



Metallic Elements Dissipation Avoided by Life cycle design for Steel

Deliverable 9.1

Data Management Plan and Implementation of
FAIR principles

Project details

Project Title	Metallic Elements Dissipation Avoided by Life cycle design for Steel		
Project Acronym	MEDALS		
Grant Agreement No.	101138516	Duration	48 months
Project Start Date	01-07-2024	Project End Date	30-06-2028

Document details

WP:	9	WP Leader:	KTH
WP Title:	Project Management		
Task:	9.1	Task Leader:	KTH
Task Title:	Task: Knowledge protection and technology transfer		
Deliverable No.	D9.1		
Deliverable Title	Data Management Plan and Implementation of FAIR principles		
Dissemination level	PU		
Written By	Björn Glaser, Sudhanshu Kuthe (KTH)		
Contributing beneficiary(ies)	KTH, SID, POLI, STEIL, TEC, LSA, UL, KOB		
Approved by	Björn Glaser		
Status	Version 1		
Date	10 December 2024		

Table of Content

Abstract	6
1 Introduction.....	7
2 The Data Management principles	8
2.1 Data proposing procedure	8
2.2 FAIR data.....	9
2.2.1 <i>Making data findable, including provisions for metadata</i>	9
2.2.2 <i>Making data openly accessible</i>	10
2.2.3 <i>Making data interoperable</i>	11
2.2.4 <i>Increase data re-use</i>	12
2.3 Allocation of resources	12
2.4 Data security	13
2.5 Ethical aspects.....	13
3 Individual Data Management Plans	14
3.1 MEDALS dataset no. 1	16
3.2 MEDALS dataset no. 2	17
3.3 MEDALS dataset no. 3	18
3.4 MEDALS dataset no. 4	19
3.5 MEDALS dataset no. 5	20
3.6 MEDALS dataset no. 6	21
3.7 MEDALS dataset no. 7	22
3.8 MEDALS dataset no. 8	23
3.9 MEDALS dataset no. 9	24
4 Conclusion	25
References.....	26
Appendix I Unified form for individual data plan.....	27

List of Tables

Table 1. Overview of data hosting and listing options	10
Table 2. User formats and compatible open software application.....	12
Table 3. Overview of the MEDALS datasets.	14
Table 4. Unified form, for collecting and presenting basic deatils about data.	27

Abbreviations and acronyms

DMP	Data Management Plan
DOI	Digital Object Identifier
IPR	Intellectual Property Rights
NDA	Non-disclosure Agreement
ORD	Open Research Data
PU	Public
SEN	Sensitive
TC	Technical Committee
tbd	To be declared (will be updated in future version)
PU	Public
URL	Uniform Resource Locator
WP	Workpackage

Abstract

The present document constitutes the deliverable D9.1 “Data Management Plan and Implementation of FAIR principles”. This Data Management Plan (DMP) defines data management policy that will be used within the MEDALS project. Moreover, the DMP lists the most important datasets that are (or will be) gathered, processed and/or generated. Individual DMP is provided for each dataset using the unified form that simplifies the collection of the most relevant dataset information. As the majority of the datasets are (or will be) derived from the existing industrial set-ups, it is reasonable to expect that Intellectual property rights (IPR) issues will severely limit the final data openness. Even though, in long term the DMP management activities are expected to positively effect on the reuse of the research data, due to improved quality and clarity of the datasets. During the project lifetime this DMP will be regularly reviewed and updated by the consortium partners, in order to reflect current project progress.



1 Introduction

This document represents the Data Management Plan (DMP) of the MEDALS project, which has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No 101058694 [1].

The DMP aims to ensure that the project data is soundly managed and follows the FAIR principles (findable, accessible, interoperable and reusable) whenever this is possible. Moreover, the DMP ensures that MEDALS activities are compliant with the Horizon Europe Open Access policy and the recommendations of the Open Research Data (ORD) pilot, which aims to improve and maximize access to and re-use of research data generated by Horizon Europe projects. The ORD pilot applies primarily to the data needed to validate the results presented in scientific publications, but also other data can be provided by the beneficiaries on a voluntary basis. Participation in the ORD Pilot does not necessarily mean to open all research data. Rather, the ORD pilot follows the principle “as open as possible, as closed as necessary”, in order to balance openness and protection of scientific information, commercialization and IPR, privacy concerns, security as well as data management and preservation questions.

