



Information Research - Vol. 29 No. 2 (2024)

Exploring scholarly perceptions of preprint servers

Shir Aviv-Reuven, Jenny Bronstein and Ariel Rosenfeld

DOI: <https://doi.org/10.47989/ir292820>

Abstract

Introduction. Preprint servers play an important role in scholarly communication. The study investigates scholars' engagement, experiences, and perceptions regarding the use of these servers, both as information sources and publishing venues. This qualitative study seeks to extend our understanding of how these servers operate within the academic ecosystem and influence scholarly communication.

Method. Data was collected through 32 semi-structured interviews with scholars from different disciplines, to explore their engagement, experiences and perceptions in using these platforms.

Analysis. The data collected from these interviews underwent thematic content analysis using ATLAS.ti software. This analysis facilitated the organization and thematic examination of the textual narratives derived from the interviews.

Results. In this study, scholars discussed their perceptions about the benefits of using preprint servers in scholarly work such as rapid dissemination of information and open access, but also raised concerns regarding the lack of peer review for the studies uploaded to these servers.

Conclusion. These findings emphasize the growing, yet diverse, role preprint servers play in scholarly communication and their differential impact across academic disciplines.

Background

The scholarly research ecosystem has evolved significantly in the last thirty years due to electronic publishing. This evolution has expanded scholarly distribution beyond traditional books and journals to include diverse formats like preprint papers, datasets, multimedia, and software (Johansson et al., 2018; Puebla et al., 2021). A preprint is a full scientific manuscript, typically awaiting peer review, uploaded by authors to a public server for a basic scientific check before being freely available online (Berg et al., 2016). These manuscripts are hosted on preprint servers, accessible to all as online archives. Once online, they are indexed by platforms such as Google Scholar, and can be accessed, cited, and commented on by readers without the need for expensive journal subscriptions.

Preprint servers play a significant role in the broader landscape of academic publishing, influencing how research findings are disseminated. They are particularly valuable in fast-paced academic disciplines, enabling the swift sharing of new research (Maggio and Fleerackers, 2023). This rapid dissemination is not only essential for advancing knowledge quickly but also promotes open access, making research findings more accessible, especially in areas with limited resources process (Biesenbender, et al., 2024; Puebla et al., 2021). According to Fraser, et al.'s (2022) findings journal articles initially shared on preprint servers tend to garner more citations and online mentions compared to articles from the same journals that weren't posted on the platform. Furthermore, these platforms facilitate early-stage scholarly discussions and collaborations, contributing to a more dynamic and iterative process in research development. This enhances the overall pace and quality of academic progress across various fields (Kodvani et al., 2022). However, preprint servers present challenges, particularly regarding the credibility and trustworthiness of non-peer-reviewed information. This raises concerns about quality assurance and the integrity of academic publishing (Fraser et al., 2022). Additionally, their influence on reshaping traditional scholarly norms and metrics cannot

be overlooked. Penfold and Polka (2020) explore how preprint servers were crucial during the COVID-19 pandemic and how Twitter facilitated the creation of a community that enhanced the visibility of preprints and provided support to their authors.

Preprint servers necessitate a nuanced understanding of their role in scholarly communication, balancing the need for swift information dissemination with maintaining research integrity (Penfold and Polka, 2020). This complex interplay highlights the evolving nature of scholarly communication in the digital era, underscoring the importance of adapting traditional academic practices to accommodate new modes of information sharing while ensuring the reliability and credibility of scholarly work. These platforms lack *official* recognition or consideration from most academic establishment due to their non-traditional review processes (Chiarelli, et al., 2019). Previous research on the use and role of preprint servers as dissemination and publication venues has predominantly employed quantitative methods. This paper presents a qualitative approach to the study of preprint servers as platforms for publishing and disseminating research, offering a distinct perspective compared to previous quantitative inquiries. This approach aims to deepen our understanding of how these servers function within the academic ecosystem and their impact on the scholarly communication process.

