A new framework for the French Open Science Monitor

Anne L'Hôte[1] and Eric Jeangirard[1]

[1]French Ministry of Higher Education, Research and Innovation, Paris, France

November 2021

Keywords: open science, open access, unpaywall, clinical trials

1. Introduction

2. Method

2.1 Publications

2.1.1 Open access dynamic

From the first edition of the French Open Science Monitor, it was clear that the estimate of the open access rate should not be static but should try to capture the dynamics of the opening (Jeangirard 2019). Indeed, the 0-day open access exists but we cannot assume it represents the totality of the open access. Therefore, for a given set of publications, say the publications published during the year Y, it makes sense to measure the open access rate at different point in time, for example at some moment in year Y+1, Y+2 …

To do so, it becomes necessary to historicize the database containing the open access information. So, instead of maintaining a database that keeps track of the opening of each publication, like Unpaywall is doing, we have to make regular snapshots of the whole Unpaywall database. Each snapshot is used as an observation date to measure the open access rate. It is important to note that this method natively embeds the potentials open access discovery errors from the underlying Unpaywall database, that can be false negative (a publication is actually open at this point in time but it not detected) or false postive (wronly seen as open whereas it is closed).

This method of analysis therefore reveals two temporal dimensions: publication dates and observation dates. Obviously, the observation date must be after the

publication date. To avoid that the proliferation of possible analyses blurs the message, we propose to look mainly at two elements :

- A main statistics that is the **1Y Open Access rate**: it represents the open access rate of the publications published during year Y and measured (observed from the snapshot of the OA discovery database) at point during year Y+1 (generally in December if the data is available).
- Also the **shape of open access curve** (open access rate function of the publication year). From an observation date to another, the evolution of the shape gives an insight of the speed of opening. An inverted U curve means the open access rate is lower for recent publications. Flat curves means the open access rate is the barely the same, whatever the age of the publication. Increasing curve instead would mean recent papers are more and more open.

2.1.2 Discipline impact

2.1.3 Publishers and dissemination platforms strategies

2.1.3.1 Identification of the dissemination platforms

2.1.3.2 Business models

2.1.3.3 Licences

2.1.3.3 Article Processing Charges (APC) estimation

2.1.4 The role of the open repositories

2.1.5 Other impacts on open access

2.1.5.1 Funding

2.1.5.2 Main authors affiliation country

2.1.5.3 Authors affiliations

2.1.6 Data collection system and architecture

2.2 Clinical trials and observational studies

3. Results

4. Discussion and conclusion

4.1 Findings

4.2 Limitations and future research

Software and code availability

Data availability

References

Jeangirard, Eric. 2019. "Monitoring Open Access at a National Level: French Case Study." In *ELPUB 2019 23d International Conference on Electronic Publishing*. OpenEdition Press. https://doi.org/10.4000/proceedings.elpub.2019.20.