

Description

Open Access Week 2024 is a good time to reflect on the challenges faced by an important component in the open access ecosystem — preprint servers.

Open Access and Preprints

The open access movement emerged in the early 1990s with the rise of the internet. Online scholarly publishing became a reality, rooted in the vision of providing universal online access to academic works without financial or legal barriers. Interestingly, preprints predate the open access movement. Offline preprints, in the form of physical mailed copies, have existed since at least the 1960s. The establishment of arXiv in 1991 ushered in the online preprint server movement, which created a new and easy way of sharing and discovering pre-peer review records.

Yin and Yang

Preprints neatly align with the principles of open access. They allow researchers the opportunity to share their findings almost instantly, offering a space to accelerate the pace of scientific discovery and dissemination. Technically, preprint servers offer researchers a chance to get quick feedback from the academic community and subsequently improve their work before formal peer review. Unfortunately, the scientific community has yet to fully embrace this promise of engagement, with community commentary present in only a minority of submitted preprints.

Preprints serve as early publications that allow the academic community to engage with research before a potentially lengthy delay in publication after peer review. However, the absence of vetting in preprints raises concerns about potentially flawed papers slipping through the cracks. This duality highlights the need for a balanced approach to leveraging the benefits of preprints while mitigating their inherent risks. The downside is that there is much potential for poor-quality science to be disseminated and circulated. The consequence of such dissemination is compounded by the fact that preprints are often picked up by news media, which then publicizes unvalidated results.

The COVID-19 pandemic is a good example of both the benefits and drawbacks of preprint servers (see also here). Preprint servers helped disseminate the latest findings quickly but also sparked controversy regarding the quality and reliability of the information shared. As researchers rapidly

Copyright: Enago Academy under the CC-BY-NC-SA 4.0 license

uploaded findings to preprint platforms, submissions were sometimes misinterpreted or misrepresented, leading to misinformation and conspiracy theories. This prompted widespread discussion on social media despite the papers being withdrawn shortly after publication due to scientific concerns.

New Threats & Solutions

So far, we have assumed that submissions are from honest, well-meaning researchers. However, we now live in a world where paper mills — nefarious for-profit entities producing legitimate-looking content — can produce submissions in large volumes, crossing ethical, legal, and moral boundaries. These entities can flood the preprint server ecosystem with bogus studies, making it harder to find genuine research.

Initiatives like Plan S and Plan U are shaking up the traditional publishing landscape by advocating for full and immediate open access to scholarly publications funded by public and private grants. They promote the use of preprints as a standard practice, which could further complicate matters in terms of submission volumes.

Publishers and journals are stepping up their game by adopting automated screening and integrity checks, which can screen submissions for aspects such as plagiarism, ethics and integrity, language quality, and metadata. It's a significant step forward in maintaining high standards in academic publishing, helping to weed out the sub-par submissions before peer review.

Preprint servers also implement screening processes to ensure that submissions are within scope and free from offensive or non-scientific content. The screening typically involves two main steps: an initial review by in-house staff to verify submission details and a subsequent evaluation by volunteer experts who assess the potential public health implications of the research (at least for biomedical preprint servers). While this process is not as rigorous as traditional peer review, it does help to maintain a level of quality and safety in the content shared.

There are many players in the field of automated screening and research integrity checks for publishers and journals, and more are emerging. So why not implement similar checks for preprint servers? These systems could help the staff and volunteer teams screen submissions in a faster, more efficient way, improving the quality of submissions that make it onto preprint platforms. Taking it a step further, preprints that have undergone screening could even be awarded a "green channel" for faster submission to traditional publishers, speeding up the publication process without compromising on quality. This would not only enhance the reliability of preprints but also lighten the load for publishers, allowing them to focus on what sets them apart from preprints: peer review.

The Way Forward

Copyright: Enago Academy under the CC-BY-NC-SA 4.0 license

To make this vision a reality, industry-wide collaboration is necessary. Researchers, publishers, and policymakers need to work hand in hand to integrate automated screening and integrity checks across all preprint platforms. Funding agencies should back initiatives that develop and fine-tune such tools, ensuring they are both effective and fair. Universities and research institutions also have a role to play by educating their researchers about the importance of maintaining integrity in preprints and encouraging the use of reputable and "verified" preprint servers.

Of course, these checks aren't a silver bullet. These systems will have problems associated with technology and automation. They may carry biases from their training data, lack applicability across different research fields, and exhibit suboptimal levels of accuracy. To address these issues, a balanced approach that combines automated checks with human oversight is essential. This way, the bulk of the screening workload is managed efficiently, while nuanced cases receive the attention they deserve from experienced reviewers.

Implementing automated checks also requires substantial financial and technical resources. The financial sustainability of preprint servers is evidently precarious, as many rely on temporary funding from philanthropic organizations and commercial partnerships. Implementing these checks with human oversight can increase the operational burden on servers, necessitating a balance between ensuring quality and maintaining rapid research dissemination. Potential solutions to address this include endowment funding from funders and support from the state for sustaining open research infrastructure. Smaller preprint servers may struggle even more with the costs, potentially leading to unequal standards across platforms. Preprint platforms could collaborate to share technical expertise and resources, reducing the burden on individual servers.

The Gates Foundation has taken the first steps toward this future with the launch of Verixiv, their new preprint server that runs submissions through a lengthy list of pre-publication checks. Currently, however, it is only accepting research funded by the foundation. It remains to be seen whether such initiatives can be scaled up for broader coverage and whether the necessary funding to do so will be available.

Despite these challenges, the integration of automated screening and integrity checks is likely the only scalable solution for the long term. Ideally, these measures will help operators uphold higher standards of research quality and integrity, and as the technology behind these checks advances, the screening workload for preprint server operators should also diminish.

Conclusion

To conclude, as preprints become an increasingly integral part of scholarly communication, it is imperative to bolster their screening processes with automation. Embracing such measures will ensure that the quest for open access and rapid dissemination does not come at the expense of integrity and trust. With these improvements, preprint servers can continue to thrive and serve their functions without succumbing to the deluge resulting from bad actors.

This article, "Upholding the Integrity of Open Science" was published by *The Scholarly Kitchen.*

Copyright: Enago Academy under the CC-BY-NC-SA 4.0 license

Category

1. Thought Leadership

Date Created
2024/11/18
Author
krishna_kumar

Copyright: Enago Academy under the CC-BY-NC-SA 4.0 license