ZonMw’s key items – background information

Requirement

ZonMw requires researchers who receive a grant from ZonMw to make a data management plan (DMP). According to ZonMw’s procedure for research data management, ZonMw monitors the outcome of a researchers data management on the basis of a set of key items (‘kerngegevens’). The key items relate to several aspects of the FAIR principles. An explanation of these items is given below.

Researchers have to deliver information on these items – as far as possible - at the start and during their project. The full information is needed by the end of their project. Some of ZonMw’s programmes only ask for a selection of items. Please check the requirements of the programme and/or call for proposals.

ZonMw may publish the information on the key items on its website.

Key items also apply to other resources that underlie research results

Besides databases, also other resources should be made available for reuse, for instance: collections of biological material, lab journals, audio files, images, software, etc. In these cases the metadata that describe the resource must be findable (FAIR) for other researchers.

Note: for the text to be readable, we will refer to ‘databases’ or ‘datasets’ without mentioning the other resources.

Some frequently asked questions

Q1: What is the purpose of the key items of ZonMw?

A: Two reasons:

• Administrative: check whether the DMP has been carried out.
• Informative: The key items are in fact a minimal dataset for the outcome of RDM (research data management and stewardship). ZonMw wants to learn from RDM methods, tools and standards that are applied in projects. In that way, ZonMw can further improve / tailor RDM procedures for its research programmes. Also, ZonMw can stimulate the development and/or use of standards within research communities.

Q2: The key items are far too ambitious. Most research is not yet ready to meet them. Or some key items simply do not fit the type of research in a research community.

How does ZonMw deal with this problem?

A: The key items are important topics for reusable data. At the same time, ZonMw realises that there is still a lot to be developed for RDM. When filling in the scheme for key items, there is an option to state “not available”/ “not applicable”, and the reason for this (‘comply or explain’). All this information is input for further, community specific developments in ZonMw’s procedures.
RDM differs in some aspects for **quantitative and qualitative research**. It is ZonMw’s intention to accommodate both types. You can find more information in the guidance about Qualitative research data.

<table>
<thead>
<tr>
<th>The key items for a dataset (or metadata about other resources) are as follows:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The persistent Identifier (e.g. DOI-code) that refers to the dataset <em>(provide the code)</em></td>
</tr>
<tr>
<td>2. The digital repository (or data archive), preferably certified, where the dataset is deposited <em>(provide the link)</em></td>
</tr>
<tr>
<td>3. The online catalogue, or web portal where the dataset is registered or indexed and can be found (this may also be the repository) <em>(provide the link)</em></td>
</tr>
<tr>
<td>4. The terms of use giving the conditions for allowing other researchers to use the dataset <em>(provide the link, or persistent identifier)</em></td>
</tr>
<tr>
<td>5. The data format of the dataset, preferably machine readable <em>(provide the name, or link)</em></td>
</tr>
<tr>
<td>6. The terminology (e.g., code, classification, ontology) for the data <em>(provide the name, or link)</em></td>
</tr>
<tr>
<td>7. A metadata scheme that provides information about (the provenance of) the research data (preferably a scheme that is specific for your research discipline. If not available, use a generic scheme) <em>(provide the link)</em></td>
</tr>
</tbody>
</table>

**Background per key item**

**Key item 1: The Persistent Identifier referring to the dataset**

A Persistent Identifier (PI) is an online permanent referral to a digital object that is independent of its storage location. The digital object in this case is the dataset itself, or metadata that describe what the dataset is about (see key item 7). This PI is a unique ‘label’ (usually in the form of a code) and is created by a (certified) data archive or repository. With a PI the dataset, or a description of it, can always be found on the internet, even when the name or location of it is changed since its creation. They are essential for ensuring Findability and sustainable archiving of your dataset. In addition, a PI enables you to cite your data in publications. Examples of PI’s are DOI, Handle, URN of ARK.

For more information on PIs, see the web pages of the International DOI Foundation (IDF), Datacite (TUDelft). The course of RDNL is also a good option to learn more about the storing, managing, archiving and sharing of data.

