



EDITORIAL

Preprints optimize research communication

Los Preprints optimizan la comunicación de investigaciones

[Abel L. Packer](#)^{1*}

¹Fundação de Apoio à Universidade Federal de São Paulo (FAPUNIFESP). Sao Pablo, Brasil.

*Corresponding author: abel.packer@scielo.org

Received: 05/07/2021. Approved: 07/07/2021

How to cite this article

Packer AL. Preprints optimize research communication. Rev haban cienc méd [Internet]. 2021 [cited 20(4):e4294. Available from: <http://www.revhabanera.sld.cu/index.php/rhab/article/view/4294>];

The Preprint is a manuscript that is publicly available in open access on a website known as Preprint Server before being published by a scientific journal. It is an option provided to authors who benefit from quickly and formally communicating their research before or in parallel with submission to journals for undergoing peer review and publication, if accepted. It is also an option for journals to deposit as preprints the manuscripts already approved that are in the editing and publishing process. This conception that we use in the SciELO Program is slightly more flexible than the one defined by the Committee on Publication Ethics (COPE): "A preprint is a scholarly manuscript posted by the author(s) in an openly accessible platform, usually before or in parallel with the peer review process".⁽¹⁾

Thus, a preprint is one of the possible communication objects of a research with two relevant periods. First, as an initial and almost always unpublished research communication and, second, as a historical version of the final article approved, edited, and published by a journal, also identified by version of record. Both versions are permanently valid, identified by a digital object identifier (DOI) and indexed by Google Scholar and other indexers. As a good practice, the preprint, when accessed, should provide a link to the final article, when there is one.

This understanding and current preprint practice as a form of advanced communication of research results for the advancement of knowledge is associated with digital texts and the infrastructure of information and communication technologies and web interoperability. Importantly, the essence of preprint design and functionality goes back more than 50 years, as said by Matthew Cobb when, in 1961, he rescued the revolutionary creation of the National Institutes of Health (NIH) Information Exchange Groups (IEG) that promoted the circulation of paper preprints on different topics and began to operate with 3,600 researchers worldwide and more than 2,500 manuscripts. This development, however, ended six years later, mainly due to the strong reaction of major journals of commercial publishers and scientific societies which considered that accepting preprints would constitute a threat to their prestige and finances.^(2,3)

The preprints reappeared renewed in the 1990s, thirty years after the creation of the IEGs at the Los Alamos National Laboratory, under the direction of Paul Ginsparg with the establishment of the e-mail service and later with the arXiv preprint platform or server on high energy physics. The growing success of arXiv was evident by means of the progressive expansion of its thematic coverage. Operated by Cornell University since 2001, arXiv currently covers eight subject areas and has accumulated 1,8 million preprints by the end of 2020.^(4,5) It is considered to be the most important and recognized preprint server. For many years, arXiv has presented itself as an isolated preprint solution alongside failed attempts in other areas.



Resistance to preprints placed the broad field of Biological, Medical and Biomedical Sciences as the last frontier of resistance to the systematic adoption of preprints in the scholarly communication ecosystem, which finally began to be overcome in 2013 with the launch of three important initiatives: the bioRxiv preprint server for Biological Sciences (<https://www.biorxiv.org/>) at the Cold Spring Harbor Laboratory (CSHL) in November 2013, followed by the medRxiv preprint server for Medical and Biomedical Sciences (<https://www.medrxiv.org/>) in June 2019, also owned by CSHL, but operated in collaboration with Yale University and the British Medical Journal (BMJ), self-proclaimed as a global provider of knowledge in healthcare, and, as of 2016, of the non-profit organization ASAPbio (Accelerating Science and Publication in Biology, <https://asapbio.org/>).