The purpose of this DMP is to provide an analysis of the main elements of the data management policy that will be used in the MEDALS project by the consortium with regard to the project research data. The DMP describes the data management life cycle for the research data to be collected, processed and/or generated. Furthermore, it describes what methodology and standards will be adopted, whether and how this data will be shared and/or made open and how this data will be curated and preserved during and after the project.

The DMP is not a fixed document, rather it is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. It is envisioned that the DMP will be updated at least twice during the project lifespan. Additional ad hoc updates are also expected when deemed necessary in order to include new datasets, reflect changes in the already identified datasets, changes in consortium policies and plans or other potential external factors. KTH is responsible for the preparation of the DMP.

2 The Data Management principles

The research data that will be generated during the MEDALS project can be divided within four classes: (i) acquired observational data (e.g. survey data, sensor data, and production overview), (ii) experimental data from laboratory equipment (e.g. experimental setup), (iii) simulation data generated from models, and (iv) derived / processed data. Most of the MEDALS's research data is envisioned to be generated within the industrial environment and existing industrial set-ups. Consequently, it is expected that not all of the data will be opened, since this could lead to leak of sensitive information that would jeopardize the competitiveness and market position of the involved industrial partners. Moreover, it is expected that not all MEDALS datasets will be available to every consortium partner as this could in some cases violate the IPR of individual partners. Therefore, bi-lateral NDA (non-disclosure agreements) may be needed in order to make this data accessible only for selected partners.

Even though, the MEDALS project will join the ORD pilot in order to support the paradigm shift towards openness. The consortium partners will strive to maximize openness and follow the FAIR principles whenever possible. It is expected from all consortium partners that if it is possible, they will tend to apply data curation measures in order to hide the sensitive IPR information and prepare the derived datasets that could be opened. All MEDALS data that underlies scientific publications shall be made available via open-access online platforms, unless subject to protection, or if release of all or part of the data to open-access platforms would jeopardize the action's main objective.

Due to the sensitive nature of the data gathered and generated within the framework of the MEDALS project, individual data management plans will be declared for each dataset (see Section 3). On the other hand, this section will provide the general data management principles that will be followed by all of the consortium partners.

2.1 Data proposing procedure

All consortium partners are responsible to identify the appropriate data that should be managed within the DMP (either internal/restricted or open data). The data owners (i.e. project partners that have been involved in the data capture/generation and preparation) will provide information about the data using the template form (see **Appendix I**). The data description form (individual data management plan) will be sent to the KTH, who will be the main editor of the DMP document.

2.2 FAIR data

2.2.1 Making data findable, including provisions for metadata

Each MEDALS dataset will get a unique name using the following convention:

[data name] (MEDALS dataset, v [ZZ]):

- [data name]: short and unique data name
- v [ZZ]: version of dataset.

With the unique data naming convention the dataset could be referenced and identified through the project documentation and internal project storage platform. Moreover, the version of the data is a part of the data name in order to differentiate data versions. Whenever the data will be made public, it will also get the Digital Object Identifier (DOI) which is a persistent and unique identifier. DOI is an alphanumeric string assigned to uniquely identify an object. It is tied to a metadata description of the object as well as to a digital location, such as a Uniform Resource Locator (URL), where all the details about the object are accessible. The DOI numbers will be provided either by the publisher when publishing data in an article (i.e. in a scientific journal) or by using the repositories that generate also DOI numbering. For this purpose, several different free services can be used, such as ResearchGate [2] and DataCite [3], (refer **Table 1**). Each specific version of data that will be declared as public will get only one DOI, which will serve as the main digital identifier of the data as well as to a digital location. This DOI should be consistently used in order to reference the data for reporting and dissemination purposes. If the data will not get a DOI (private/internal data), a unique dataset name should be used for referencing purposes, following the convention described above. When uploading the data either on a public or an internal (i.e. owned by one of the MEDALS partners) repository, the data should be accompanied by information that will facilitate their understanding and re-use by interested stakeholders. The basic information that will assist interested stakeholders and a systematic overview of the data properties will be described using the form presented in Appendix I. The form consists of a set of questions and serve as a template to collect and present basic details about data, including its format and file type as well as meaningful information about who created or contributed to the dataset, its name and reference, date of creation, under what conditions it may be accessed etc.

