



SCIENCE EUROPE

PRACTICAL GUIDE TO
**THE INTERNATIONAL ALIGNMENT OF
RESEARCH DATA MANAGEMENT**

November 2018

'Practical Guide to the International Alignment of Research Data Management':
D/2018/13.324/4

Author: Science Europe

Co-ordination: Science Europe Working Group on Research Data

For further information please contact office@scienceeurope.org

© Copyright Science Europe 2018. This work is licensed under a Creative Commons Attribution 4.0 International Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited, with the exception of logos and any other content marked with a separate copyright notice. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Icons made by monkik and Gregor Cresnar from www.flaticon.com



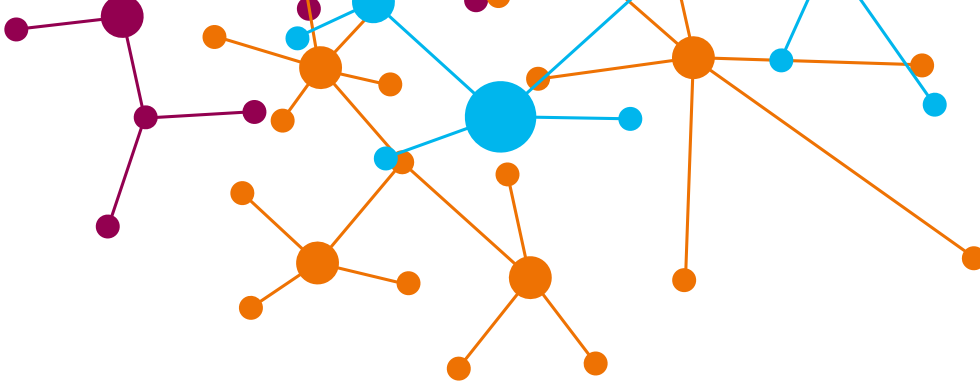


Table of Contents

Foreword by Professor Stan Gielen	2
Introduction	4

CORE REQUIREMENTS FOR DATA MANAGEMENT PLANS

7

CRITERIA FOR THE SELECTION OF TRUSTWORTHY REPOSITORIES

11

GUIDANCE

Translating the Core Requirements into a DMP template
Guiding the Selection of Trustworthy Repositories

15

Notes and References	31
Annex: Compatibility with the FAIR Data Principles	32

Foreword by Professor Stan Gielen

*Member of the Science Europe Governing Board
and President of the Netherlands Organisation
for Scientific Research*



At the European Open Science Cloud (EOSC) Summit in June 2017, I made a commitment to champion the alignment of research data management (RDM) among research funding organisations in Europe. This commitment was the origin of an initiative for that purpose, launched by Science Europe and the Netherlands Organisation for Scientific Research (NWO) in January 2018. The aim of the initiative was to develop a set of core requirements for data management plans (DMPs), as well as a list of criteria for the selection of trustworthy repositories where researchers can store their data for sharing. In light of the development of the EOSC and an increasing tendency towards data sharing, these requirements and criteria should help to harmonise rules on data management throughout Europe. This will aid researchers in complying with RDM requirements even when working with different research funders and research organisations.

Less than a year after its launch, I am pleased to introduce the results of this endeavour. These core requirements for DMPs and criteria for the selection of trustworthy repositories have been developed by experts from Science Europe's Member Organisations, who have sought additional input from external stakeholders to ensure a broad consensus.

Science Europe and NWO will promote these core requirements and criteria in order to ensure they are taken up and supported by as many stakeholders as possible.



I am proud that NWO is among the pioneer organisations within Science Europe that intend to implement these core requirements and criteria in their policies in the course of 2019. With support from other research organisations, both inside and outside of Science Europe, this guide can serve as a reference document for the development or revision of RDM policies throughout Europe and beyond.

November 2018

A handwritten signature in blue ink, appearing to read 'C.C.A.M. Gielen', is positioned above the name.

Professor Dr C.C.A.M. Gielen

Introduction

Research funding organisations, research organisations,¹ and individual researchers have different needs and requirements when it comes to research data management (RDM). The core requirements for data management plans (DMPs) and criteria for the selection of trustworthy repositories presented in this guide provide organisations and communities with a common basis from which they can develop RDM policies. These should be considered as minimum requirements that can be supplemented with more specific ones according to the needs of the community or organisation in question.