You are required to provide ZonMw with the PI that is linked to your dataset at the end of your project.
Key item 2: The digital repository where the dataset is deposited

At the end of your project, you are required to deposit your data sustainably in a data repository (or a data archive). Preferably, this should be done in a certified repository (core trust seal), as listed here. A certified repository ensures that data can be shared in the long run.

Examples of Dutch data repositories with background information about data archiving are DANS, SURF and 4TU.Centre for research data.

If you are used to deposit your data on international repositories, you can keep on doing so. Especially when this is required by co-funders or scientific papers. In that case you do not have to deposit your data on a Dutch repository as well.

In case you decide to deposit your data in the local IT facility at your institution, be sure that the data set is curated, retrievable and accessible. In this case you are required to register your data in an online catalogue, or web portal to ensure that they are findable (key item 3).

Read more in the guidance themes Data repositories and catalogues and Overview of platforms, catalogues and repositories for searching and reusing databases and collections.

Read more in the guidance for Qualitative research data about facilities for archiving.

You are required to provide ZonMw with the link of the digital archive at the end of your project.

Key item 3: The online catalogue or web portal where the dataset is registered

In addition to the persistent identifier, the findability of your dataset will be enhanced by registering the dataset in an online catalogue or web portal with a search engine. In this way, the dataset is findable for other potential users. Research Data Alliance has formulated a clear definition.

You can register the dataset on such a catalogue. This does not mean however that the data themselves are stored there. Rather, you register information about your dataset (metadata), and provide a reference through a persistent identifier.

The information about the dataset may include title, description, research goal and contact information and conditions for getting access to the data (key item 4), etc. It may be generic information, or specific information aimed at a research community.

An example of an international catalogue of databases can be found on FAIRsharing.org.

If possible, you use a catalogue that is common in your field of research. A good example of a specialised catalogue is zorggegevens.nl. This online catalogue lists a large number of datasets/databases from health research, along with descriptive metadata.

More examples are listed in the guidance theme Overview of platforms, catalogues and repositories for searching and reusing databases and collections.
It is important to register your dataset in a catalogue. In that way other researchers know that certain data already exist. They can decide to contact the data producer to agree on reusing the data in new research, instead of producing similar data again. Also if it is not possible to share your dataset, for instance because of privacy-sensitive data or because of intellectual property, it is still important to show that the data are already there.

**Note:** if you deposit your data in a repository (**key item 2**) that also has a search engine and functions as an online catalogue, it is not necessary to register your dataset in another searchable resource.

You are required to provide ZonMw with link of the searchable resource in your dataset is registered, at the end of your project.

**Key item 4: the terms of use for other researchers to use the dataset**

ZonMw aims at reusable data according to its grant requirements. Therefore, ZonMw promotes researchers to create FAIR data. FAIR data, however, are not necessarily OPEN for anybody. With the ‘A’ of accessible in FAIR, you as a researcher can state the conditions by which the data will be shared.

If the reuse of your dataset is bound to specific conditions, or in other words there is restricted access to your dataset, other researchers must be able to view the terms of use. These must be findable online, e.g. through the website of your institute, or the catalogue or repository (**key item 2 and 3**).

The terms of use have to be made available by your institute or research group and should not be personal. Other researchers should be able to find out who they need to contact if they want to make use of the data.

The legal status of the licenses and conditions for reusing the data have to be clear. You can use international standards for the terms of use, or you can formulate them yourself together with a legal advisor.

Some of the criteria you can think of are:
- Collaboration in using the dataset, including agreements on publication and authorship
- The sharing of data for commercial purposes, taking into account the provisions of state aid law
- Conditions related to data security
- The approval of the participants allows for further research using this dataset, including linkage (privacy, informed consent)
- Agreements on methodology
- The way in which the data is made available
- The period of access
- Reimbursement for costs, e.g. in obtaining the data
- A committee decides whether access is granted
Read more information about privacy issues and access to qualitative data in the guidance about Qualitative research data.