Table 1. Overview of data hosting and listing options

Service provider	Type of service	Costs	Intended use	Retention period
KTH	Microsoft Teams server on private server at KTH	Free of charge	Upload and backup of internal and open data, internal and public links, access restrictions	at least 3 years after the end of project
KTH	MEDALS website hosted on server at KTH	Free of charge	Data listing	at least 3 years after the end of project lifespan
ResearchGate[2]	Add your research data	Free of charge	Data hosting (up to 512MB), data listing, DOI generation	unlimited
DataCite [3]	Registering DOIs	organization must be a member of DataCite	DOI generation and registering	unlimited

2.2.2 Making data openly accessible

In order to expand the impact of the MEDALS project, the consortium partners will maximize their effort to facilitate sharing of results and deliverables within and beyond the consortium. Selected data and results will be shared with the scientific community and other stakeholders through publications in scientific journals and presentations at conferences, as well as through open access data repositories. All data owners should agree on data openness. If the data cannot be made open, the data owners will clearly justify why data cannot be made openly accessible. The main elements when considering confidentiality of datasets are:

- To protect intellectual property regarding underlying processes, products and technologies where the data could be used to derive sensitive information that would impact the competitive advantage of the consortium or its members.

- Commercial agreements as part of the procurements of components or materials that might foresee the confidentiality of data.
- Personal data that might have been collected in the project where sharing them is not allowed by the national and European legislation.

A subset of data or modified data could be generated out of the original data in order to provide an open version of original data. Generally, all research data underpinning publications should be made available for verification and re-use unless there are justified reasons for keeping specific datasets confidential. For data storage and backup different services can be used that offer open access. The list of the services that will be used within the MEDALS project are overviewed in **Table 1**. Specific data can be stored in more than one repository. Consortium partners will have a possibility to store their datasets in a repository governed by the project coordinator (i.e. Microsoft Teams service on private server at KTH). Together with the data, supplementary material should be stored, that allows interpretation and use of the data (e.g. data summary, the associated metadata, documentation and code). If data accessibility problems or restrictions occur, the interested parties should directly contact a contact person declared within the individual data management plan. The contact person is also in charge to grant additional data access permissions of the internal data or declare the individual conditions under which the data can be accessed.

2.2.3 Making data interoperable

To maximize the data interoperability (i.e. allowing data exchange and re-use between researchers, institutions, organizations, countries, etc.) the MEDALS data will be provided in standard formats that could as well be read by free and/or open software applications. That will simplify the data uptake and facilitate re-combinations with different datasets from different origins. **Table 2** lists some of the standard formats and the accompanying free and/or open software applications that could be used. If some other formats are used, the author of the data should also provide information of the appropriate software application that could be used to view or edit the data.

Table 2. User formats and compatible open software application.

Standard format	Free and/or open software application
CSV, TXT	Notepad++
SVG, EPS	Inkscape
PDF	Adobe Reader DC, WPS office
DOC, DOCX, XLS, XLSX, ODF, PPT, PPTX	OpenOffice, WPS office
AVI	VLC, opencv-python
DAT	ibaAnalyzer

Complementary data documentation should also be provided to encompass details on vocabularies and units of measurements. Wherever possible consortium partners will identify and use standard data and metadata vocabularies (e.g. IEC62264/ISA95 terminology, domain vocabulary) to maximize the data interoperability.

2.2.4 Increase data re-use

The permitted re-use of the MEDALS open data will be defined by the choice of the appropriate user license. The consortium partners are encouraged to use one of the following Creative Commons Licenses:

Creative Commons Attribution (CC BY): lets others distribute and copy the article, to create extracts, abstracts, and other revised versions, adaptations or derivative works of or from an article (such as a translation), to include in a collective work (such as an anthology), to text or data mine the article, even for commercial purposes, as long as they credit the author(s), do not represent the author as endorsing their adaptation of the article, and do not modify the article in such a way as to damage the author's honor or reputation.

