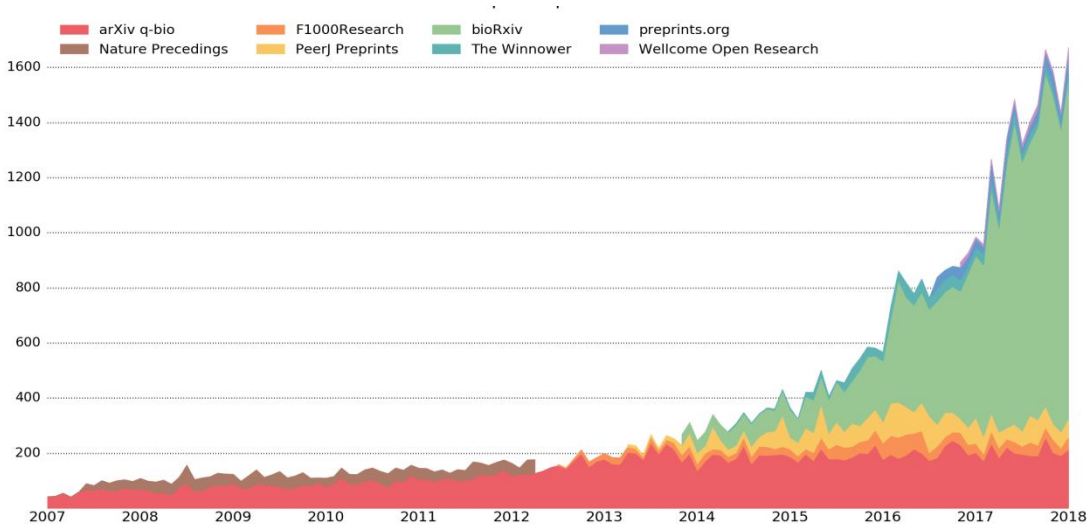
MedRxiv(https://www.medrxiv.org)

5. Current States of Preprints

In the last 2-3 years, there has been a fast growth of the preprint ecosystem, based on mutual efforts from advocacy groups, research funders, researchers, and hosting platforms and services. The Open Science Framework (OSF) permits searching across 25 different providers each with its own policies, guidelines, content, governance, financial structure, and communities.

The number of preprint submissions has been speedily growing since the mid-2010's in the Life

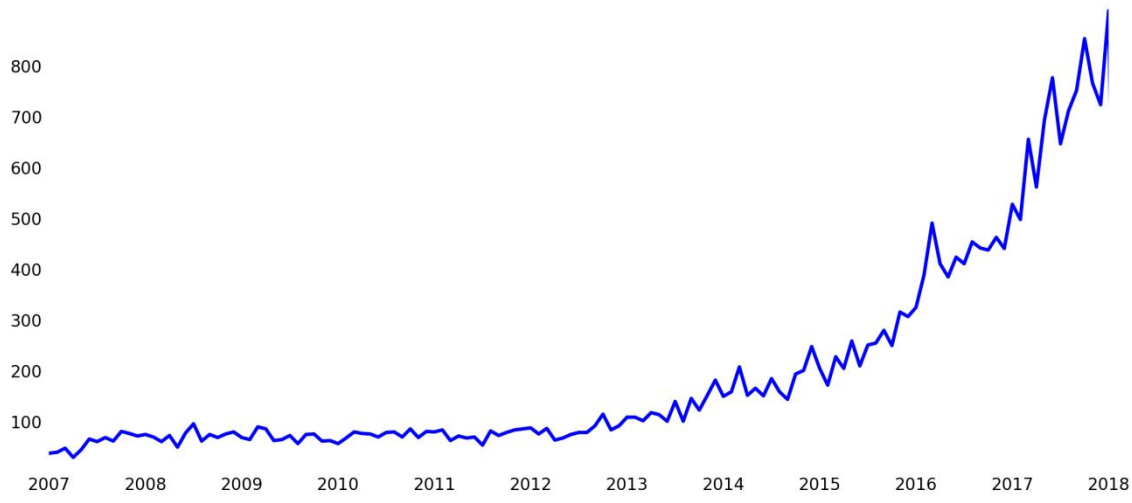
Sciences, based on data gathered by PrePubMed, frequently in relation with the arrival of the bioRxiv server hosted by Cold Spring Harbor Laboratory.



Preprints per month

Source: PrePubMed

This growth is approximately reflected in the number of new senior/first authors per month, suggesting wider contributions from researchers in the Life Sciences is a key driving issue. The European Commission's Open Science Monitor also shows a visualization and presentation of the temporal and geographical distribution of preprints from diverse fields.



