Overview of Metadata-elements to be provided to METIS for (shared) Datasets from DR.

Author: Ed Simons, project leider RDS/RIS.

Legenda:

- **Termen in bulletlist in de linkerkolom:** the actual metadata-elements are the entries listed in the bullet list. For the actual harvesting format, meaning the XML-schema corresponding to this metadata-list e.g. the one used to communicate between METIS and the RU Repository can be used; this a technical implementation aspect that the technical people can handle.
  In red: compulsory elements. Black: optional but recommended.
- *Italic texts: titles or headers to group metadat-elements*s.
- **Grey cells (on top):** automatically filled in by METIS, so not to include in the supplied metadata files..

Fields	*Value/Explanation*	*Remarks regarding the input from DR (Donders)*
• Result category	Has always the value: "dataset"	Does not have to be provided, will automatically be added in METIS.
• Status	During registration: Has the value: "input completed" / "submitted"; once a dataset has been published (via DANS EASY or the DR): Has the value: "publicly available" per default.	Does not have to be provided, since the metadata of all shared datasets are supposed to be "publicly available". ATTENTION: this does not mean that the dataset cannot have "Restricted Access", it just means that the metadata about the data can be exposed publicly.
Researcher metadata: • U or Z-number • ORCiD • Researcher ID • Surname • Initials • Prefix (van, de, etc…) • Title	The personnel number of the researcher (if RU). International person ID International person ID as used by WoS	DOI and Researcher ID if available. This would facilitate the linking of publications in METIS to the researchers and their datasets as wel as the linking of publications to the correct researchers harvested from WoS into METIS.

Dataset Metadata: • Title dataset • Subtitle • Archive • Year of archiving • Date available • Rightsholder *Dataset identifiers* • DOI • Handle • URL at archive • Description • Keywords		
	Name of the archive where the dataset is stored	
	The year in which the dataset is archived	
	(dd/mm/yyyy) The date from which you want the metadata to become available in DANS/Narcis (default, the date of registering the metadata).	
	The name of the institut(e)(ion) that formally holds the rights to the data and is allowed to grant access to the data.	Up to ow in METIS the default here is "Radboud University"
	International persistent Identifier for the dataset. Also (another kind of) identifier for the dataset In case of no persistent id, the URL pointing to the stored dataset.	In practice the dataset will have one of these three identifiers. So ONLY ONE of these has to be provided, but with the correct label (so that it is clear whether it is a DOI or a handle)
	Free text. (content) Description of the dataset (which research subject, the nature of the data, the way the data were collected and analysed)	
	Free text. List of terms, separated by ";" characterizing the data or its research contentwise.	This field can be used by DR to provide the MESH-classification terms/strings or any other keyword list they may use.

• Audiences	Predefined. The group of people (scientific discipline) the dataset is relevant for (a list of pre-defined choices is available, based on KNAW categories).	For DR-datasets some default choices out of the list could be used, e.g. "Health Sciences". Of course to be decided by Donders which to be used.
• Languages	Free text. The language(s) of the dataset and the accompanying documents, separated by ";"	
Temporal coverage • Collection period: from – to	(dd/mm/yyyy) Two date fields indicating the start and end date of the collection of the data.	
• Subject period	Free text. The (historical) period the data refers to (e.g. 1950-1960 or "Second world war", etc…)	
Spatial coverage • Area	Free text. Description of the area, region or country the dataset applies to.	
• Coordinates	Lattitude and longitude coordinates of either: the geographical spot (two coordinates) or the geographical region (box: 4 coordinates) the dataset applies to.	
• Related results	Id's (DOI, handle) of the publications related to the dataset.	The DOI's will be used as a key to link the bibliographical metadata of the publication to the dataset metadata, so that the datasets and the related publications can be reported and jointly exposed on the Internet, e.g. as intended by NARCIS.

		ATTENTION: the link between the dataset and publications based on the DOI as key, can only be realized if the publication-info is (already) present in METIS.
• Access level	In case of RIS datasets: "open access" or "restricted access". If a dataset has at least one "restricted access" file, the general level of the dataset becomes "restricted access". "Restricted access" means that access is granted (or denied) by the researcher (via RIS). "Open access" means that files are accessible, without interference of the researcher.	For the DR-datasets this means one entry "Restricted access" or "Open access" has to be provided, per dataset since for datasets registered elsewhere (i.e. not at DANS through RIS), we do not register separate file info.
• Format	Provides information on the type of files in a dataset. In RIS this is automatically detected by the system using file extensions without the "." (for instance: DOC, PDF, POR, SAV etc).	In its simplest form, there is only one format (if all files have the same format), if not a list of the various file formats present in the dataset has to be provided.
Project identification • Project id • Project title	Identifier of the project the dataset is a product/result of Title of the project the dataset is a product/result of Providing the project info allows to create overviews of datasets per project and could be of substantial value, if not necessary, for reporting to the project funder.	In case of externally funded projects, the ID provided/defined by the funder (e.g. EU, NWO). Otherwise the internal code as used within the Donders Institute.