XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTIONAL METHODS IN ENGENEERING

Preprints as an emerging practice in Brazilian scientific communication: analysis of adoption by institutions and areas of knowledge

Dênnyo A. Souza[1], Thiago M. R. Dias[2], Davi F. S. Guimarães[1], Patrícia Mascarenhas Dias[3]

[1]*Technical High School Program, Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG)*
Av. Amazonas, 5.253, Nova Suiça, Belo Horizonte, 30.421-169, Minas Gerais, Brazil
dennyoarcanjo@gmail.com, guimaraes.davi.fs@gmail.com
[2]*Dept. of Computing, Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG)*
Av. Amazonas, 5.253, Nova Suiça, Belo Horizonte, 30.421-169, Minas Gerais, Brazil
thiagomagela@cefetmg.br
[3]*Dept. of Computing, Universidade Estadual de Minas Gerais (UEMG)*
Av. Presidente Antônio Carlos, 7545, São Luiz, Belo Horizonte, 31.275-083, Minas Gerais, Brazil
patriciamdias@gmail.com

Abstract. The Open Science movement has encouraged the use of preprints as a way to accelerate scientific communication, increase transparency in research processes, and expand access to knowledge. This study investigates the production of preprints registered in the Lattes Platform, analyzing approximately 12.8 million articles published in journals declared by almost 9 million Brazilian researchers. Based on the identification of 5,920 records explicitly classified as preprints, it was possible to characterize the profile of the authors, the areas and major areas of knowledge involved, the institutions to which they are affiliated, and the temporal evolution of these publications. The results reveal a recent growth in the adoption of preprints, although their presence in national scientific production remains incipient. Most authors with preprints are affiliated with public universities and work in the areas of Exact, Biological and Health Sciences. The analysis also highlights a positive correlation between the use of preprints and open science practices, such as the use of the ORCID identifier. The study contributes to understanding how preprints have been incorporated into Brazilian scientific practice and points out ways to strengthen them as a strategic component of open scientific communication.

Keywords: Scientific Production, Preprints, Lattes Platform.

1 Introduction

Preprints represent one of the fundamental pillars of open science, enabling the rapid dissemination of knowledge. In emergency contexts, such as the COVID-19 pandemic, this publishing modality has played an essential role in the rapid communication of discoveries about prevention, treatment, and epidemiology of the disease. In addition to accelerating the circulation of knowledge, preprints expand access to scientific information, promoting greater collaboration among researchers.

The publication of preprints favors accelerated scientific communication and open access to research results, allowing discoveries to be disseminated and used by other researchers, contrasting with the traditional system, which is closed, slow, and restrictive in the peer review of articles [1].

Despite the growing adoption of preprints in various fields of knowledge, there is still little information about their systematic use by Brazilian researchers, especially regarding the formal registration of this scientific production in national databases. This gap is significant, as this modality can expand the impact of national research and strengthen open research practices. Thus, the question arises: how are Brazilian researchers

adopting this form of publishing? This study aims to identify, quantify, and characterize preprints registered on the Lattes Platform, analyzing areas of activity, affiliated institutions, and the temporal evolution of publications. By mapping this adoption, the research contributes to the understanding of the national open science landscape, providing support for initiatives that encourage the use of preprints as a tool for democratizing knowledge.

2 Methodology

The research methodology consisted of analyzing data from CVs registered on the Lattes Platform. The LattesDataXplorer framework [2] was used to collect and process CV data.

A total of 12,881,308 records of articles published in journals were analyzed, including almost 9 million CVs. To identify preprints, the criterion adopted was to explicitly indicate that the work was a preprint in the "Nature" field.

From this identification, a total of 5,920 articles classified as preprints were extracted. Based on these records, the authors' profiles, their fields of expertise, and their institutions were characterized. An analysis was also conducted of the temporal evolution of the number of registered preprints since the first one was identified, dating back to 1973.

Additionally, data on the presence of persistent identifiers such as ORCID were extracted, allowing us to observe the degree of adherence to practices associated with open science among preprint authors. Among the almost 9 million CVs analyzed, 3,177 authors submitted at least one registered preprint, of which 1,833 had an ORCID identifier associated with their CV.