Quality-assured research data are key building blocks of the research process. Research data should be permanently, publicly, and freely available for re-use. In recent years, diverse stakeholders from research funders to publishers have endorsed a concise set of principles, known as the FAIR Data Principles,² to enhance the re-use of data. The core requirements for DMPs and criteria for the selection of trustworthy repositories presented in this guide are compliant with the FAIR Data Principles, and even go beyond them on aspects such as storage and backup during the project and long-term preservation.³ Data management based on these core requirements and criteria will therefore support researchers in ensuring that data are FAIR, where appropriate. There may be legitimate reasons (including project-specific or privacy-related ones) for delayed or restricted access, which call for a balanced approach towards openness to research data.

This guide has been developed by experts from Science Europe's Member Organisations. Throughout the development process, they compared existing templates and current practices. Stakeholders from the larger research community were also consulted to take their various needs into account.⁴

This guide was developed in a way that makes it useful for a large number of organisations and researchers. It focuses on content-related questions and does not refer to any procedural aspects⁵ of using DMPs and selecting a repository, as those may differ significantly among organisations.

This guide is divided into three parts:

Core Requirements for Data Management Plans: six aspects that every DMP should cover, with detailed guiding questions.

Criteria for the Selection of Trustworthy Repositories: four topics detailing criteria that every trusted repository should meet.

Guidance: more detailed information and examples to support the implementation of the requirements and criteria into an organisation's policies.

HOW TO USE THIS GUIDE

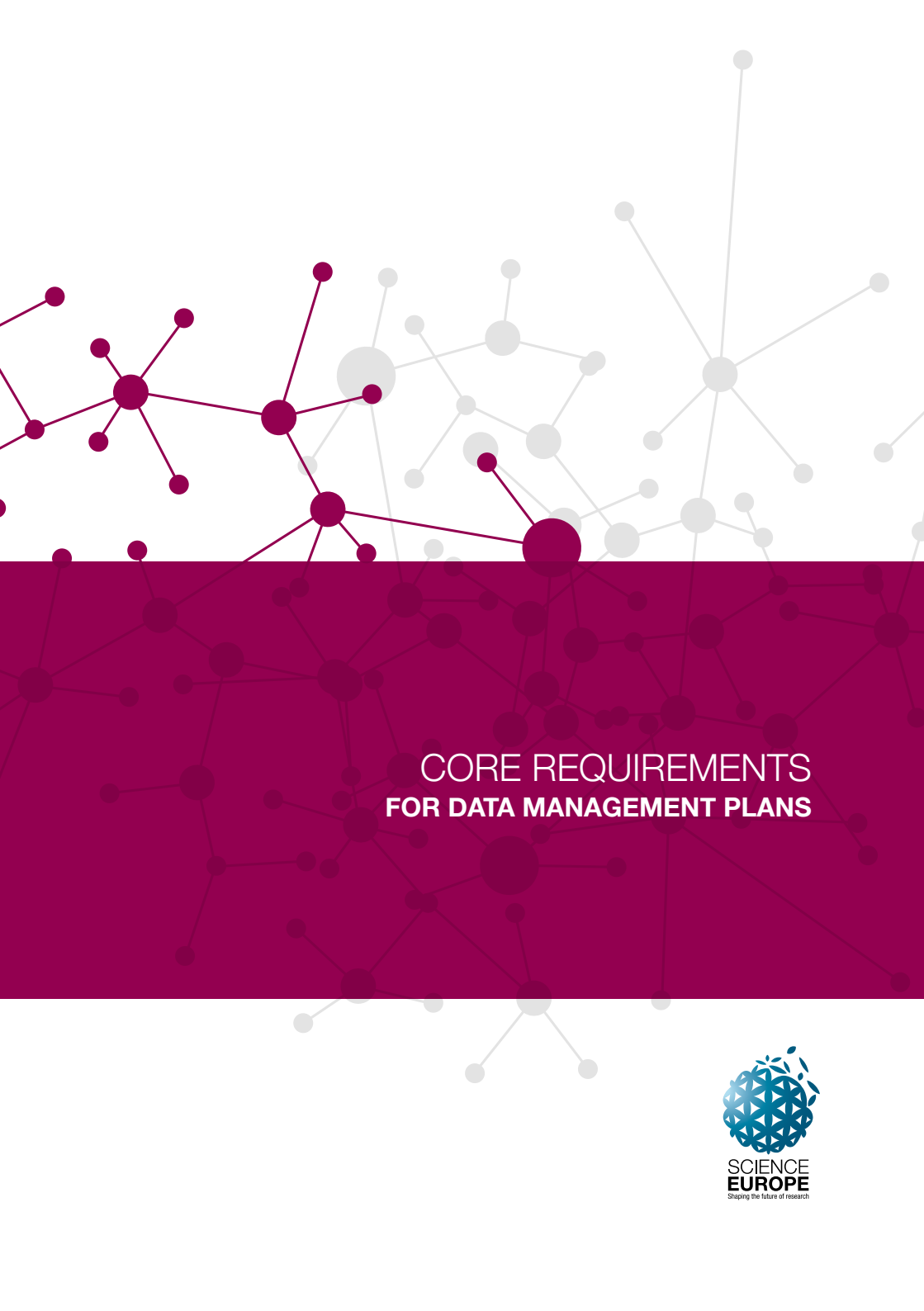
Research funding organisations, research organisations, and research communities are encouraged to use the **core requirements for data management plans** as a basis to set up their own DMP templates. At a later stage, the core requirements can also serve as reference document for the evaluation of DMPs.

