

# Research Data Management paragraph - Horizon Europe

Radboud University Template

v0.3 – September 2022

## **Need help?**

For support in drafting your data management paragraph or feedback, please get in touch with RDM Support (<https://www.ru.nl/rdm/support-training/>) or the data steward of your institute (<https://www.ru.nl/rdm/vm/policy-documents/>).

In this document we provide you with guidelines and examples on what to mention in the Horizon Europe Data Management Paragraph.

The following sections (indicated with the yellow headings) should be part of the Horizon Europe application form:

## **1. Types of data/research outputs**

Try to describe as detailed as possible what types of data you will produce or (re-)use in your research. This could include:

- Whether you create/collect your own data or that you re-use data from other sources
- The methods used, for example, questionnaires, interviews, literary review, etc.
- The data formats that result from these methods. For example, text files (.txt, .pdf), audio files, video files, spreadsheets and tables (.dat, .xlsx, .csv), etc.
- That you are aware of it when you collect personal data. For example, names, addresses, IP-addresses
- That you are aware if you collect special categories of personal data. For example, data concerning religion, health, sexual orientation, etc.
- For more information on personal data specific to Europe Horizon grants see section 4 on personal data of this file: [https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/how-to-complete-your-ethics-self-assessment\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/how-to-complete-your-ethics-self-assessment_en.pdf)

## **2. FAIR principles**

*Corresponds to: Findability of data/research outputs, Accessibility of data/research outputs, Interoperability of data/research outputs, and Reusability of data/research outputs*

Here we shall give you a description of all items regarding the FAIR principles and we have provided you with an example text at the end of the section.

In this section you should mention in which (trusted) repository you will store your data after your research and how this data complies to the FAIR principles. These principles most clearly apply to your dataset when you publish your data open access (i.e., open to the public), however there may be good reasons not to make all your data available for reuse, or at least not right away. Perhaps some of your data will not be of interest to others, or perhaps you do not have the ownership. It

might also be the case that your data is privacy- or competition-sensitive. In most research projects data can be made available for reuse, but if this is not the case, please explain clearly why. Also, this does not mean that the FAIR principles do not apply to these types of datasets. For example, if you use a repository with restricted access, the dataset can still receive a persistent identifier and the metadata can be visible and indexed by search engines.

So, whether you publish open access or not, the following points regarding the FAIR principles can be mentioned in your data management paragraph:

- **Findability:** How can others easily find your dataset?
  - Repository: Start out by mentioning the name of the repository you are planning to use for long-term storage.
  - Persistent identifier: Does the repository that you have chosen assign a persistent identifier (e.g., a DOI) to your dataset? That way your dataset is uniquely and persistently identifiable. If so, you could specify the type of identifier.
  - Metadata: Metadata describe the data, for example by keywords, a summary, authors, codebooks etc. This can be done by the repository, but also by you. In case of the former: often, repositories with a certificate adhere to specific metadata standards. You can mention these standards. In case of the latter: you can add documentation on the content, context and structure of the dataset. Mention what you are planning to add.
  - Indexed in a searchable resource: Check if the repository of your choice is indexed by regular search engines, such as Google Scholar. Repositories with a certificate (such as the CoreTrustSeal) are often automatically indexed. An overview of data repositories can be found [here](#).
  
- **Accessibility:** How have you regulated who can access the dataset?
  - Intellectual Property Right (IPR) considerations: In general, when you are employed at the Radboud University, the Radboud University is the legal owner of the data. You can mention this. However, in some circumstances other arrangements have been made regarding the property rights. For example, some consortia have documents stating who owns the data. In that case, mention the arrangements made and who has the intellectual property rights. If these are not clear yet, mention that arrangements shall be made with all parties involved.
  - Limitations on and protocols for the use of data are made explicit. Data should be retrievable by authorized persons with a computer and an internet connection through a well-defined protocol. Note that accessible data does not automatically imply open or free access. Data published with restricted access can also be FAIR. Although at Radboud University open access is stimulated, there can be ethical or legal reasons for not making (parts of your) data open access This means that if you have good reasons to publish your data under restricted access or not at all, mention those reasons and the protocols that are in place.
  - Related to the previous point. You can mention any data use agreements or licences under which you will share your data.
  - Timeline: When you decide to share the data, mention when they are accessible. For example, at the end of your project or when a corresponding article is published.
  
- **Interoperability:** Is it possible for people and computers to interpret the data and combine it with other datasets?

- Include proper documentation (see also the third bullet under findability).
  - If existing in your discipline, make use of standard vocabularies, ontologies and thesauri in your (meta)data, or provide mapping of your data to these vocabularies, ontologies and thesauri.
  - Mention that you will try to use interoperable file formats where possible. In general, most preferred interoperable file formats are those that are widely used and free to access. For example, .odt (Open Document Text) or .pdf has a preference over .docx files, seeing that Microsoft Word is no free software and possibly not available to everybody (now or in the future).
- **Reproducibility:** How will you make your data ready for (re-)use by others?
- You can mention any data use agreements or licences under which you will share your data. For example, the CC BY 4.0 licence is a commonly used license. Some repositories have standard licences that will accompany your dataset.
  - Software: When you work with specific software, for example for your analyses, explain how and where the software is available and if it is not commonly available, how you will deal with that. If you don't use special software, mention that all the data can be opened with generally available software tools.

**Example text:**

The following example text is based on the above-mentioned bullet points. Please keep in mind that the example is very specific and tailored to a mock dataset and the Radboud Data Repository. Every project and every repository need tailored explanations. Read the example carefully and adjust where necessary. **Be aware: just copy-pasting this example would likely not be enough!** For example, when you are not able to share your data open access, remove those sentences and explain why your data is not suited for open access data sharing.

*“It is the policy of Radboud University in general and our research institute [name research institute] in particular to comply with the FAIR principles and share with the scientific community any data obtained in research projects, as long as ethical and legal regulations permit it. Where possible, data will be archived via the Radboud Data Repository. Via these archiving facilities, data will be (1) Findable by indexing data by search engines on the internet, including rich metadata according to the Dublin Core and DataCite schemas, and receiving a persistent identifier (DOI), (2) Accessible by using an open internet protocol, including clear authorisation procedures, and, where possible, the data will be shared when related articles are published under an open access license, (3) Interoperable by using standards for metadata (Dublin Core/Datacite), by adding documentation (codebook and readme), using preferred formats, and using the standard vocabulary BIDS, and (4) Reusable by including rich metadata, making sure that all data can be opened and used by generally available software (analysis) tools, by adding documentation with instructions for reuse, and by publishing it under an open access license.”*

### **3. Curation and storage/preservation costs**

- Mention the persons involved in the data management of your data. This includes you, but perhaps also supervisors, colleagues, or consortium members.

- Quality assurance: Mention all persons that are involved in helping you with data management, such as checking the quality of your dataset. For example, data stewards, the RDM support team, or anyone who does curation on your dataset.
- Costs: Standard costs for data management (e.g., RU network drives, "werkgroepfolders" and the Radboud Data Repository) are covered by the RU. Mention these standard costs and add any possible extra costs related to data management. For example: *“Standard data management costs (such as storing data on the RU network drives and archiving data for the long term) are covered by Radboud University. I do not foresee any additional costs.”* Or, when working with consortia: *“Standard data management costs (such as storing and archiving data) are covered by the consortium [or replace by name of consortium partner who pays]. Additional costs, namely [specify] are covered by [specify].”*