

Review of: "DMP for the Open Science Project 2023/24 - Team Harkonnen"

Erica Andreose¹

¹ University of Bologna

Potential competing interests: No potential competing interests to declare.

- **Assessment of existing data**

The DMP provides a thorough assessment of existing data sources, including the Open Citation Meta and Crossref datasets, which are utilized to create the OpenCitations Peer Review Index (OCPRI). I would suggest being more specific about the provenance by inserting some links. The description of the dataset specifies that it does not support any software, but I imagine it will support the software the team is going to implement. I would suggest inserting it and linking the repository.

- **Information on new data**

Clear information is provided regarding the new data generated by the software, referred to as the OpenCitations Peer Review Index Software (OCPRIS).

- **Quality assurance of data**

The DMP demonstrates a commitment to data quality assurance through various documented procedures, including the use of tools for automatic checks (on which it would be appropriate to have more specifics in the future) and adherence to format specifications. This ensures the reliability and accuracy of the output dataset. Given the use of metadata and controlled vocabularies, it would be appropriate to report what type of vocabularies were chosen.

- **Backup and security of data**

Data preservation and security are addressed adequately in the DMP, with plans to preserve the dataset through a DOI for long-term access. While specific details on backup mechanisms are not provided, the use of a DOI ensures persistent identification and accessibility, mitigating risks associated with data loss.

- **Expected difficulties in data sharing**

The DMP acknowledges the potential for difficulties in data sharing, particularly regarding the sensitivity of personal data contained within the dataset. However, since the dataset does not contain sensitive information, the impact on data sharing is minimal, and no significant challenges are anticipated in this regard.

- **Copyright/intellectual property rights**

The plan adheres to open science principles by licensing the dataset under Creative Commons Attribution 4.0, facilitating reuse by third parties. Additionally, metadata is provided under the Creative Commons Zero license, ensuring accessibility and interoperability while minimizing restrictions on use.

- **Responsibilities**

Clear responsibilities are outlined for individuals involved in the management of the described output.

- **Preparation of data for sharing and archiving**

The DMP demonstrates a comprehensive approach to preparing data for sharing and archiving, including the provision of metadata and compliance with interoperability best practices (FAIR principles). The Data Management Plan is well done and is a great starting point, being a draft.