Research organisations and individual researchers can refer to this guide for support when drafting their DMPs.

For other actors in the research sector, this guide serves as a reference document on how a DMP should be structured and used.

The **criteria for the selection of trustworthy repositories** will help research funding organisations, research organisations, and individual researchers to identify repositories for storing and sharing data.

The core requirements for DMPs and the criteria for the selection of trustworthy repositories can be seen as stand-alone documents and used independently. It is however recommended to take both into consideration when developing or amending institutional or discipline-specific policies in order to reach the best possible alignment among institutions.



CORE REQUIREMENTS FOR DATA MANAGEMENT PLANS

Introduction to the Core Requirements for Data Management Plans

Research funding organisations and research organisations increasingly require researchers to develop data management plans. These plans support the researcher in considering all relevant aspects of data management from the very beginning of a research project. A DMP should stimulate researchers to think about optimal handling, organising, documenting, and storing of their data.

Currently, there is a lot of variation in research data management policies. Many research funding organisations, research organisations, and research communities have developed their own rules and templates. This can be confusing for researchers and is especially problematic as many researchers acquire their funding from different sources; they are increasingly confronted with different grant requirements and institutional policies. There is an urgent need to align data management policies in order to provide more clarity for researchers. DMPs should not be a bureaucratic burden for them, but a useful means of support when planning and conducting a research project.

The following list presents six topics that should be covered in DMPs, each specified with several guiding questions. These topics and questions for setting up a DMP form the core requirements that every research funding organisation should request in order for the researcher to develop a useful DMP. The order of the core requirements can be changed according to specific needs and organisational focal points. However, all six core requirements need to be addressed in a DMP.



An example template providing guidance on which aspects to further consider in a DMP can be found on Page 17 of this guide.



CORE REQUIREMENTS FOR DATA MANAGEMENT PLANS

When developing solid data management plans, researchers are required to deal with the following topics and answer the following questions:

- **1. Data description and collection or re-use of existing data**
 - a. How will new data be collected or produced and/or how will existing data be re-used?
 - b. What data (for example the kinds, formats, and volumes) will be collected or produced?
-

- **2. Documentation and data quality**
 - a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?
 - b. What data quality control measures will be used?
-

- **3. Storage and backup during the research process**
 - a. How will data and metadata be stored and backed up during the research process?
 - b. How will data security and protection of sensitive data be taken care of during the research?
-

- **4. Legal and ethical requirements, codes of conduct**
 - a. If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?
 - b. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?
 - c. How will possible ethical issues be taken into account, and codes of conduct followed?
-





5. Data sharing and long-term preservation

- a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?
 - b. How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?
 - c. What methods or software tools will be needed to access and use the data?
 - d. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?
-



6. Data management responsibilities and resources

- a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?
 - b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?
-

**CRITERIA FOR
THE SELECTION OF
TRUSTWORTHY
REPOSITORIES**

Introduction to the Criteria for the Selection of Trustworthy Repositories

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 are 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.