3 Results

The results reveal a still nascent but growing adoption of preprints among Brazilian researchers. Of the more than 12 million articles analyzed, only 5,920 were registered as preprints, representing less than 0.005% of the total scientific production declared in Lattes CVs. Despite this low proportion, significant growth has been observed over time, especially since the 2010s, as evidenced by the graph of publications by year (Figure 1).

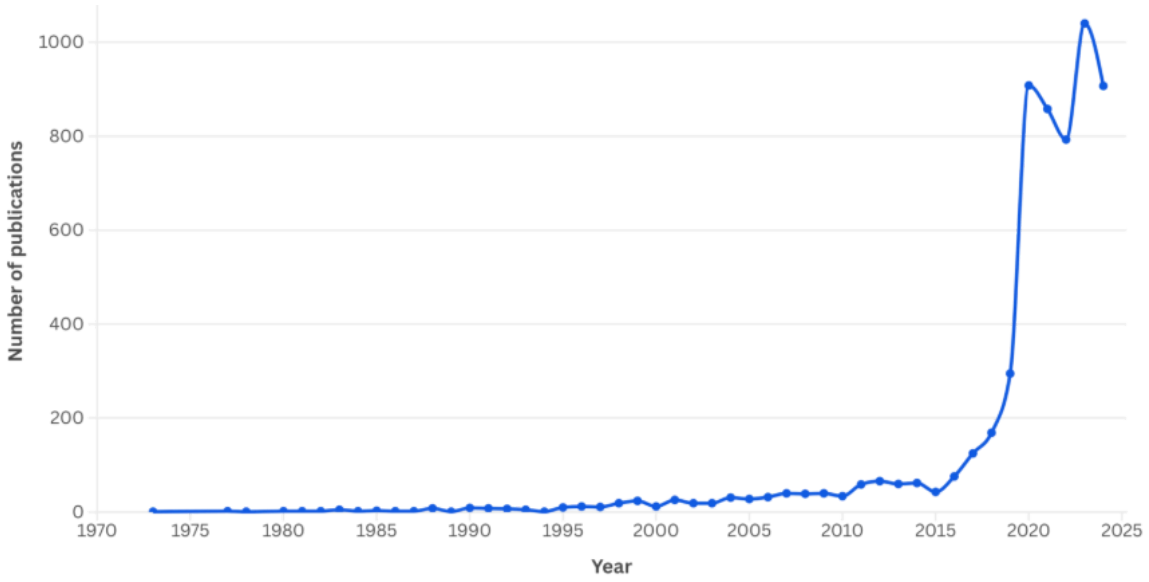


Figure 1 – Temporal evolution of the number of published preprints.

A temporal analysis of preprint production reveals a marked increase, accelerating since the mid-2010s. Although the first preprint registered on the Lattes Platform dates back to 1973, there was a prolonged period of

low incidence until approximately 2015, when the number of registrations began to grow more consistently. This increase coincides with the consolidation of thematic and multidisciplinary preprint repositories, such as bioRxiv, arXiv, and SciELO Preprints [1], as well as the expansion of the debate surrounding Open Science in Brazil and worldwide. The graph of publications by year indicates that, between 2019 and 2025, there was a significant increase in the number of registered preprints, one of the reasons being the greater acceptance of preprints by funding agencies and scientific journals.

Regarding author profiles, it was identified that 3,177 researchers had registered at least one preprint in their CVs. Most authors with preprints work in the broad fields of Exact and Earth Sciences, Biological Sciences and Health Sciences, with particular emphasis on subfields such as Geometry and Topology, Algebra, Applied Mathematics, Physics of Elementary Particles and Fields and General Physics. Physics and Biology were among the first fields of knowledge to record initiatives involving preprint use [1], which may explain this performance.

The ratio of preprints to total publications also varies significantly across fields and institutions. Figure 2 shows the distribution of authors by field. Regarding institutional affiliation, the authors with the highest absolute number of preprints are concentrated primarily in federal teaching and research institutions, highlighting the role of these institutions in the adoption of open practices. The University of São Paulo (USP), the Federal University of Rio de Janeiro (UFRJ), and the State University of Campinas (Unicamp) stand out, with the highest numbers of preprint authors, respectively.

