Review of: "Availability of Open Citations from Open Journals in Crossref v1"

Constance Dami[1]

1 University of Bologna

Introduction

This review is made in the context of the 2021-22 Open Science course by S. Peroni at the University of Bologna. It is aimed at reviewing the first version of the protocol established by Davide Brambilla, Chiara Catizone and Giulia Venditti for their research on the availability of Open Citations from Open Journals in Crossref.

The protocol has been published on protocols.io on April 12, 2022 at this link.

This review follows the PLOS ONE guidelines for reviewing a Study Protocol.

It should be noted that the reviewed protocol is only at primary stage of research and is only a draft of what the final one for the author's research will be.

Research questions.

The research questions are clearly stated in the abstract preceding the methodology: The authors' research is about articles from Open Access journals in DOAJ and the availability of metadata of these articles in Crossref.

More precisely, the authors want to:

- Scout the availability of the articles' reference lists,
- Examine the presence of DOIs and the entities responsible for their specification
- Making further analysis to verify the distribution of the provenance of the articles' DOI.

In order to insert the research in an academic context, it would be relevant to present what is the state of research in this field and how the findings on this question will benefit to the comunity. I would advise the authors to insert some literature on the subject as references, and some links to the organizations mentioned (DOAJ, Crossref,…) to allow the reader to find more information on the subjects if needed.

Protocol

Although the reviewed protocol is only the first version, not yet providing full details of the steps, each one is described in a precise manner. The sequence of steps makes sense, starting from data gathering through the concerned organization's APIs to publishing the data, followed by data analysis and visualization. Links to the mentioned tools are provided, except in the case of step 3 which could be transformed in a link.

The file formats obtained in output are also clearly stated at each step.

On another note, the PLOS ONE guidelines suggest to include the word "Protocol" in the title of the manuscript.

Methodology

As previously stated, the protocol is easy to follow, but there is a minor difference between what has been announced in the abstract and what is actually carried out in the protocol. Indeed, the abstract states that the analysis will be carried out by querying the DOAJ API, while the first step of the protocol mentions the public data dump of the journals. The abstract or the first step should be revised such as to have one corresponding to the other.

Some steps could already be more precise in order to make the protocol as replicable as possible. For example, after step 3, there is no indication on where the information about the agency responsible for the DOIs will be stored: will it be added to one of the preexisting two datasets or will it be stored in a new one? Such a precision was present in the previous steps and should be added here to ensure that someone replicating the experiment would obtain the same datasets in the same formats as the authors at each step.

The final part about data analysis and data visualization is still relatively vague, which is very understandable at this stage of research, yet the idea behind it is clear and makes sense. I would suggest adding a link to the test analysis mentioned, to make it more explicit to a non-initiated reader.

Data

The data used for this research is extracted from publicly available repositories and queryable APIs, so there are not limits to reproducibility. At this stage of research there are no outputs, but there is no indication in the protocol on how and where the data found will be published and this information should be added.

Conclusions

In conclusion, the protocol is clear and provides a sound methodology.

Some minor changes would already improve its quality and make it clearer before the next version at a more advanced step of the research, such as referring to literature to put the new research questions in a scholarly context, providing more links to tools or organizations mentioned, and making sure that the methodology stated in the abstract reflects the reality of the steps described in the protocol.

References

- Davide Brembilla, Chiara Catizone, Giulia Venditti 2022. Availability of Open Citations from Open Journals in Crossref. https://dx.doi.org/10.17504/protocols.io.kxygxz7ywv8j/v1
- Ivan Heibi, Silvio Peroni 2022. A methodology for gathering and annotating the raw-data/characteristics of the documents citing a retracted article. Version 2. https://dx.doi.org/10.17504/protocols.io.bqqumvww
- PLOS ONE guidelines for Reviewing Study Protocols