More detailed explanations on the criteria for the selection of trustworthy repositories can be found on Page 26 of this guide.

CRITERIA FOR THE SELECTION OF TRUSTWORTHY REPOSITORIES



Trustworthy repositories should meet the following minimum criteria:

○ **1. Provision of Persistent and Unique Identifiers (PIDs)**

- a. Allow data discovery and identification
- b. Enable searching, citing, and retrieval of data
- c. Provide support for data versioning

○ **2. Metadata**

- a. Enable finding of data
- b. Enable referencing to related relevant information, such as other data and publications
- c. Provide information that is publicly available and maintained, even for non-published, protected, retracted, or deleted data
- d. Use metadata standards that are broadly accepted (by the scientific community)
- e. Ensure that metadata are machine-retrievable

○ **3. Data access and usage licences**

- a. Enable access to data under well-specified conditions
 - b. Ensure data authenticity and integrity
 - c. Enable retrieval of data
 - d. Provide information about licensing and permissions (in ideally machine-readable form)
 - e. Ensure confidentiality and respect rights of data subjects and creators
-





4. Preservation

- a. Ensure persistence of metadata and data
 - b. Be transparent about mission, scope, preservation policies, and plans (including governance, financial sustainability, retention period, and continuity plan)
-



GUIDANCE

Translating the Core Requirements into a DMP template

The following example of a data management plan template is based on the core requirements for DMPs.⁶ These core requirements should be considered as a minimum standard, leaving the flexibility to formulate additional guidelines according to the needs of specific domains or to national or local legislation.

The template presented below refers to the 15 questions covering six core requirements for good data management. Additional guidance and explanations are provided to help researchers fill out such a template and to assure that all relevant aspects of research data management are covered. The below table is an example of how the core requirements can be transformed into a DMP template. It will be up to the individual organisations and disciplines to develop templates that fit their needs.

GENERAL INFORMATION

Administrative information

- Provide information such as name of applicant, project number, funding programme, version of DMP.

1 DATA DESCRIPTION AND COLLECTION OR RE-USE OF EXISTING DATA

1 a

How will new data be collected or produced and/or how will existing data be re-used?

- Explain which methodologies or software will be used if new data are collected or produced.
- State any constraints on re-use of existing data if there are any.
- Explain how data provenance will be documented.
- Briefly state the reasons if the re-use of any existing data sources has been considered but discarded.

1 b

What data (for example the kind, formats, and volumes), will be collected or produced?

- Give details on the kind of data: for example numeric (databases, spreadsheets), textual (documents), image, audio, video, and/or mixed media.
- Give details on the data format: the way in which the data is encoded for storage, often reflected by the filename extension (for example pdf, xls, doc, txt, or rdf).
- Justify the use of certain formats. For example, decisions may be based on staff expertise within the host organisation, a preference for open formats, standards accepted by data repositories, widespread usage within the research community, or on the software or equipment that will be used.
- Give preference to open and standard formats as they facilitate sharing and long-term re-use of data (several repositories provide lists of such 'preferred formats').
- Give details on the volumes (they can be expressed in storage space required (bytes), and/or in numbers of objects, files, rows, and columns).

2 DOCUMENTATION AND DATA QUALITY

2a

What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data?

- Indicate which metadata will be provided to help others identify and discover the data.
- Indicate which metadata standards (for example DDI, TEI, EML, MARC, CMDI) will be used.
- Use community metadata standards where these are in place.
- Indicate how the data will be organised during the project, mentioning for example conventions, version control, and folder structures. Consistent, well-ordered research data will be easier to find, understand, and re-use.
- Consider what other documentation is needed to enable re-use. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, and so on.
- Consider how this information will be captured and where it will be recorded for example in a database with links to each item, a 'readme' text file, file headers, code books, or lab notebooks.

2b

What data quality control measures will be used?

- Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeated samples or measurements, standardised data capture, data entry validation, peer review of data, or representation with controlled vocabularies.

3 STORAGE AND BACKUP DURING THE RESEARCH PROCESS

3a

How will data and metadata be stored and backed up during the research?

- Describe where the data will be stored and backed up during research activities and how often the backup will be performed. It is recommended to store data in least at two separate locations.
- Give preference to the use of robust, managed storage with automatic backup, such as provided by IT support services of the home institution. Storing data on laptops, stand-alone hard drives, or external storage devices such as USB sticks is not recommended.