This study examines scholars' engagement, experiences and perceptions in using these servers both as information sources and as publishing venues. We formulate the following research question:

- How do scholars engage with, experience, and perceive preprint servers as both sources of information and venues for publishing their research?

Method

Researchers from different disciplines participated in thirty-two semi-structured interviews. The study used purposive sampling to ensure diversity across various fields, aiming

to capture a broad spectrum of perspectives on the issues under exploration. To achieve a diverse and comprehensive sample, participants were recruited from universities throughout Israel. For each discipline, an equal number of male and female researchers were interviewed.

The research methodology involved semi-structured interviews, which provided a structured set of themes and questions but also allowed for the flexibility to explore in-depth insights and professional viewpoints as they naturally arose during the conversations. All interviews lasted between 30 and 60 minutes, were conducted using Zoom, and each session was recorded and transcribed. To maintain confidentiality, identifying information was removed from the transcripts, and each recording and transcript were assigned a unique serial number to ensure the anonymity of the participants. The data gathered from these interviews was analysed using thematic content analysis with the help of AtlasTi¹ software. This analysis facilitated the organization and thematic examination of the textual narratives derived from the interviews.

Results

Participants in this study shared diverse perspectives on their interactions with preprint servers, reflecting the dynamic role of these platforms in academic research and highlighting both the benefits and challenges associated with their use. This brief exploration offers preliminary insights into how preprint servers are perceived within the scholarly community, shedding light on their impact on research dissemination and communication in the academic field. Findings from the content analysis show that preprint servers advance scientific research by facilitating rapid dissemination and open access to research findings, enabling immediate peer feedback and accelerating scholarly communication. Scholars in the medical field are increasingly utilizing preprint servers like BioArXiv and MedArXiv to expedite the dissemination of their findings. Participants emphasized that the

COVID-19 pandemic hastened the shift towards immediate public access to research, highlighting the value of prompt feedback in the academic process. Participant number 35 a researcher in the field of Medicine shared his thoughts

I think there's two things. It can be used for two purposes and people use them for those two purposes. The first one is to plant your flag, to say 'I'm the first one. Maybe my paper's not out there but I'm the first one who's doing this.' And the second one is, you have data that, you know, Covid was a good example. You know, at some point I had data that I thought can really help, you know, move things forward in the way patients are treated, are managed, and I didn't think it was ethical to wait 6 months for the results to be out for people to see. I thought, you know, people need to see it and maybe it's going to change, maybe it's not perfect, etc. but at least people can see it.

Participant number 35, further explained that uploading his findings to a preprint server significantly broadened the reach of his work, extending beyond the academic community, and that the feedback received further advanced his research (Biesenbender et al., 2024; Puebla, et al., 2021).

I put a paper on a preprint server, which was about the role of vaccination on long Covid and it really, it really got picked up, you know. It was re-shared, I don't know how many thousands of times, journalists, etc. etc. but the most, the most interesting thing was actually people reaching out to me and say 'your paper is really good and really important, but it's also flawed. Did you think about this, this and, this?' and I said 'huh, actually, I hadn't' and then I changed the paper and then we went back and realized and added and, in the end, we ended up with I think a better paper than we would have published in the first place.

Past studies have shown that a compelling motive for scholars to upload their papers to a

preprint server is the opportunity to gain visibility in a popular research area (Maggio and Fleerackers, 2023). Participant number 27 from Computer Science explains her need for using these servers

I had a paper which I submitted with my students, which I knew was with a very very popular topic and it's important we publish first so that it is clear we have done it, so immediately after submitting to a conference, within a week we placed it on ArXiv and indeed, only a week later more work (in the same topic) was published.

Participant number 22 from Computer Science explained how uploading papers to preprint servers addresses the issue of lengthy publication times in traditional journals, which can compromise the relevance and impact of his work

I currently have a paper in a journal (under review) where in a journal the time until you get a response is a year and it's a survey, it's something that will quickly become less relevant, so it was very important to me and my colleagues to upload it as quickly as possible.

Echoing these statements, findings from prior studies suggest that uploading papers to preprint servers enables scholars to establish priority in publishing their findings, thereby securing their place and recognition in the specific research area and making their work readily accessible to other researchers. (Biesenbender, et al., 2024; Fraser et al., 2022; Waltman et al., 2021).

