

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

Alessandro Bertozzi¹

¹ University of Bologna

Potential competing interests: The author(s) declared that no potential competing interests exist.

Introduction

Venditti. This protocol was developed in the Open Science research area, during the Open Science 2021/2022 course at the University of Bologna. The full protocol is available at the following link.

Before proceeding to the heart of the review, some initial clarification. I am currently a student in the Digital Humanities and Digital Knowledge course at the University of Bologna. My level of experience in the field of research treated is at the same level as the authors. Therefore, the writer of the review has the same level of experience as the person who produced the research. Furthermore, while reading the review, it is good to remember that the protocol being analysed is a first version of the final product. Therefore, hasty judgements will be avoided.

The review will be structured as follows: I will deal with the individual points highlighted by the authors and express the positive aspects, indicating possible critical points. As mentioned above, given the state of progress of the research, the judgements will be recommendations on how to improve the presentation of the content and not judgements on the completeness of the research.

The abstract part

The protocol begins by presenting the abstract of the research: it is correctly formulated, giving a general idea of what the research is going to cover and what the objectives are. The only suggestion I would make is to better specify the context and motivations of the research. Furthermore, I would add some references to some external resources, in order to give the reader the possibility to consult directly the pages mentioned within the section.

The first section: gathering data

The central part of the protocol consists of the various steps taken to obtain the search results. There are four main sections in total, divided by other subsections. This nested division allows for easy and fluent reading. Above all, it allows a quick visualisation of the central tasks. The only two observations to be made are in the numbering of the tasks and the clarification of the variables taken into account.

The first problem can be found in point 2: the text is comparable in importance to the texts of the following sub-points. So I recommend respecting the priority marked by the numbering. The text, in fact, should be a summary of its subsequent sub-tasks.

The other problem concerns points 1 and 2. The second point specifies the use of ISSNs to send a request to the API. However, it does not specify that the ISSN is part of the unique identifiers downloaded previously. I recommend adding

this specification to make it easier for non-experts to read.

Other sections

For convenience I will condense the other steps into one section. As reiterated above, the protocol is still in the state of rough draft. So the state of development is inevitably still incomplete. From the point of view of the structure, the steps are well defined. Moreover, there are important clarifications regarding the data format and possible implementation of programming languages.

The only note that I feel like specifying is an advice. It would be correct to specify the reasons for certain choices: in point 6 it is specified a formula to calculate the "effect size", but without any reflection on the reasons for this choice. Therefore, in order to illustrate precisely the operational steps to obtain the research objectives, it would be valuable to specify why this particular formula was chosen at the expense of others. In addition, it would be good to add an external reference to illustrate what precisely this type of formula does.

Conclusion

For a first version, the analysed protocol is well structured. The only two aspects that needs to be taken into account, in addition to the points specified above: a more structured link with all external resources mentioned, or used, during the writing of the research and a more specific description of why certain choices were made.