Creative Commons Attribution-Non Commercial-No Derives (CC BY-NC-ND): for non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.

2.3 Allocation of resources

Costs related to open-access to research data in Horizon Europe are eligible for reimbursement under the conditions defined in the Horizon Europe Grant Agreement, in particular Article 6 and Article 6.2.D.3, but also other articles relevant for the cost category chosen. Costs cannot be claimed



retrospectively. Project beneficiaries will be responsible for applying for reimbursement of costs related to making data accessible to others beyond the consortium.

2.4 Data security

MEDALS will use methods that emphasize easy access and trust among the participants. The following guidelines will be followed to ensure the security of the data:

- Store data in at least two separate locations to avoid loss of data (e.g. MEDALS repository and local storage at data owner's location for backup).
- MEDALS repository (Own-Cloud) use a RAID array to provide data redundancy.
- Encrypt data if it is deemed necessary by the participating researchers.
- Label files and versions in a systematically structured way to ensure coherence of the final data.

2.5 Ethical aspects

Currently, no additional ethical or legal issues have been identified. This will be updated in case if any ethical or legal issues arise during the project lifespan.

Informed consent for gathering the data using the survey is included. Raw survey data will be shared only among the consortium partners as defined in the informed consent. Aggregated, summarized and processed data could be made open and also included within the publicly open project deliverables.

3 Individual Data Management Plans

This section will develop as the project progresses and reflects the current status within the consortium about the primary data that will be produced. This section provides meaningful information in this respect per each identified dataset. Short overview of the datasets is presented in **Table 3** and detailed overview of the MEDALS datasets is presented in the following subsections. Note, that not all listed datasets are available at the moment. The datasets that are planned to be collected and organized during the project lifetime have listed empty dataset version (i.e. v[xx]). When the data will be collected and organized the actual dataset version will be provided.

Table 3. Overview of the MEDALS datasets.

Data no.	Data name	Version	Data type	Origin	Open [y/n]
1	Users survey	v[xx]	Questionnaire	MEDALS validation survey	y
2	Non-users survey	v[xx]	Questionnaire	MEDALS validation survey	y
3	LSA specification and requirement of the process	v[xx]	Laboratory data	Analysis	n
4	TEC Trials and results on the use of pyrometallurgical technologies for the valorization of residues from new metallurgical processes for the tramp elements removal of the liquid steels	v[xx]	Acquired observational and experimental data	Characterization tests, Experimental trials	n
5	KTH model-related data	v[xx]	Derived process data	Models	n

6	STE specification and requirement of the process	v[xx]	Historical process lags	Production	n
7	KOB specification and requirement of the process	v[xx]	Derived process data	Models	n
8	SID specification and requirement of the process	v[xx]	Laboratory data	Production	n
9	UL computational data	v[xx]	Derived process data	Models	n

3.1 MEDALS dataset no. 1

Data identification	
Data name	Users Survey (MEDALS dataset, v[xx])
Data type	Questionnaire answers
DOI	/
Date of creation	tbd
Contact person	Björn Glaser
Data summary	
Data description, purpose and relation to the project objectives	tbd
Data origin	MEDALS validation survey
Contributors and roles	KTH (organizing and collecting the survey data)
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	questionnaire
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	y
Level of accessibility	Raw survey data will be accessible to consortium partners, aggregated and pre-processed survey results will be open to public
Justification	/
License type	tbd
Searchable keywords	tbd

3.2 MEDALS dataset no. 2

Data identification	
Data name	Non-users survey (MEDALS dataset, v[xx])
Data type	Questionnaire answers
DOI	/
Date of creation	tbd
Contact person	Björn Glaser
Data summary	
Data description, purpose and relation to the project objectives	tbd
Data origin	MEDALS validation survey
Contributors and roles	KTH (organizing and collecting the survey data)
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	quationnaire
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	y
Level of accessibility	Raw survey data will be accessible to consortium partners, aggregated and pre-processed survey results will be open to public
Justification	/
License type	tbd
Searchable keywords	tbd