New Senior Author Per month

Source: PrePubMed

One wider result of this growth is that most(around 78%) major publishers permit, or even inspire, work to be shared as preprints (counter to a common interpretation of the 'IngelFinger Rule', whereby the same research should not be published twice). The reasoning for this is likely two-fold: articles have not yet been authenticated through peer review, therefore

publishers still have the opportunity to establish their 'added-value' in scholarly communication; and if publishers were to cancel submissions that had been formerly shared as preprints, this would reduce an ever-growing proportion of their potential submission pool. As such, expansions of preprints have extensive consequences on how scholarly research is circulated, and therefore on the wider scholarly communication ecosystem as a whole.

6. Research on preprints

Recent research validates that preprints shared on bioRxiv gained more online craze and citations than similar journal articles available without preprints (Serghiou et al., 2018). This result might not be directly causal (e.g., due to other factors like extensive sharing on social media), but guides that people are at least sharing on bioRxiv research of adequately high quality, as judged by their peers, and that they are interrelating with and citing this work. According to Google Scholar, the most-highly cited source in Economics is the NBER Working Papers platform, with a h-5 index of 165 , and in Physics and Mathematics 4out of 5of the top-cited sources are sub-categories of arXiv. Such cross-disciplinary usage infers that not only are preprints becoming widely re-used by researchers, but that their adoption is becoming increasingly respected as a manner of scholarly communication. In some specialised sectors/fields/disciplines within Maths and Physics, preprints are broadly used and now have become the standard for communication (Gentil-Beccot et al., 2010; Lariviere et al., 2014).

The gradually widespread and planned adoption of preprints (and preprint servers) has the potential to dramatically affect the dispersal of research. In the future, journals would go on important in handling peer review to authenticate research articles, but such authentication and references would be openly assessed by a wider pool of readers, and their capacity to digest the content. Ultimately, it recommends that preprint servers and journals fulfill discrete roles for readers, and also have diverse effects within various research communities (e.g., Davis and Fromerth, 2007; Moed, 2007; Larivière et al., 2014; Ginsparg, 2016). Further research has lately confirmed that there are virtually no differences between articles published on the arXiv (and to a lesser extent, bioRxiv), and the final published versions, which could have weighty effects on the 'value add' claims of publishers, and economic decisions regarding scholarly communication (Klein et al., 2018).

7. Conclusion

Preprint repositories, once confined to specific academic fields, have become an essential part of scholarly publishing. They expedite the dissemination of knowledge and promote collaboration, though challenges like ensuring quality and combating misinformation remain. The COVID-19 pandemic highlighted their significance, driving wider acceptance and integration into research practices. As the preprint landscape continues to evolve, stakeholders must navigate its complexities while harnessing its transformative potential for scholarly communication. Journals should implement open preprint policies that uphold transparency while maintaining rigorous peer review standards. By using preprints responsibly, the academic community can complement traditional publishing methods, fostering both innovation and collaboration.(S,Pippa.2022)

Acknowledgment: The author is thankful to the editor and the referee for their valuable suggestions and comments.

References

[1] Davis, P.M., & Fromerth, M.J. (2007). Does the arXiv lead to higher citations and reduced publisher downloads for mathematics articles? *Scientometrics*, 71(2), 203-215. https://doi.org/10.1007/s11192-007-1661-8

[2] Gentil-Beccot, A., Mele, S., & Brooks, T.C. (2010). Citing and reading behaviours in high-energy physics. *Scientometrics*, 84(2), 345-355. https://doi.org/10.1007/s11192-009-0111-1

[3] Ginsparg, P. (2016). Preprint Déjà Vu. *The EMBO Journal*. https://doi.org/10.15252/embj.20169553

[4] Hettne, K. M., et al. (2021). A practical guide to preprints: Accelerating scholarly communication. https://doi.org/10.5281/zenodo.5600535

[5] Klein, M., et al. (2018). Comparing published scientific journal articles to their pre-print versions. *Int J Digit Libr*. https://doi.org/10.1007/s00799-018-0234-1

[6] Larivière, V., et al. (2014). arXiv e-prints and the journal of record: An analysis of roles and relationships. *J Assn Inf Sci Tec*, 65(6), 1157-1169. https://doi.org/10.1002/asi.23044

[7] Moed, H. F. (2007). The effect of “open access” on citation impact: An analysis of arXiv's condensed matter section. *J Am Soc Inf Sci*, 58(13), 2047-2054. https://doi.org/10.1002/asi.20663

[8] Pippa, S. (2022). The evolution, benefits, and challenges of preprints and their interaction with journals. *Science Editing*, 9(1), 83. https://doi.org/10.6087/kcse.269

[9] Serghiou, S., & Ioannidis, J.P.A. (2018). Altmetric scores, citations, and publication of studies posted as preprints. *JAMA*, 319(4), 402-404. https://doi.org/10.1001/jama.2017.211683