Guidance Qualitative research data

As is the case for quantitative research, also data from qualitative research (and from mixed methods) need to be reusable for verification and – ideally – for new research. However, at various points during your research data management planning you need to take different measures as compared to quantitative research. This is indicated in ZonMw’s DMP template. In case you come across DMP issues that do not fit the type of research you are planning to do, please indicate this in your DMP.

The following tools and procedures may be helpful to create reusable qualitative data, complying with (a number of) FAIR principles.

1. Data collection, processing and storage

Qualitative data may be textual, graphical, audio and video data. Several websites and tools provide an overview of aspects to consider during collection and analysis of qualitative data.

- Guidance Describing the data collection
- The Quality Handbook of EMGO
- The UK Data Archive
- The ‘Platform voor kwalitatief onderzoek’ KWALON (in Dutch)
- The research infrastructure for social sciences CLARIN and its Dutch node CLARIAH

The raw data as well as the transcripts / processed data need to stored according to the Gedragscode Wetenschappelijke integriteit.

2. Metadata

Metadata are ‘data about data’, describing e.g. the content of the database, and how it has been created (provenance). There are specific metadata schemes for qualitative data collections:

- The Data Documentation Initiative Alliance is for the documentation of, among other things, social science and health research.
- For designing a metadata scheme that fits your research community and type of qualitative data, it may be helpful to look into the framework for describing and reusing metadata blueprints: the Component MetaData Infrastructure (CMDI), initiated by CLARIN.
  The “components” are building blocks for description of your data. They can include field definitions, and they can be grouped into a ready-made description format (a “profile”). Both are stored and shared with other users in the Component Registry to promote reuse.

3. Privacy and access to qualitative data

Your data should be protected from misuse. On the ZonMw webpage for ‘wet- en regelgeving’ (in Dutch), you find more information on types of data, the corresponding privacy issues and protective measures. Also the ELSI Servicedesk of Health-RI may help.
For sensitive personal data (as is the case with health data) you should prevent that they lead to the identification of a natural person.

Anonymization and pseudonymization are possible when processing / transcribing qualitative data. However, in some cases it may not be possible to fully exclude identification. Identification of a person may come for instance from the context of the data, or voice recognition, etc.

If there is a risk for identification that you cannot prevent, please explain this in your DMP. In that case you still need to save your data, and describe them with metadata. However, you will need to restrict the options for reuse. Nevertheless, access to the data may still be possible under special safety conditions in e.g. CLARIAH’s media suite.

4. Software and interoperability

There are several software tools for the qualitative analysis of large bodies of textual, graphical, audio and video data. An example is ATLAS-ti.

For exchanging processed qualitative data between QDA software programs the Rotterdam Exchange Format Initiative (REFI) developed an interoperability standard (the QDA-XML standard). The software programmes include: ATLAS.ti, Dedoose, f4Analyse, MAXQDA, NVivo, QDAMiner, Quirkos and Transana.

See also KWALON’s news on ‘codebook exchange’.