3.3 MEDALS dataset no. 3

Data identification	
Data name	LSA specification and requirement of the process (MEDALS dataset, v[xx])
Data type	Historical process logs, database table/views export
DOI	/
Date of creation	01.07.2024-30.06.2028
Contact person	
Data summary	
Data description, purpose and relation to the project objectives	Data will be used to identify and verify the deep learning based models
Data origin	Sensor data
Contributors and roles	LSA
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	tbd
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	n
Level of accessibility	Raw data will be accessible only for internal use in LSA
Justification	/
License type	tbd
Searchable keywords	LSA, verification

3.4 MEDALS dataset no. 4

Data identification	
Data name	TEC Trials and results on the use of pyrometallurgical technologies for the valorization of residues from new metallurgical processes for the tramp elements removal of the liquid steels (MEDALS dataset, v[xx])
Data type	Characterization, tests and trials results
DOI	/
Date of creation	01.07.2024-30.06.2028
Contact person	Teresa Gutierrez Seco
Data summary	
Data description, purpose and relation to the project objectives	Results obtained in WP5 and WP6 from waste characterization, agglomeration tests and metallurgical trials. These data will serve as a basis for the definition of critical parameters for the recovery of Cu and other high value elements from steelmaking residues
Data origin	TEC tests and trials
Contributors and roles	TEC (define/prepare/perform the characterization/tests/metallurgical trials activities) and SID (provide residues and define requirements)
Size of data	tbd
Data utility	tSteel makers, recyclers, etc..
Standards and conventions	
Methodologies for collection or generation	tbd
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	TECNALIA server
Retention period(s)	at least 3 years after the end of the project
Data exploitation and sharing	
Data openly available	n
Level of accessibility	Raw data will be accessible only for internal use in TEC and SID
Justification	Process confidentiality/ IPR
License type	tbd
Searchable keywords	TEC, steelmaking waste/residues/by_products valorization, copper recovery, pyrometallurgical technologies

3.5 MEDALS dataset no. 5

Data identification	
Data name	KTH model related data (MEDALS dataset, v[xx])
Data type	Derived process data
DOI	/
Date of creation	01.07.2024-30.06.2028
Contact person	Björn Glaser
Data summary	
Data description, purpose and relation to the project objectives	Data from output of the models
Data origin	Process data from models
Contributors and roles	KTH
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	tbd
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	n
Level of accessibility	Raw data will be accessible only for internal use in KTH
Justification	/
License type	tbd
Searchable keywords	KTH, model, verification

3.6 MEDALS dataset no. 6

Data identification	
Data name	STE specification and requirement of the process (MEDALS dataset, v[xx])
Data type	Historical process logs, database table/views export
DOI	/
Date of creation	01.07.2024-30.06.2028
Contact person	Philipp Döhr
Data summary	
Data description, purpose and relation to the project objectives	Data will be
Data origin	STE production
Contributors and roles	STE
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	tbd
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	n
Level of accessibility	Raw data will be accessible only for internal use in STE
Justification	/
License type	tbd
Searchable keywords	STE, process data, scrap

3.7 MEDALS dataset no. 7

Data identification	
Data name	KOB specification and requirement of the process (MEDALS dataset, v[xx])
Data type	Historical process logs, database table/views export
DOI	/
Date of creation	01.07.2024-30.06.2028
Contact person	Rutger Gyllenram
Data summary	
Data description, purpose and relation to the project objectives	Data will be used to identify and verify the model of KOB process
Data origin	KOB data
Contributors and roles	KOB
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	tbd
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	n
Level of accessibility	Raw data will be accessible only for internal use in KOB
Justification	/
License type	tbd
Searchable keywords	KOB, process data

3.8 MEDALS dataset no. 8

Data identification	
Data name	SID specification and requirement of the process (MEDALS dataset, v[xx])
Data type	Database table/views export
DOI	/
Date of creation	01.07.2024-30.06.2028
Contact person	Asier Arteaga Ayarza
Data summary	
Data description, purpose and relation to the project objectives	
Data origin	SID analysis
Contributors and roles	SID
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	tbd
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	n
Level of accessibility	Raw data will be accessible only for internal use in SID
Justification	/
License type	tbd
Searchable keywords	SID, Laboratory data, sampling analysis