3b

How will data security and protection of sensitive data be taken care of during the research?

- Explain how the data will be recovered in the event of an incident.
- Explain who will have access to the data during the research and how access to data is controlled, especially in collaborative partnerships.
- Consider data protection, particularly if your data is sensitive for example containing personal data, politically sensitive information, or trade secrets. Describe the main risks and how these will be managed.
- Explain which institutional data protection policies are in place.

4 LEGAL AND ETHICAL REQUIREMENTS, CODES OF CONDUCT

4a

If personal data are processed, how will compliance with legislation on personal data and on security be ensured?

- Ensure that when dealing with personal data data protection laws (for example GDPR) are complied with:
 - › Gain informed consent for preservation and/or sharing of personal data.
 - › Consider anonymisation of personal data for preservation and/or sharing (truly anonymous data are no longer considered personal data).
 - › Consider pseudonymisation of personal data (the main difference with anonymisation is that pseudonymisation is reversible).
 - › Consider encryption which is seen as a special case of pseudonymisation (the encryption key must be stored separately from the data, for instance by a trusted third party).
 - › Explain whether there is a managed access procedure in place for authorised users of personal data.

4b

How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?

- Explain who will be the owner of the data, meaning who will have the rights to control access:
 - › Explain what access conditions will apply to the data? Will the data be openly accessible, or will there be access restrictions? In the latter case, which? Consider the use of data access and re-use licenses.
 - › Make sure to cover these matters of rights to control access to data for multi-partner projects and multiple data owners, in the consortium agreement.
- Indicate whether intellectual property rights (for example Database Directive, *sui generis* rights) are affected. If so, explain which and how will they be dealt with.
- Indicate whether there are any restrictions on the re-use of third-party data.

4c

What ethical issues and codes of conduct are there, and how will they be taken into account?

- Consider whether ethical issues can affect how data are stored and transferred, who can see or use them, and how long they are kept. Demonstrate awareness of these aspects and respective planning.
- Follow the national and international codes of conducts and institutional ethical guidelines, and check if ethical review (for example by an ethics committee) is required for data collection in the research project.

5 DATA SHARING AND LONG-TERM PRESERVATION

5a

How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

- Explain how the data will be discoverable and shared (for example by deposit in a trustworthy data repository, indexed in a catalogue, use of a secure data service, direct handling of data requests, or use of another mechanism).
- Outline the plan for data preservation and give information on how long the data will be retained.
- Explain when the data will be made available. Indicate the expected timely release. Explain whether exclusive use of the data will be claimed and if so, why and for how long. Indicate whether data sharing will be postponed or restricted for example to publish, protect intellectual property, or seek patents.
- Indicate who will be able to use the data. If it is necessary to restrict access to certain communities or to apply a data sharing agreement, explain how and why. Explain what action will be taken to overcome or to minimise restrictions.

5b

How will data for preservation be selected, and where data will be preserved long-term (for example a data repository or archive)?

- Indicate what data must be retained or destroyed for contractual, legal, or regulatory purposes.
- Indicate how it will be decided what data to keep. Describe the data to be preserved long-term.
- Explain the foreseeable research uses (and/or users) for the data.
- Indicate where the data will be deposited. If no established repository is proposed, demonstrate in the data management plan that the data can be curated effectively beyond the lifetime of the grant. It is recommended to demonstrate that the repositories policies and procedures (including any metadata standards, and costs involved) have been checked.

5c

What methods or software tools are needed to access and use data?

- Indicate whether potential users need specific tools to access and (re-)use the data. Consider the sustainability of software needed for accessing the data.
- Indicate whether data will be shared via a repository, requests handled directly, or whether another mechanism will be used?

5d

How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

- Explain how the data might be re-used in other contexts. Persistent identifiers should be applied so that data can be reliably and efficiently located and referred to. Persistent identifiers also help to track citations and re-use.
- Indicate whether a persistent identifier for the data will be pursued. Typically, a trustworthy, long-term repository will provide a persistent identifier.

6 DATA MANAGEMENT RESPONSIBILITIES AND RESOURCES

6 a

Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?

