

# Review of: "Data Management Plan of Retrieving SSH Journals Citation Information from three datasets (COCI, META and ERIH-PLUS)"

Seyedali Ghasempouri<sup>1</sup>

<sup>1</sup> University of Bologna

**Potential competing interests:** No potential competing interests to declare.

Overall, the data management plan appears comprehensive and well-structured. It covers essential aspects such as data summary, reused data, metadata provisions, data accessibility, interoperability, data reuse, allocation of resources, data security, and ethical aspects. However, there are a few points that need improvement or clarification:

**Clarity and organization:** The data management plan is well-structured and covers essential aspects such as data summary, reused data, metadata provisions, data accessibility, interoperability, data reuse, allocation of resources, data security, and ethical aspects. However, it could benefit from a more explicit outline or a table of contents to help readers navigate the document more easily.

**Metadata and standardization:** The plan would benefit from the use of standardized vocabularies and controlled vocabularies for metadata. Using standard vocabularies enhances the discoverability, interoperability, and reusability of the dataset. It is recommended to consider using a controlled vocabulary and providing mappings to more commonly used ontologies to improve the interoperability of the data.

**Quality assurance:** The plan currently lacks documented procedures for quality assurance of the data. Incorporating such procedures is essential to ensure the reliability and accuracy of the dataset. It is recommended to include a dedicated section that outlines the quality assurance process, including data validation, cleaning, and any other necessary steps to ensure the data's quality.

**Data storage and preservation:** While the plan mentions that data will be stored securely with backup and recovery procedures, it would be helpful to provide more details about the specific storage solution and clarify the duration of the limited storage time. Providing this information will help ensure long-term data preservation and accessibility.

**Potential difficulties in data sharing:** Although the plan states that there are no ethical or legal issues that can impact data sharing, it does not explicitly address potential difficulties in data sharing. It is recommended to include a section discussing any potential barriers to data sharing and strategies for overcoming these challenges.

**Support for data reuse:** The plan could benefit from a more comprehensive description of the support provided for data reuse. While it mentions that workflow and method descriptions will be provided and that data reuse will be supported for up to 5 years, it would be helpful to elaborate on the types of support that will be offered, such as user guides, example



use cases, or a helpdesk for user inquiries.