

RDMO Questionnaire (generic RDMO catalog)

General

Topic

What is the main research question of the project?

Please give some keywords describing the research question.

Research field

To which discipline / disciplines is the project assigned?

Project schedule

When does the project start?

When does the project end?

If applicable: give reasons for deviations in the project start date or project end date.

Project coordination

Which persons or institutions are responsible for the project coordination?

Project partners

Who is/are your project partner(s)?

Does your institution have rules or guidelines for the handling of research data collected in the project? If so, briefly outline them and refer to further information if necessary. Please also indicate how binding they are.

Who is the data management contact for this project partner?

Funding

Who is funding the project?

Under which funding line and/or programme is the project funded?

Does the research funding organisation have any specifications or guidelines regarding the handling of research data collected in the project? If so, please briefly outline them and refer to further information if necessary. Please also indicate the extent to which they are binding.

Other requirements

Are there other data management requirements (e.g. from the scientific community) that need to be considered?

What data management requirements are these?

Content classification

Data sets

Please give some keywords describing the data set.

What kind of data set is it?

Data origin

Is the data set generated or reused?

If reused, who created the data set?

If reused, at what address, PID or URL is the data set available?

Reuse

Who might be interested in this data set (for reuse)? What are the implications of the potential for reuse for the provision of the data?

By whom and in what context is this data set reused?

Reproducibility

Is the data set reproducible, i.e. could it be re-created or recovered if it were lost?

Technical classification

Data collection

When does the collection or generation of data begin?

When does the collection or generation of data end?

When does the data cleansing/preparation start?

When does the data cleansing/preparation end?

When does the data analysis start?

When does the data analysis end?

If applicable, give reasons for deviations between planned and actual dates.

Data size

What is the size of the data set?

How much data are produced per year?

If applicable, give reasons for deviations between expected and actual values.

Formats

What file formats are used?

Tools

Which tools, software, technologies or processes are used to generate or collect the data?

What software, processes or technologies are required to use the data?

Is documentation about relevant software needed to use the data?

Versioning

Are different versions of the data set created?

What versioning strategy is used for this data set?

What technology or tool is used for versioning?

Name all versions created and indicate the differences to the previous version in each case.

Data usage

Usage scenarios

What will this data set be used for/how will it be used during the project?

How often will this data set be used?

To what extent will infrastructure resources be required (e.g. CPU hours, bandwidth, storage space... etc.).

Are there actual or potential usage scenarios that could benefit from support by a data management or IT expert, or that even require such support?

Data organisation

Where is the data set stored during the project?

Under which URL can the data set be accessed during the project?

Are there internal project guidelines for consistent data organisation? If so, which and where are they documented?

Is there an internal project guideline for naming the data? If so, please briefly outline the naming conventions and, if necessary, link to the documentation.

Data storage and security

Who can access the data set?

How and how often will the data be backed up?

Who is responsible for the backups?

What measures or provisions are in place to ensure data security (e.g. protection against unauthorised access, data recovery, transmission of sensitive data)?

Interoperability

Is this data set interoperable, i.e. can it be shared and reused between researchers, institutions, organisations, countries, etc.?

Data sharing and reuse

Will this data set be published or shared?

If so, where are the data published?

If not, please give reasons for this and distinguish between legal and/or contractual reasons and voluntary restrictions.

If so, under which terms of use or license is the data set published or shared?

If there are any restrictions on the reuse of this data set, please explain why.

When will the data be published (if they are)?

Enter the persistent identifiers of your data set publications.

Collaborative work

Will the data be collaboratively used?

Which platform / tools is / are used for collaboratively working on data and publications?

How is the collaborative work on the same files organised?

Quality assurance

What quality assurance measures are in place for this data set?

Data integration

Is integration between reused and newly created data ensured? If so, how?

Costs

What data management staff costs will be incurred as part of the collection, creation or acquisition of data in the project?

What material costs are incurred for data management as part of the collection, creation or acquisition of data in the project?

What personnel costs are incurred for data management in connection with the use of the data in the project?

What material costs for data management arise in connection with the use of the data in the project?

What personnel costs are incurred in connection with data storage and data security measures during the project?

What material costs are incurred in connection with the storage of data records during the project?

If applicable, give reasons for deviations between estimated and actual values.