Findings revealed additional benefits to uploading papers to a preprint server, including rapid online access to the papers, free public availability aligning with the Open Access (OA) approach, and enhanced opportunities for collaboration through open reviews (Strcic, et al., 2022). Participant number 25 from Computer Science focused on the Open Access approach, 'I think that people have ideals about moving towards all-inclusive open publication model, so some of them are really excited about making sure everything is in arXiv ...and the conferences will have to deal.'

Preprints servers and other open access initiatives have enabled the creation of a new type of publication venue called overlay journal. These publications conduct editorial assessment or peer review of publicly available papers, typically preprints, without hosting the articles themselves, instead linking to the original documents (Puebla, et al., 2021).

Participants raised concerns and described negative views and perceptions regarding preprint servers. These concerns echo findings from recent studies (Desjardins-Proulx et al., 2013; Sarabipour et al., 2019) and stated that posting a preprint negatively affected their work (Sever et al., 2019). Participant number 24 from Computer Science described an instance by which uploading a paper in a Preprint server had a negative impact on their attempts to get accepted in formal publication venues.:

We have submitted a paper to a conference that was very important to me and the paper was very important to me and placed in the ArXiv an early version of the paper with a slightly different name...The reviewer took our papers title and searched it in Google Scholar found our early paper and rejected our submitted paper because it has been already published.

The content analysis revealed that some participants considered early publication on preprint servers as a disadvantage, fearing that it might allow other researchers to access their ideas and subsequently "beat them" to formal publication, as Participant number 6 from Information Science explains: 'You sort of reveal your cards perhaps too early if you show what you are doing and you don't want to reveal it before it's published.'

Maggio and Fleerackers (2023) highlight the risks associated with early posting on preprint servers, including potential loss of novelty and the chance of others publishing similar results first. They also note the possibilities of misuse by special interest groups and copyright-related queries from publishers. An additional concern revealed in the content analysis is the lack of peer review when a paper is uploaded to a preprint server. Participant number 7 from Information Science describes this problem:

'People can upload trash to ArXiv and because it wasn't peer reviewed, it's trash that's also cited and so you create trash upon trash and it creates an infinite loop, so it's very problematic.'

Participant number 5 from Information Sciences perceives the lack of peer review in a different manner

It doesn't require peer review, and in essence, its quality is done by citing it. If it's cited it's probably good and that's proof that it is scientifically sound and then i think that it is unclear how much this whole issue of peer reviews will last.

Puebla et al. (2021) discuss in their paper the multifaceted issue of peer review in preprint servers, that allows on one hand the rapid publication and dissemination of research findings to larger audiences and in the other hand it lacks the formal, structured process typical of traditional academic publishing, leading to concerns about the credibility and

quality of the disseminated research. Findings of this study parallel these concerns presenting different perspectives on this issue.

Conclusion

This study offers an exploratory analysis of scholars' views and perceptions regarding the role of preprint servers. Results indicate a general awareness of preprint servers and understanding of their terminology. Despite varying levels of familiarity, the study highlights common reasons scholars use preprint servers, such as rapid publication, early citations, and free public access. Participants identified several benefits, including quick dissemination, open access, establishing research precedence, receiving feedback, and fostering collaborations. Nonetheless, the lack of peer review remains a significant concern, though emerging practices in commenting and open peer review are beginning to address this challenge.

