

Review of: "La Chouffe DMP New"

Umut Kucuk¹

¹ University of Bologna

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

This review addresses La Chouffe DMP New. Which it is the very first version of a data management plan (DMP) for an ongoing research project, authored by Giulia Venditti (orcid:0000-0001-7696-7574); Davide Brembilla (orcid:0000-0002-9481-5053); Chiara Catizone (orcid:0000-0003-2445-2426), under the [Open Science](#) course held by [Prof. Silvio Peroni](#), at the University of Bologna. Before involving into the details of the produced DMP it is crucial to give some preliminary information about the aim of the research, research questions and the purpose to have a DMP.

The ongoing project that held by the above mentiond authors has the purpose to answer the following questions: How many DOAJ articles are included in Crossref, how many of these articles contain information about reference lists, how many of these references are DOI specified, How many of these DOIs are specified by publishers and how many are specified by Crossref? Starting from these research questions, the researchers aims to assess the data related to the articles and possessing a notion about the communication between the aggregators regarding the information sharing.

For a project that will consist of different types of data, it is essential to create a data management plan right at the beginning of the project. It is a very important approach for the whole project to make clear definitions of the data collected, obtained or created, and to create a realistic roadmap of how this data will be managed and how it will be stored. In addition to these, answering the questions of the standards of the data and how this data will be preserved during and after the project are among the purposes of the data management plan. Readers who require more detailed information on the subject are strongly recommended to visit the [github repository](#) of the Open Science course, which also includes the whole project.

In the dataset description section, the reason for collecting the data and the reason for its creation are briefly explained. A brief explanation is perfectly acceptable as it is about an ongoing project and the first version of the DMP. However, it is beneficial to enrich these particular part with more detailed description that may provide required information also for the scholars that have a limited knowledge about the domain or for the ones who are not familiar with these particular concepts.

Types, formats, and sizes of data are clearly and satisfactorily expressed. Information on who the datasets can be useful for is given. At this point, brief predictions about the data likely to emerge at the end of the project could be included. Apart from this, a very realistic and reasonable dataset description section has been created according to a research that

has not been completed yet. The scope of the dataset belonging to the software is clearly revealed, and information is given about which program it will be produced with.

Reusability is an extremely important aspect for a research in this domain. A more descriptive and comprehensive narrative on the reuse of data, both used in the content of the project and that will emerge as a result of the project, could make a very positive contribution to the project.

The metadata and the metadata schema to be used are specified in the data definition, the standardized vocabularies and naming conventions sections are left to be updated later. This is quite normal as it is progressed through the envisioned datasets. It is also stated that the metadata will be harvestable. In this section, more detailed information can be given in the light of the methodological approach followed, as the study progresses.

Since the data does not contain personal or sensitive information, there is no ethical or legal problem. For this reason, a comprehensive explanation is not required. On the other hand carrying out a detailed approach to data security and backup is important for the sustainability of the project. Additional information on this topic would improve the project.

It is also quite normal for this stage that the distribution of tasks among researchers is not specified. As the study progresses, a more specific view may emerge on this subject.

I think that the small changes mentioned above can contribute positively to the project. In addition, a proof reading to be made at the last stage of the documentation may be useful to prevent typos that may have been overlooked. Despite the creation of an ongoing study at the initial stage, a very satisfactory, realistic and applicable data management plan has emerged. As the work progresses and deepens, I believe that the project and the data management plan will evolve into a much richer and more comprehensive point as a result of the updates to be made with the progress that will occur.

References

Venditti G., Brembilla D., Catizone C. (2022). La Chouffe DMP New (Version 0). Zenodo.
<https://doi.org/10.5281/zenodo.6411449>