- Outline the roles and responsibilities for data management/stewardship activities for example data capture, metadata production, data quality, storage and backup, data archiving, and data sharing. Name responsible individual(s) where possible.
- For collaborative projects, explain the co-ordination of data management responsibilities across partners.
- Indicate who is responsible for implementing the DMP, and for ensuring it is reviewed and, if necessary, revised.
- Consider regular updates of the DMP.

6 b

What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

- Explain how the necessary resources (for example time) to prepare the data for sharing/preservation (data curation) have been costed in. Carefully consider and justify any resources needed to deliver the data. These may include storage costs, hardware, staff time, costs of preparing data for deposit, and repository charges.
- Indicate whether additional resources will be needed to prepare data for deposit or to meet any charges from data repositories. If yes, explain how much is needed and how such costs will be covered.

Guiding the Selection of Trustworthy Repositories

The following table provides guidance for the selection of trustworthy repositories by criteria structured according to four main topics.

1 PROVISION OF PERSISTENT AND UNIQUE IDENTIFIERS (PIDS)

A trustworthy repository should:

1a Allow data discovery and identification

- ensure that PIDs are included in the corresponding metadata.

1b Enable searching, citing, and retrieval of data

- consistently assign PIDs (for example a DOI,⁷ URN,⁸ ARK⁹) to the data it holds, allowing the corresponding data and metadata to be found, referred to, and retrieved, even if the location where the data are stored changes.

1c Provide support for data versioning

- ensure that the version of the data stored in the repository is clearly specified and documented via a permanent audit trail in order for the provenance to be traced.

Note: Not all repositories use an accepted and universal PID system such as the ones mentioned above. Instead, they use a local identifier or administrative number that the repository itself maintains. This increases the risk that the data cannot be found anymore if they are moved to another location, or if the repository ceases to exist, reorganises, or changes its governance.

2 METADATA

The data should be accurately described with rich metadata. The metadata should document how the data were generated, under what license and how they can be re-used, and provide the context for proper interpretation by other researchers.

A trustworthy repository should:

2a Enable finding data

- ensure interoperability and re-use of data by others by providing the data and metadata in an accessible language, based on a well-established formalism. Data and metadata should be described using standard vocabularies and formats allowing computer systems to search for them, combine them in an automatic way, and distinguish the metadata from the research data file(s).

2b Enable referencing to related relevant information

- ensure that in the metadata information it is possible to declare links to other relevant or associated information by providing the PID and a description of the scientific relation. One particular kind of information is details on the associated researcher, for which permanent research IDs exist (such as ORCID, ¹⁰ ISNI, ¹¹ or DAI ¹²).

2c Provide information that is publicly available and maintained, even for non-published, protected, retracted, or deleted data

- ensure that metadata are archived for the long term and that metadata always remain retrievable, even if the corresponding research data are not or no longer available (for example due to privacy restriction, legal obligations, or other protective measures).
- ensure that retracted data due to poor research practices or misconduct are still findable through the metadata and preserved in order to allow examination of the research record.

2d Use metadata standards that are broadly accepted (by the scientific community)

- ensure that the metadata maintained by the repository are machine-retrievable and use standards that are broadly accepted (by the scientific community).

- ensure that community standards or best practices for data handling are followed if these exist. Note that repositories that are specialised in a particular research field may have community standards regarding the data and metadata that are uploaded.

2e Ensure that metadata are machine-retrievable

- encourage that the information contained in the metadata are structured in a way that allows machines to retrieve it, for example by providing a form with specific fields to be completed.

3 DATA ACCESS & USAGE LICENSES

A trustworthy repository should:

3a Enable access to data under well-specified conditions

- be clear about the terms under which the data can be re-used. Such (license) information is usually included in the metadata.

3b Ensure data authenticity and integrity

- ensure that the metadata contain detailed information about the provenance of data, including how they were generated, how they were processed, in which context they may be re-used, and how reliable they are.

3c Enable retrieval of data

- allow retrieval of data or at least metadata using an open and standardised protocol (not a proprietary communication protocol).

3d Provide information about licensing and permissions (in ideally machine-readable form)

- allow license information to be referred to in a structured way, so that the conditions of use are clear, preferably to humans as well as to machines. Where possible, common or broadly accepted licensing systems should be used (such as Creative Commons) which can be referred to by URL.

