

Best Practice: Assurance of metadata quality, version 1.0

A **high metadata quality** is crucial, **because** this ensures that your data

- can be found, reproduced and reused for secondary analysis in terms of good scientific practice.
- can be linked with other data and data types (such as associated publications).
- can be correctly cited in your own interest.
- becomes visible after the forwarding by da|ra to other international portals (DataCite Metadata Store) and citation indices (e. g. Data Citation Index).

To increase the international visibility it is important to complete the metadata fields in **English** as well. In addition, other **criteria** of quality metadata are to be noted¹:

- Accuracy, completeness, provenance, conformance to expectations, logical consistency and coherence, timeliness and accessibility of the metadata.

As the findability also significantly depends on the metadata quality it is advisable to consider the view of potential users of the research data and to reassess the provided metadata thereupon. Metadata quality is not to be confused with data quality (see also best practice paper data quality), the transition between the two is however fluent. Thus data can become metadata and vice versa.

The da|ra metadata schema 3.0 (doi:10.4232/10.mdsdoc.3.0) is the central prerequisite for securing metadata quality as it offers the possibility to add additional information about the resources in excess of the mandatory properties. This additional information is recommended by da|ra and is also in the interest of the publication agents. The da|ra metadata schema 3.0 is based on the **Data Documentation Initiative**² (DDI) **standard**. A formal assessment of metadata correctness is effected automatically through the reconciliation of the submitted metadata in xml format (via API, xml-upload) with the **xsd schema**. The assessment for users of the web interface takes place in the background and extensive help texts are available.

A **core set of ten mandatory fields** is the basic condition for the DOI-registration. The provision of **additional optional elements** as well as their child elements is recommended explicitly for the above mentioned reasons. These metadata fields include:

Classification | Keywords | Contributor | Alternative Identifier | Description | Geograph. Coverage

Authority files, e. g. the Integrated Authority File (GND) and further controlled vocabularies of the da|ra metadata schema give support to simplify and speed up the input.

High quality metadata examples from da|ra:

European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008) (Dataset)

http://bit.ly/metadata_EVS_1981-2008

Going on Pilgrimage Online: the Representation of Shia Rituals on the Internet (Text)

http://bit.ly/metadata_Going_on_Pilgrimage_Online

Statistical Information System GeroStat (Collection) http://bit.ly/metadata_GeroStat

xml examples can be found at <http://www.da-ra.de/en/technical-information/doi-registration/>.

In addition we offer [advice](#) on questions concerning the quality of your metadata.

¹ Bruce, Thomas R./Hillmann, Diane I. (2004): The continuum of metadata quality: defining, expressing, exploiting. In: Metadata in Practice. Chicago: American Library Association.

² Homepage of the Data Documentation Initiative: <http://www.ddialliance.org> (accessed 17.09.2014)