Sciety and the Rise of the Overlay Service

Hannah Drury

Recent global events have given rise to explosive growth in preprint deposition. With the subsequent response from volunteers willing to evaluate the increased volume and with software development to connect the pipeline, is there a realistic way forward for a "publish first" model of publication?

Preprinting in the Biomedical and Life Sciences

Preprinting in general is not a new concept. Briefly, it entails an author or authors depositing an early version of their scientific manuscript to a public server as soon as they deem it ready, thereby disseminating the results to the widest possible audience earlier than would often be achievable, for example, by waiting for publication in a journal.

There are many well documented advantages to preprinting for both the authors and readers of scientific publications.[1] Authors retain control over when and where their manuscript is available, along with any priority claim to the reported results, while readers are rewarded with free and immediate access to the latest developments in their field. Furthermore, posting a preprint ahead of journal submission has been shown to increase the number of subsequent citations.[2]

Indeed, the notion of preprinting as a practice is gathering momentum, even within the biomedical and life science fields for which it is a relatively recent phenomenon. It will perhaps be news to no one that in recent years the number of biomedical preprints posted has steadily risen to finally reach a peak in response to the threat from COVID-19 (Figure 1).

While the preprint format arguably represents a major shift in the narrative of scientific publication, it is not without criticism. Most notably, preprints are yet to undergo the cross-examination levelled at those manuscripts published in journals; the virtues of which it is beyond the scope of this article to scrutinize. Detractors have argued that preprints not only dilute the scientific record, but also contribute to

HANNAH DRURY (0000-0001-8393-3153) is a Product Manager for eLife Sciences, currently working on the Sciety project.

the spread of misinformation amongst the public. Above all, most preprints lack the value afforded by the process of peer review itself, whereby experts in one's field offer feedback and recommendations for improvement.

Preprint Evaluation

If preprints are truly to be accepted as research artifacts by the scientific community and beyond, then they must be subject to the same scrutiny: not to determine if they are suitable for publication in a particular journal, but if their conclusions are supported robustly within the context of the manuscript itself.

To address this issue, several initiatives have begun to openly evaluate preprints, making their opinions publicly available. The advantage of reviewing already published work via a so-called "overlay journal" or other "overlay review service" in this way is that there is no need for the opaque editorial selection that characterizes traditional journal publication. This can foster more constructive reviews that are focused on the improvement of the science within the context of the article itself, in turn serving the interests of both authors and readers alike.

Open evaluations also mean readers can gain more context around the articles they are interested in, not only increasing their understanding but helping to validate and shape their opinions. Given that these readers will in turn be expected to contribute to the community as peer reviewers themselves, such transparency would also seem to offer a valuable learning tool, and this has borne out in several user interviews undertaken with researchers who are early in their scientific careers.

The format that these open evaluations take is as varied as the initiatives themselves, hence why the umbrella term "evaluation" itself is more appropriate than "review". Some efforts, it is true, are reminiscent of the journal process and produce multipart reviews from multiple experts overseen by a handling editor. Others are briefer; their output perhaps comprising answers to standardized questions about the manuscript's strengths and weaknesses. Still others take the form of reports generated from results produced by automated tools. Such diversification represented by different models is encouraging because it implies a spirit of experimentation, where popular models attract a greater readership or lead to a greater effect, and less popular options are forced to adapt to better serve the scientific community as a whole.