The bioRxiv and medRxiv servers have been very well received, as shown by the growth in the number of deposited preprints: bioRxiv went from an average of 800 preprints in the first half of 2017 to reach 3,000 in 2021; while medRxiv has received, on average, monthly deposits of over 1,150 manuscripts in 2020 and in the first half of 2021.^(6,7) Both servers have been playing a decisive role in communicating research on SARS-CoV-2 and COVID-19, with a total of 17,000 preprints deposited between January 2020 and June 2021 and, therefore, contributed to overcoming resistance to its use.⁽⁸⁾

The consolidation of the use of preprints in Biological, Medical and Biomedical Sciences, considering the success of bioRxiv and medRxiv, has been enriched with the action of ASAPbio in concepts, policies, methodologies, ethics, and indicators on preprints.

The progressive presence of preprints in the global ecosystem of science communication is revealed in the ASAPbio directory of 56 preprint servers (<https://asapbio.org/preprint-servers>)⁽⁹⁾ and on Wikipedia's list of 63 active preprint repositories (https://en.wikipedia.org/wiki/List_of_preprint_repositories) in June 2021. With a greater or lesser volume of documents, preprint servers are present in all regions with scientific output: Africa, Asia, Europe, Latin America, and North America.

The servers are characterized by a wide thematic coverage, whether multidisciplinary or from disciplines of Physical and Technological Sciences, and Social and Human Sciences. They are operated on a non-profit basis: academic institutions and research communities (arXiv, AfricArxiv, bioRxiv, ChinaXiv, medRxiv, etc.) or by research funding agencies (Gates Open Research, SciELO Preprints, Wellcome Open Research, etc.), scientific societies (ChemRxiv, PsyArXiv, etc.), major commercial publishers (Authorea/Wiley; F1000 Research; Taylor & Francis; PeerJ PrePrints; O'Reilly Media & SAGE Publishing; Preprints.org; MDPI, Research Square; Spring Nature, SSRN-First Look, Elsevier, etc.), as well as by individual journals or article platform. Most servers are relatively young and have a low number of preprints: 50% of the servers listed are three years old or less in operation and each of them accumulates less than 500 preprints. The SciELO Preprints server, a reference in Latin America, was created in April 2020; it is on the list of youngest servers and, in June 2021, it has accumulated just over 1,000 deposited preprints.

Preprint servers fulfill essential functions in managing the manuscripts that were progressively specified by the arXiv server and by those who followed it over the last 30 years. The servers are equipped with specific technological capabilities, but typical of a bibliographic information system: quality control, management of full texts and their metadata, dissemination, and availability of performance metrics. Quality control is applied by specialized editors who organize themselves in councils and networks to verify the text structure according to the sections required in the manuscripts sent to scientific journals, the research relevance and the results obtained. Text management includes DOI assignment, possibility of creating new versions to improve preprints, indexing from metadata extracted from full texts, and digital preservation. Dissemination operates indexing, based on metadata extracted from full texts, exposure for indexing by Google Scholar and other indexers, and connection of preprints with the final version of the article when it is published in a journal. Metrics provisioning includes number of downloads, presence in social media and citations received.

The servers operate free of charge for preprint authors and open access for users under Creative Commons Licenses, defined by the authors or required by the server. The software platforms of most preprint servers are proprietary, and few are publicly available, as is the case of the Open Preprints Systems (OPS/PKP) that operates SciELO Preprints.

Preprints provide authors, especially young researchers, with new features and benefits.⁽¹⁰⁾ First, authors gain greater control and responsibility in communicating their research, with the benefits of quickly sharing results, ensuring authorship of findings, promoting their visibility, and enhancing manuscripts with new versions before submitting them to a journal. Second, they allow them to participate in the peer review process with known scientific identities and production and preprint performance indicators. These features and benefits tend to overcome fears regarding acceptance by journals.

Preprints contribute to the alignment of journals with open science practices. The vast majority of journals from commercial publishers and scientific societies already accept manuscripts previously deposited as preprints, as indicated in the list of scientific publishers, in accordance with Wikipedia's preprint acceptance policies (https://en.wikipedia.org/wiki/List_of_academic_publishers_by_preprint_policy).

In the SciELO Network, the indexing criteria indicate the end of 2022 as the deadline for journals to align their policies with the open science *modus operandi* that includes the acceptance of preprints for evaluation and publication, which implies abandoning the policies of publishing only the results of unpublished researches and use only the double-blind peer review. The expectation is that journals will contribute decisively to enable the benefits that preprints, as an open science practice, bring to researchers and to the progress of science.