About the authors

Shir Aviv-Reuven is a PhD candidate at the Department of Information Science at Bar -Ilan university in Israel. She can be reached at shir.aviv-reuven@biu.ac.il

Jenny Bronstein is an Associate Professor in the Department of Information Science at Bar -Ilan university in Israel. She can be reached at Jenny.Bronstein@biu.ac.il

Ariel Rosenfeld is an Associate Professor in the Department of Information Science at Bar -Ilan university in Israel. He can be reached at Ariel Rosenfeld@biu.ac.il

References

- Aviv-Reuven, S. & Rosenfeld, A. (2021). Publication patterns' changes due to the covid-19 pandemic: a longitudinal and short-term scientometric analysis. *Scientometrics*, 126(8), 6761–6784. <https://doi.org/10.1007/s11192-021-04059-x>
- Berg, J. M., Bhalla, N., Bourne, P. E., Chalfie, M., Drubin, D. G., Fraser, J. S., ... & Wolberger, C. (2016). Preprints for the life sciences. *Science*, 352(6288), 899–901. DOI: [10.1126/science.aaf9133](https://doi.org/10.1126/science.aaf9133)
- Biesenbender, K., Toepfer, R., & Peters, I. (2024). Life scientists' experience with posting preprints during the COVID-19 pandemic. *Scientometrics*, 1–28. <https://doi.org/10.1007/s11192-024-04982-9>
- Chiarelli, A., Johnson, R. Pinfield, S., & Richens, E., (2019). Preprints and scholarly communication: An exploratory qualitative study of adoption, practices, drivers and barriers. F1000Research 8. DOI: [10.12688/f1000research.19619.2](https://doi.org/10.12688/f1000research.19619.2)

Desjardins-Proulx, P., White, E. P., Adamson, J. J., Ram, K., Poisot, T., & Gravel, D. (2013). The case for open preprints in biology. *PLoS Biology*, 11(5), e1001563. <https://doi.org/10.1371/journal.pbio.1001563>

Fraser, N., Mayr, P., & Peters, I. (2022). Motivations, concerns and selection biases when posting preprints: A survey of bioRxiv authors. *PLoS One*, 17(11), e0274441.

Kodvanj, I., Homolak, J., Virag, D., & Trkulja, V. (2022). Publishing of COVID-19 preprints in peer-reviewed journals, preprinting trends, public discussion and quality issues. *Scientometrics*, 127(3), 1339-1352. <https://doi.org/10.1007/s11192-021-04249-7>

Johansson, M. A., Reich, N. G., Meyers, L. A., & Lipsitch, M. (2018). Preprints: An underutilized mechanism to accelerate outbreak science. *PLoS Medicine*, 15(4), e1002549. <https://doi.org/10.1371/journal.pmed.1002549>

Maggio, L. A., & Fleerackers, A. (2023). Preprints in health professions education: raising awareness and shifting culture. *Academic Medicine*, 98(1), 17-20. DOI: 10.1097/ACM.00000000000005001

Penfold, N. C., & Polka, J. K. (2020). Technical and social issues influencing the adoption of preprints in the life sciences. *PLoS Genetics*, 16(4), e1008565. <https://doi.org/10.1371/journal.pgen.1008565>

Puebla, I., Polka, J., & Rieger, O. Y. (2021, February 18). *Preprints: Their Evolving Role in Science Communication*. <https://doi.org/10.31222/osf.io/ezfsk>

Sarabipour, S., Debat, H. J., Emmott, E., Burgess, S. J., Schwessinger, B., & Hensel, Z. (2019). On the value of preprints: An early career researcher perspective. *PLoS Biology*, 17(2), e3000151. <https://doi.org/10.1371/journal.pbio.3000151>

Sever, R., Roeder, T., Hindle, S., Sussman, L., Black, K.-J., Argentine, J., Manos, W., & Inglis, J. R. (2019). bioRxiv: The preprint server for biology. *BioRxiv*, 833400. <https://doi.org/10.1101/833400>

Strcic, J., Civljak, A., Gloznic, T., Pacheco, R. L., Brkovic, T., & Puljak, L. (2022). Open data and data sharing in articles about COVID-19 published in preprint servers medRxiv and bioRxiv. *Scientometrics*, 127(5), 2791-280. <https://doi.org/10.1007/s11192-022-04346-1>

Waltman, L., Pinfield, S., Rzayeva, N., Henriques, S. O., Fang, Z., Brumberg, J., & Swaminathan, S. (2021). *Scholarly communication in times of crisis*. Research on Research Institute, 10, m9. DOI: 10.6084/m9.figshare.17125394