3e Ensure confidentiality and rights of data subjects and creators

- provide a way for authentication and authorisation of human and machine-users, allowing to set user (or group) specific access rights to account for data with confidentiality issues and other restrictions.

4 PRESERVATION

A trustworthy repository should:

4a Ensure persistence of metadata and data

- ensure the preservation and continued availability and access to the data and metadata entrusted to it by its users.

4b Be transparent about mission, scope, preservation policies, and plans (including governance, financial sustainability, retention period, and continuity plan)

- manage the preservation of data and metadata in a documented way. In particular, it should have a preservation policy that details the mission and scope of the repository, governance aspects, financial sustainability, outsource partners, and retention periods (the timeframe of preservation).
 - have a publicly available contingency plan and ensure preservation of data and metadata beyond the lifetime of the repository (for example by allowing easy extraction and transfer of data and metadata to another repository).
-

Notes & References

- 1 The term research organisations refers to research performing organisations, universities, and research institutes.
- 2 Making data Findable, Accessible, Interoperable, and Re-usable. Please see the annex for further information or visit <https://www.force11.org/group/fairgroup/fairprinciples>
- 3 For further information on how the FAIR Principles are translated into the core requirements and criteria, please see the Annex.
- 4 The concept was presented at open event on 30 January 2018, and two consultation rounds were organised in April 2018 and August/September 2018: <https://scieur.org/rdm-initiative>
- 5 For procedural elements of implementing DMPs: RDA DMP Common Standards WG <https://www.rd-alliance.org/groups/dmp-common-standards-wg>
- 6 The core requirements for data management plans were developed as part of the initiative for the voluntary international alignment of research data management requirements, led by Science Europe and the Netherlands Organisation for Scientific Research (NWO). Detailed information about the initiative is available at <http://scieur.org/rdm-initiative>
- 7 Digital Object Identifier
- 8 Uniform Resource Name
- 9 Archival Resource Key
- 10 Open Researcher and Contributor ID
- 11 International Standard Name Identifier
- 12 Digital Author Identifier

Annex:

Compatibility with the FAIR Data Principles

THE FAIR DATA PRINCIPLES				Core Requirements for DMPs (CR)	Criteria for the Selection of Trustworthy Repositories
To be Findable					
F1	(meta)data are assigned a globally unique and eternally persistent identifier			CR 5d	Criterion 1
F2	data are described with rich metadata			CR 2a	Criterion 2
F3	metadata clearly and explicitly include the identifier of the data they describe			CR 5d	Criterion 1, Criterion 2
F4	(meta)data are registered or indexed in a searchable resource				Criterion 2
To be Accessible					
A1	(meta)data are retrievable by their identifier using a standardised communications protocol			CR 5c	Criterion 1, Criterion 2
A1.1	the protocol is open, free, and universally implementable			CR 5c	Criterion 2
A1.2	the protocol allows for an authentication and authorisation procedure, where necessary			CR 4b, CR 5a, CR 5c	Criterion 3
A2	metadata are accessible, even when the data are no longer available			CR 4c, CR 5a, CR 5d	Criterion 2c

THE FAIR DATA PRINCIPLES

Core Requirements
for DMPs (CR)

Criteria for the
Selection of
Trustworthy
Repositories

To be Interoperable			
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation	CR 1b, CR 2a	Criterion 2d
I2	(meta)data use vocabularies that follow FAIR principles	CR 2a, CR 2b	Criterion 2
I3	(meta)data include qualified references to other (meta)data	CR 2a, CR 5a, CR 5c	Criterion 2b
To be Re-usable			
R1	meta(data) are richly described with a plurality of accurate and relevant attributes	CR 2a, CR 2b	Criterion 2
R1.1	(meta)data are released with a clear and accessible data usage license	CR 4b, CR 5a	Criterion 3d
R1.2	(meta)data are associated with detailed provenance	CR 1a, CR 1b, CR 2b	Criterion 1c, Criterion 2, Criterion 3b, Criterion 4a
R1.3	(meta)data meet domain-relevant community standards	CR 1b, CR 2a	Criterion 2d

Notes

[illegible]

[illegible]

[illegible]

Science Europe

Rue de la Science 14
1040 Brussels
Belgium

Tel +32 (0)2 226 03 00
Fax +32 (0)2 226 03 01
office@scienceeurope.org
www.scienceeurope.org

