

Criteria for the Selection of Trustworthy Repositories

TEMPLATE FROM THE SCIENCE EUROPE PRACTICAL GUIDE TO

THE INTERNATIONAL ALIGNMENT OF  
RESEARCH DATA MANAGEMENTA black and white logo

Description automatically generated with low confidence

# Introduction

Providing access to data is one of the pillars of sound, reproducible scientific research. More and more research funding organisations, institutions, and journals require researchers to deposit their research data in repositories. Researchers need to be able to identify trustworthy repositories where they can store their data for sharing. There is currently no generally accepted list of such repositories, whereas general registries of repositories list more than 2,000 of them. However, the maturity and trustworthiness of these repositories is difficult to assess. Many repositories have not yet sought to be certified by an acknowledged certification body. Identifying an appropriate repository can therefore be a challenging task for researchers, their organisations, and funding organisations.

In some disciplines, researchers work with discipline-specific repositories which already have certain policies and standards in place that meet the needs of the specific community. Other repositories serve a more general research public, and their policies and standards are necessarily more generic as well.

Some repositories have been certified as trustworthy repositories by one of several acknowledged certification bodies. In order to facilitate the recognition of trustworthy repositories for researchers, it is strongly recommended that repositories that have not yet been certified seek certification by such a body.

It is always recommended to refer to broadly recognised discipline-specific or certified repositories in the first place. The criteria for the selection of trustworthy repositories presented in this guide should be used in cases where no such repository can be identified.

The list of criteria presented in this guide consists of a number of minimum criteria, organised on four major topics that all trustworthy repositories should fulfil. This list does not prioritise one criterion over another.

# Criteria for the Selection of Trustworthy Repositories

Trustworthy repositories should meet the following minimum criteria:

1. Provision of Persistent and Unique Identifiers (PID)
2. Allow data discovery and identification
3. Enable searching, citing, and retrieval of data
4. Provide support for data versioning
5. Metadata
6. Enable finding of data
7. Enable referencing to related relevant information, such as other data and publications
8. Provide information that is publicly available and maintained, even for non-published, protected, retracted, or deleted data
9. Use metadata standards that are broadly accepted (by the scientific community)
10. Ensure that metadata are machine-retrievable
11. Data Access and Usage Licences
12. Enable access to data under well-specified conditions
13. Ensure data authenticity and integrity
14. Enable retrieval of data
15. Provide information about licensing and permissions (in ideally machine-readable form)
16. Ensure confidentiality and respect rights of data subjects and creators
17. Preservation
18. Ensure persistence of metadata and data
19. Be transparent about mission, scope, preservation policies, and plans (including governance, financial sustainability, retention period, and continuity plan)