3.9 MEDALS dataset no. 9

Data identification	
Data name	UL computational data (MEDALS dataset, v[xx])
Data type	Acquired observational data
DOI	/
Date of creation	01.07.2024-30.06.2028
Contact person	Norma.Bargary
Data summary	
Data description, purpose and relation to the project objectives	
Data origin	Models, simulations and analysis
Contributors and roles	UL
Size of data	tbd
Data utility	tbd
Standards and conventions	
Methodologies for collection or generation	tbd
Format of data	tbd
Metadata	tbd
Archiving and preservation	
Location(s)	tbd
Retention period(s)	tbd
Data exploitation and sharing	
Data openly available	n
Level of accessibility	Raw data will be accessible only for internal use at UL
Justification	/
License type	tbd
Searchable keywords	UL, model, simulation

4 Conclusion

This Data Management Plan provides procedures for managing the data that will be produced within the framework of the MEDALS project. Moreover, the DMP provides the unified form that will be used to collect and represent the basic information about the MEDALS's data. A list of identified data is provided and will be updated throughout the project timespan. Since the data will mostly be collected from the industrial set-ups, it is expected that IPR issues provide the main obstacle to open the data. Nevertheless, the MEDALS will join the ORD pilot. The data will be managed according to this document and periodic evaluation will be performed to check if any identified data sets or parts of these datasets could be made open.



References

1. MEDALS grant agreement Project no. 101138516 HORIZON-CL4-2023-TWIN-TRANSITION-01-45
2. ResearchGate. Available at: <https://www.researchgate.net/> (Accessed: November 25, 2024).
3. DataCite Support. Available at: <https://support.datacite.org/> (Accessed: November 25, 2024).



Appendix I Unified form for individual data plan

The Table 4 represent the template that should be used in the MEDALS project to collect and represent the information of the project data.

Table 4. Unified form, for collecting and presenting basic deatils about data.

Data identification	
Data name	[short data name] (MEDALS dataset, v[xx])
Data type	[user input]
DOI	[user input]
Date of creation	[user input]
Contact person	[user input]
Data summary	
Data description, purpose and relation to the project objectives	[user input]
Data origin	[user input]
Contributors and roles	[user input]
Size of data	[user input]
Data utility	[user input]
Standards and conventions	
Methodologies for collection or generation	[user input]
Format of data	[user input]
Metadata	[user input]
Archiving and preservation	
Location(s)	[user input]
Retention period(s)	[user input]
Data exploitation and sharing	
Data openly available	[y/n]
Level of accessibility	[user input]
Justification	[user input]
License type	[user input]
Searchable keywords	[user input]

More details about the expected content of the form are given below:

- *Data name* – Short and unique name of the data following the convention
- *Data type* – The type of study in the frame of which the dataset is produced.
- *DOI (if available)* – If the dataset is public DOI number must be provided
- *Date of creation* – Date of creation of the dataset.
- *Contact person* – Contact details of the person in charge of the dataset responsible to provide additional information and support about the data.
- *Data description, purpose and relation to the project objectives* – A concise description of the data together with references to the MEDALS objectives, WPs and Tasks.
- *Data origin* – From where the data was captured/generated?
- *Contributors and roles* – List of project partners that were involved in data capture and preparation and their roles.
- *Size of data (actual/estimated)* – Size of the data in GB (actual or estimated if the data is still to be created/finalized).
- *Data utility* – Potential external stakeholders for whom the data may prove useful (to whom it might be useful).
- *Methodologies for collection or generation* – The methodology and tools employed for collecting/generating the data.
- *Format of data* – The format and volume of the dataset.
- *Metadata* – Metadata reference or clear presentation of the metadata that could be used to reason the data.
- *Location* – Where will the data and associated metadata, documentation and code be deposited (storage and backup)?
- *Retention period* – Length of time for which the data will remain re-usable
- *Data openly available (y/n)* – Will the data be openly available?
- *Level of accessibility* – If data is not openly available define who and under what conditions can access and use the data (provisions regarding the confidentiality of the data).
- *Justification* – Justification why data is not openly available.
- *License type* – Under which license the data will be openly available.
- *Searchable keywords* – General keyword that could be used for search engines.