CONTINUED

Biomedical preprints per month through 2020-04
Sources: Jordan Anaya (PrePubMed), Naomi Penfold, EuropePMC, arXiv, Crossref, SSRN
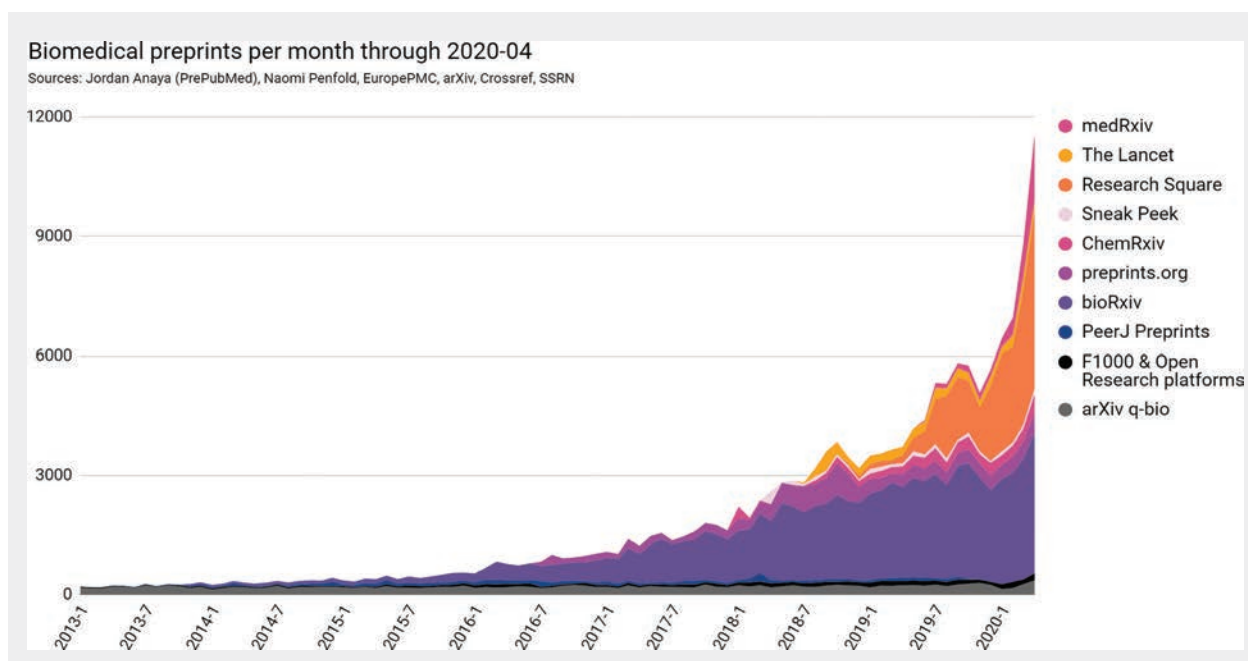
Figure 1. The growth of life science preprint deposition over time.[3]

One side effect of such diversification is a level of sporadic distribution. The unprecedented nature of these overlay activities means that there is no agreement on a suitable technical infrastructure to support them, so initiatives have been forced to either build their own or use tools that were not intended for such purposes. Furthermore, the evaluations themselves are scattered across the internet, often in public repositories or on individual websites, divorced from the preprints themselves as well as the full overlay service spectrum. While this doubtless allows each initiative to provide suitable context and transparent information about their processes to garner trust from their readers, those same readers struggle to find the evaluations in the first place.

Sciety as a Solution

Sciety emerged as a solution to the problems posed by the "publish first" model of scientific publication. It is a freely available online application that enables researchers, policy makers, funders, and other stakeholders to search for preprints and discover and share the kind of additional information that may have once been limited to journal publications. This information is presented in the form of an ever-evolving activity feed associated with each preprint (Figure 2), challenging the notion that the output of research is ever truly "finished". Activities currently include evaluations and version history but will be extended in line with Sciety's growing functionality to include richer information about the preprint's history.

Our vision for Sciety is for it to become the primary tool for navigating the growing preprint landscape. We aim to aggregate the outputs of these diverse services and make them more discoverable, increasing their reach while providing researchers with crucial trust indicators that help them decide in which preprints to invest their limited reading time. In this way, we question the traditional definition of "overlay" as an indicator of secondary tier review, and you will not see this term while navigating around the website.

Because the article is publicly available, there is no limit to the number of activities its feed may accumulate, or indeed the number of groups contributing evaluations. In this way, it is possible for a single preprint to attract multiple, diverse opinions that work together to inform a reader's overall interpretation. When interleaved with notifications of subsequent version updates, and even journal publication events, the feed becomes a living document of a preprint and its evolution.

Sciety Groups

The entities providing evaluations on Sciety are known as groups. We found this term to be the least overloaded with existing expectations, which has allowed us to be completely agnostic with regards to the processes each group follows and the outputs they produce.

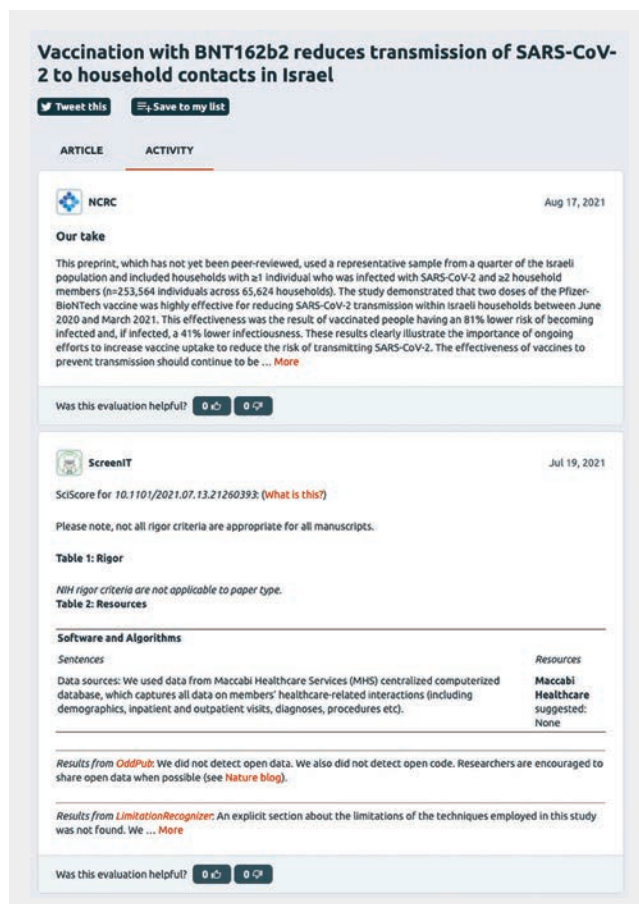Each group we partner with has its own homepage on Sciety, and we work with representatives from the group to