You are required to provide ZonMw with a link or a PI to the terms of use at the end of your project.

EMBARGO?
It could be that an embargo period needs to be observed, during which your data will not be accessible, for example for publication, intellectual property belonging to certain companies or commercial interests associated with exploitation of the research results.

ZonMw’s grant terms and conditions stipulate a 3-month embargo, and a maximum of 9 months for patents.

Key item 5: The data format (preferably machine readable) of the dataset

To ensure long-term usability, accessibility and preservation of data, ZonMw recommends you to use a ‘preferred’ file format. DANS has made an extensive list of preferred formats. DANS is confident that preferred formats will offer the best long-term guarantees in terms of usability, accessibility and sustainability.

As a general guideline, DANS believes that the file formats best suited for long-term sustainability and accessibility:
• Are frequently used
• Have open specifications
• Are independent of specific software, developers or vendors

In practice, it is not always possible to use formats which satisfy all of these criteria.

Non-preferred formats are file formats that are widely used in addition to the preferred formats, and which will be moderately to reasonably usable, accessible and robust in the long term. DANS favours the use of preferred formats and recommends depositors to try to deposit data as much as possible in preferred formats.

ZonMw follows the recommendations of DANS for datasets from ZonMw’s projects.

Read more information about software and interoperability of qualitative data in the guidance about Qualitative research data.

You are required to provide ZonMw with a link to (or name of) the data format dataset at the end of your project.

Key item 6: The terminology (e.g., code, classification, ontology) for the data

For your data to be interoperable, understandable and usable for humans, but more importantly also for computers, you have to ‘speak a common language’. For instance, we agree to speak English or
Dutch. Another example is that we agree to register weight in kilograms. Moreover, an interoperable datasets can be linked or integrated with other datasets.

For your research data this means: use a terminology for recording your data (codes, classifications, ontologies) that are commonly used in your field of research.

This key item relates to a complex part of your DMP. **Make good arrangements for this aspect of your study from the start of the project.** It is much more work to do it in retrospect!

**For clinical research,** you will find more information at Nictiz ('third layer of interoperability'), and CASTOR (with tools for creating reusable data in clinical research, including the various layers of metadata, and coding metadata). A commonly used terminology in health care is SNOMED CT.

**For biomedical research** you can find more about ontologies on Bioportal.

**For qualitative research** you can find more information about software and interoperability of qualitative data in the guidance about Qualitative research data.

Within a single research project, multiple standards for data terminology can be used. For example, standards used in a small subfield of a scientific discipline, standards derived from a specific protocol or standards required by a renowned register for data archiving.

The most important thing is that you try to use – as much as possible - terminology standards etc. that are commonly used in your specific field of research, type of experiment, protocol and so on.

You are required to provide ZonMw with a link of the terminology standard(s) you use in your project, preferably linking to a document or webpage describing the standards.

**Key item 7: Metadata scheme (preferably specific for your research discipline)**

To make your dataset reusable, you should use a metadata scheme to describe your data. It is advised that you start with this fairly early in your project. After the project it is generally a lot harder and more work to do. There are already multiple standards available for different fields of research:

**FAIRsharing.org** provides an extensive catalogue of standards used in biomedical research. When you register your dataset in an online catalogue (**key item 3**) you are in fact already asked to provide a set of metadata. The metadata asked for an online catalogue are mainly focused on improving the findability and accessibility of your data. Also the ZonMw key items listed in this document, are an example of metadata that describe certain features of your dataset.

However, the metadata in schemes like those listed on FAIRsharing.org are far more detailed and domain specific. These specific metadata schemes are aimed at improving the interoperability and reusability of the data.
Finally, if there is no metadata scheme available yet for your research discipline, you can use a generic metadata scheme, for instance the one prescribed by the repository or catalogue, or one from Dublin Core.

More detailed information on metadata, see the guidance on Metadata on this site.

You are required to provide ZonMw with a link of the metadata scheme(s) that (you aim to) use with your dataset.