The growing use of preprints, as an established practice of scientific communication, has coexisted in recent years with the advancement of the open science movement and has been identified as one of its practices, along with the availability in web repositories of data, computer program codes, and other materials from preprint texts, published articles, and the opening of the manuscript review process. Thus, the use of preprints is one of the dimensions of the open science construct in favor of transparency in the entire process of conducting and communicating research, expanding cooperation between researchers, facilitating research replicability and strengthening the social function of research and knowledge, which is expected to be formalized in national and institutional policies for the promotion and evaluation of research, as indicated in the UNESCO open science recommendation document.⁽¹¹⁾

REFERENCES

1. COPE Council. COPE Discussion document: Preprints. March 2018 [Internet]. United Kingdom: COPE Council; 2018 [Cited 25/05/2021]. Available from: https://publicationethics.org/files/u7140/COPE_Preprints_Mar18.pdf
2. Cobb M. The prehistory of biology preprints: A forgotten experiment from the 1960s. PLoS Biol [Internet]. 2017 [Cited 25/05/2021];15(11):e2003995. Available from: <https://doi.org/10.1371/journal.pbio.2003995>
3. Nassi Calò L. La (pre)historia de los preprints en ciencias biológicas [Internet]. Brasil: SciELO en Perspectiva; 2017 [Cited 25/05/2021]. Available from: <https://blog.scielo.org/es/2017/12/20/la-prehistoria-de-los-preprints-en-ciencias-biologicas/>
4. Álvarez GR, Caregnato SE. Prácticas de comunicación científica en la Física de Altas Energías: potencialidad de los preprints [Internet]. Brasil: SciELO en Perspectiva; 2018 [Cited 25/05/2021]. Available from: <https://blog.scielo.org/es/2018/02/15/practicas-de-comunicacion-cientifica-en-la-fisica-de-altas-energias-potencialidad-de-los-preprints/#.YOLrouhKhdg>
5. Cornell University. arXiv Annual Report 2020 [Internet]. NuevaYork: arXiv; 2020 [Cited 25/05/2021]. Available from: https://static.arxiv.org/static/arxiv.marxdowm/0.1/about/reports/2020_arXiv_annual_report.pdf
6. Vale RD. Accelerating scientific publication in biology. PNAS [Internet]. 2015 [Cited 25/05/2021];112(44):13439-46. Available from: [doi:10.1073/pnas.1511912112](https://doi.org/10.1073/pnas.1511912112)
7. Polka J. Preprints as a complement to the journal system in biology. Information Services Use [Internet]. 2017 [Cited 25/05/2021];37(3):277-80. Available from: <http://doi.org/10.3233/ISU-170849>
8. Fraser N, Brierley L, Dey G, Polka JK, Pálffy M, Nanni F, *et al.* The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. PLoS Biol [Internet]. 2021 [Cited 25/05/2021];19(4):e3000959. Available from: <https://doi.org/10.1371/journal.pbio.3000959>
9. Kirkham JJ, Penfold NC, Murphy F. Systematic examination of preprint platforms for use in the medical and biomedical sciences setting. BMJ Open [Internet] 2020 [Cited 25/05/2021];10:e041849. Available from: <http://doi.org/10.1136/bmjopen-2020-041849>
10. Sarabipour S, Debat HJ, Emmott E, Burgess SJ, Schwessinger B, Hensel Z. On the value of preprints: An early career researcher perspective. PLoS biology [Internet]. 2019 [Cited 25/05/2021];17(2):e3000151. Available from: <https://doi.org/10.1371/journal.pbio.3000151>
11. UNESCO. Draft text of the UNESCO Recommendation on Open Science [Internet]. Paris: UNESCO; 2021 [Cited 25/05/2021]. Available from: https://unesdoc.unesco.org/ark:/48223/pf0000376893_spa.locale=en

Conflict of interests

The author has no conflict of interests to declare.