Metadata and referencing

Metadata

What information is necessary for others to understand and reuse the data (i.e., to understand its collection or creation, analysis, and research results based on it)?

What standards, ontologies, classifications, etc. are used to describe the data and contextual information?

If it is unavoidable to use project-specific or rare ontologies, metadata schemas or vocabularies, are mappings to common ontologies etc. created?

What metadata is collected automatically?

Which metadata are collected semi-automatically?

Which metadata are collected manually?

Are metadata and contextual information checked for correctness and completeness?

Who is responsible for documenting and checking the metadata and contextual information for accuracy and completeness?

Metadata costs

What personnel costs are incurred in connection with the creation of metadata and contextual information in the project?

What material costs or non-personnel costs are incurred in connection with the creation of metadata and contextual information in the project?

If applicable, give reasons for deviations between estimated and actual values.

Structure, granularity, and referencing

Please briefly explain the structure and the relationship between your data:

Persistent Identifiers (PIDs)

Are persistent identifiers (PIDs) used for this data set?

What system of persistent identifiers is used?

Which sub-units / partial data sets should be given their own identifiers? Which of these identifiers are permanent and citable?

Who is responsible for the maintenance of the PIDs and the object maintenance (i.e. the long-term archiving of the object and thus for informing the PID service of an object move and the new address)?

PID costs

What are the personnel costs associated with using persistent identifiers in the project?

What non-personnel or material costs are incurred in connection with persistent identifiers in the project?

If applicable, give reasons for deviations between estimated and actual values.

Legal and ethics

General legal issues

Does the legal situation of different countries have to be considered?

Personal data

Does this data set contain personal data?

Data protection

In addition to the European GDPR, what other data protection laws are applicable to the project?

Sensitive data

Does the data set contain special categories of personal data?

Will the data be anonymised or pseudonymised?

To what extent is "informed consent" obtained from the persons concerned?

If no "informed consent" is obtained, please give reasons for this.

Where and how are "informed consent" forms kept?

By when will the (non-anonymised or non-pseudonymised) original data be safely deleted/destroyed at the latest?

Other sensitive data

Does this data set contain non-personal sensitive data?

If so, what non-personal sensitive data are involved?

Sensitive data costs

What personnel costs are incurred for the anonymisation of sensitive data in the project?

What are the non-personnel costs associated with anonymising sensitive data in the project?

What personnel costs are incurred in connection with other (non-technical) security measures for sensitive data in the project?

What are the non-personnel costs of other (non-technical) security measures for sensitive data in the project?

If applicable, give reasons for deviations between estimated and actual values.

Official approval

Has the research project been reviewed by an ethics committee?

Is a legal approval / permit needed for the research?

If so, which permit?

If so, who is the issuing agency?

Is a data access committee required to decide whether or not to grant access to data published by the project in response to access requests?

Intellectual property rights I

Is data used and/or created that is protected by copyright or related rights?

Intellectual property rights II

Do copyrights exist or arise in this data set?

Do other intellectual property rights exist or arise for this data set?

Has the rights owner been researched?

Intellectual property rights costs

What personnel expenses are incurred in connection with copyrights or related rights in the project?

What material costs are incurred in connection with copyright and related rights in the project?

If applicable, give reasons for deviations between estimated and actual values.

Storage and long-term preservation

Selection

What criteria/rules are used to select data for archiving (after the end of the project)?

By whom is the selection made?

Long-term preservation

Does this data set have to be preserved for the long-term?

What are the reasons for this data set to be preserved for the long-term?

How long must the data be stored?

How long should the data be (re)usable after the end of the project?

Where will the data (including metadata, documentation and any relevant code or software) be stored or archived after the project is completed?

Is the repository or data centre chosen certified (e.g. Data Seal of Approval, nestor Seal or ISO 16363)?

Have appropriate archiving solutions already been discussed with the repository or data centre?

Shall the data only be made accessible after an embargo period has expired?

How is the identity of persons accessing the data verified?

When will the data be archived?

Who is the long-term contact person who can make decisions about the data? Please enter their name.

Please enter the email address of the permanent contact named above.

Long-term preservation costs

What personnel costs are incurred in connection with long-term archiving for this project?

How will the data management costs of the project be covered?

What material costs are incurred in connection with long-term archiving for this project?

If applicable, give reasons for deviations between estimated and actual values.