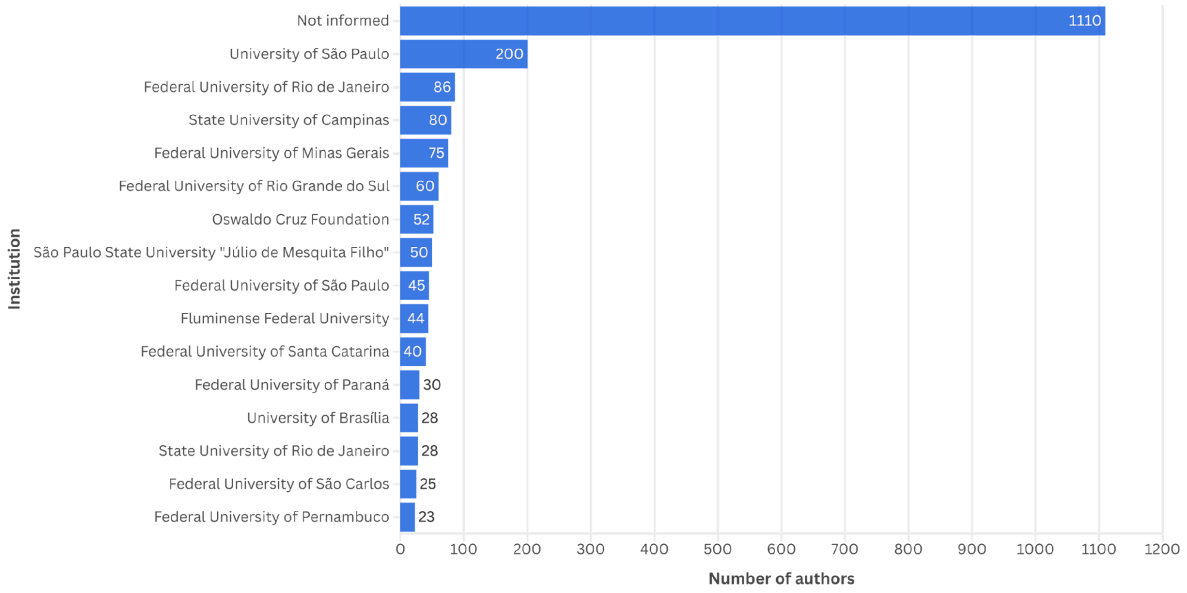


Figure 2 – Distribution of authors by affiliation.

These institutions are among those with the highest absolute number of authors who have registered preprints on the Lattes Platform. This predominance may be associated with both their greater research tradition and their involvement in international networks and programs that encourage Open Science. Recent studies confirm Fiocruz as a leading institution in Brazilian federal government initiatives on open science [3], which corroborates its position among researchers with the highest use of preprints.

The ORCID analysis indicated that approximately 58% of authors with preprints have this identifier, suggesting a positive correlation between open science practices and the use of persistent identification tools.

4 Conclusion

This study provides an unprecedented overview of the presence and characteristics of preprints in Brazilian scientific production, as reported on the Lattes Platform. The data point to a still timid, but clearly growing,

adoption of this type of publication among Brazilian researchers. Identifying preprints in CVs not only allows us to map their distribution but also signals the potential for integrating these practices into national scientific evaluation systems.

Publishing preprints represents an important strategy for aligning with the principles of Open Science, promoting the rapid and transparent dissemination of results, fostering reproducibility, and stimulating scientific collaboration. The use of identifiers such as ORCID by a significant portion of the authors analyzed reinforces the commitment to traceability and visibility of production. Thus, the results presented here not only characterize the practice of preprints in Brazil but also offer relevant insights for formulating policies to encourage their adoption, contributing to a more open and collaborative scientific culture that is aligned with contemporary scientific challenges. Research developments will consider the comparative analysis of annual distribution, repositories used, author gender, and preprint languages by field of knowledge.

5 Acknowledgments

This work was supported by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG)

References

[1] Principe, E. (2021). Prática da ciência aberta: os preprints em movimento. Páginas a&b. S.3, nº especial ConfOA. pp 59-70.

[2] Dias, T. M. R. (2016). Um Estudo Sobre a Produção Científica Brasileira a partir de dados da Plataforma Lattes. 2016. 181 f. Tese (Doutorado) - Curso de Programa de Pós-graduação, Modelagem Matemática e Computacional, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte.

[3] Dias, C. G. S. (2024). Políticas públicas e institucionais de Ciência Aberta no Brasil no período recente (2020-2023). In: Encontro Brasileiro de Bibliometria e Cientometria, v.9, Brasília. EBBC, pp 1-10.