CONTINUED

Figure 2. A feed of evaluation activity for a preprint on Sciety.

ensure interested readers are provided with a rich summary, an explanation about its review model, and the people involved. This transparency should foster trust in the group's evaluation activities.

The Novel Coronavirus Research Compendium (NCRC),[4] an initiative that provided a rapid response to the rise in COVID-19-related literature, is one of the more recent groups to join Sciety. They currently maintain their own website to which their evaluations are published, and we use the same pipeline to display the evaluations in Sciety activity feeds alongside any other information we have relating to that preprint. An evaluation from the NCRC takes the form of answers to standardized questions related to the study, and this model is laid out on the group's Sciety homepage (Figure 3).

Since partnering with our first group last year, we have begun ingesting evaluations from 14 different groups of experts, bringing our total corpus to approximately 18,500 evaluations of over 14,000 preprints at the time of writing.

Curation

Beyond evaluation, journal publication provides researchers with an additional level of curation that helps them keep track of particularly important developments in their field. We feel it is important that Sciety supports the same intention to collect different preprints together based on a shared context or relationship, however that may be defined by the curator.

In speaking to researchers, we identified two distinct use cases for curation activity. The first we might call "personal curation," and it involves saving or bookmarking particular preprints for easy future reference. We anticipate supporting this form of curation by extending Sciety's current functionality, a single default list per user that may be added to one preprint at a time, to allow for dynamic organization across multiple lists with customizable titles and descriptions to provide additional context.

The second use case requires the same level of discovery and subsequent organization, but with the purpose of informing the wider community. Such "social curation" may require Sciety to help users to achieve additional goals such as sharing via a range of networks or media, or the ability to follow others' curated lists in order to be notified when new preprints are added. Collaboration is another important aspect of this form of curation, and we expect multiple experts will eventually wish to contribute to a single list. This level of shared responsibility is likely to increase the collection's longevity and persistent relevance for followers. In the future, there may well be additional scope to support computer-assisted curation, enabling a curator to add preprints from a list of suggestions based on titles already known.

What's Next?

Sciety can currently only be used to search for preprints deposited in bioRxiv or medRxiv, but we plan to surface content from other preprint servers in the future. Similarly, we continue to build upon the number of groups we work with, adding to the diversity of opinions within the network. We remain deliberately agnostic with regard to the review model and subsequent output each group employs and produces, meaning Sciety poses no barrier to entry besides a commitment to transparency.

As the application continues to evolve, we will bolster our outreach efforts to help build a community of researchers, policy makers, and funders who share the same passion for a publish-first model of scholarly communication. Ultimately, Sciety's success will depend on behavioral and cultural changes that can only ever be supported, rather than driven, by software development.

The Future of Preprint Evaluation

It seems that recent global events have worked to shift the perception of preprints in the research community's collective consciousness. Now that we have seen what is possible ahead of, and even instead of, journal publication, it is not certain that we can return to that previous state of low-level tolerance. Preprints are becoming first class

CONTINUED

Figure 3. The Novel Coronavirus Research Compendium's (NCRC) group page summarizing the group's process and recent activities.

citizens of scientific output, particularly when considered alongside the scrutiny of so many overlay services, and I predict that the trend will only continue to flourish.

References and Links

1. https://asapbio.org/wp-content/uploads/2021/05/ASAPbio-what-are-preprints-english-aw.pdf

2. Fu DY, Hughey JJ. Meta-research: releasing a preprint is associated with more attention and citations for the peer-reviewed article. eLife 2019;8:e52646. https://doi.org/10.7554/eLife.52646

3. Polka JK, Penfold NC. Biomedical preprints per month, by source and as a fraction of total literature (version 3.0). [Data set]. Zenodo 2020. http://doi.org/10.5281/zenodo.3819276

4. https://sciety.org/groups/62f9b0d0-8d43-4766-a52a-ce02af61bc6a