

Data Management Plan

Contents

1. General information	4
1.1. When to prepare a DMP	4
1.2. Types of data	4
1.3. Metadata	4
1.4. Data preservation	4
1.4.1. Data selection	5
1.4.2. Formats of data for long-term preservation	5
1.5. Data sharing	6
1.5.1. Personal and Sensitive data	6
1.5.1.1. Swiss law	6
1.5.1.2. European law: GDPR	7
1.5.1.3. Informed consent	7
1.5.2. Licenses	7
1.6. Data repositories	8
2. DMP for SNSF projects	9
2.1. Preparation of the DMP	9
2.2. Eligible costs under SNSF grants	9
2.3. Structure of the DMP	9
2.3.1. Data collection and documentation	10
2.3.2. Ethics, legal and security issues	10
2.3.3. Data storage and preservation	11
2.3.4. Data sharing and reuse	12
2.4. Examples	14
2.4.1. DMP – ARC	14
2.4.2. DMP – COM	16
2.4.3. DMP – COM	19
2.4.4. DMP – ECO	23
2.4.5. DMP – INF	25
3. DMP for H2020 projects	27
3.1. Preparation of the DMP	27
3.2. Structure of the DMP	27

Introduction

The aim of a Data Management Plan is to plan and manage the whole life cycle of data¹. It offers a long-term perspective by outlining how data will be generated, collected, documented, shared and preserved.

Managing and sharing research data as openly as possible is one of the principles of good scientific practice. Following the SNSF Funding Regulations “*grantees are obliged to make available to the public in an appropriate manner the research results obtained with the help of SNSF funding*”. Similarly, Beneficiaries of EU funding are expected to share their data according to the FAIR Data principles. Data should be Findable, Accessible, Interoperable and Reusable.

In the following sub-chapters you will find general information on data and data management. Chapter 2 provides the structure and required content of the SNSF Data Management Plan, with standard answers for questions 2.2 and 3.1 and reference to the respective information sections of chapter 1. Chapter 3 includes some examples of DMPs prepared for projects of various disciplines.

¹ Data life cycle: data management planning – data creation – data processing – data analysis – data preservation – data publication (archiving) – data reuse.

1. General information

1.1. When to prepare a DMP

The preparation of a DMP should start during the project-planning phase. At this stage it is an useful instrument to properly plan your data creation and analysis, as well as to address relevant issues related to data storage and data sharing. Nevertheless, the DMP is a living document that remains editable for the whole duration of a research project. Information can be specified or updated accordingly to the project progresses. At the latest the DMP should be revised by the end of the project in view of the data preservation, publication and reuse. For further information on SNSF and H2020 deadlines for the DMP, see [§2.1](#) and [§3.1](#) respectively.

1.2. Types of data

Research data are data collected or produced in the course of scholarly activity which are used for the purposes of academic research or which document research findings.

There are different types of data:

- **Observational Data:** data captured in-situ, can't be recaptured, recreated or replaced.
Examples: sensor readings, sensory (human) observations, survey results, interview notes, transcripts.
- **Experimental Data:** data collected under controlled conditions, in situ or laboratory-based; should be reproducible, but can be expensive.
Examples: gene sequences, chromatograms, spectroscopy, microscopy.
- **Simulation Data:** result from using a model to study the behavior and performance of an actual or theoretic system, models and metadata, where the input can be more important than output data.
Examples: climate models, economic models, biogeochemical models.
- **Derived/Compiled Data:** reproducible; but can be very expensive.
Examples: derived variables, compiled database, 3D models.
- **Reference or canonical Data:** static or organic collection (peer-reviewed) datasets, most probably published and/or curated.
Examples: gene sequence databanks, chemical structures, census data, spatial data portals.

1.3. Metadata

Metadata is documentation that describes data. Properly describing and documenting data allows users (yourself included) to understand and track important details of the work. Metadata contains generally descriptive information like the creator, date and subject of the dataset, information that tells how the dataset relates to other documents and information on the format, access rights or preservation aspects.

It is important to capture this information right from the start of the project. Sometimes metadata is embedded in the dataset as it is automatically created by the software or tool that generates the data. If not, metadata can be registered in a separate file, like a readme.txt, spreadsheet or xml document.

Many academic disciplines have formalized specific metadata standards. You can check them out on the [Digital Curation Center](#) website. Another list of technical metadata standards is provided by the [research data Alliance](#).

1.4. Data preservation

By the end of the project you have to decide which of your research data should be preserved and how. If possible, research data should be archived and published in a dedicated data repository (see [§Error! Reference source not found.](#) and [§1.6](#)). It might o

occur that some data cannot be made publicly available and should be preserved only on institutional file servers with restricted access. In both cases data preservation must be properly planned.

Personal and sensitive data should be handled and preserved in line with the relevant legal framework and informed consent (see §1.5.1.3).

1.4.1. Data selection

It is not always feasible nor appropriate to preserve all data generated or collected during a research project. You might focus on the following:

- data which are needed to validate research findings;
- data that cannot easily be recreated or produced;
- data that is expensive to reproduce;
- experimental data;
- third party or acquired data for which the future availability is uncertain;
- data needed for further reuse.

In addition, you should also consider whether you will need to preserve multiple versions of a file or whether the most recent version will be sufficient for preservation.

You should also think about for how long should the data be preserved. In general, it is best to commit to a minimum amount of time than a maximum. However, there is no fixed amount of time for which data must be preserved. A good practice to establish a minimum amount of time may be to consider the amount of time it takes for a research paper to be cited and then to add 5 years.

It is also useful to set a timeframe after which the preservation of your research data should be re-evaluated.

1.4.2. Formats of data for long-term preservation

Since hardware and software may become obsolete over time, data should be converted into standard or open formats for long-term accessibility and preservation. The following table provides a list of appropriate data formats that you could use and formats that you should try to avoid for data preservation:

Type of data	Appropriate	Acceptable	Not suitable
Tabular data with extensive metadata	.csv, .hdf5	.txt, .html, .tex, .por	
Tabular data with minimal metadata	.csv, .tab, .ods	.xml, .xlsx, SQL script	.xls, .xlsb
Textual data	.pdf, .txt, .odt, .odm, .tex, .md, .htm, .xml	.pptx, .pdf with embedded forms, .rtf, docx	.doc, .ppt
Code	.m, .R, .py, .iypnb, .rstudio, .rmd, NetCDF	.sdd	.mat, .rdata
Digital image data	.tif, .png, .svg, .jpeg	.jpg, .jp2, .tif, .tiff, .odf, .gif, .bmp	.indd, .ait, .psd
Digital audio data	.flac, .wav, .ogg	.mp3, .mp4, .aif	
Digital video data	.mp4, .mj2, .avi, .mkv	.ogm, .webm	.wmv, .mov
Generic data	.xml, .json, .rdf		

1.5. Data sharing

Research data should be shared as soon as possible, but at the latest together with the relevant scientific publication. Besides responding to public requirements, sharing your data helps generating more visibility to your research.

Datasets must always be carefully documented with associated metadata, such that other researchers understand how the data was collected, as well as under which conditions and how it can be re-used.

The FAIR Data Principles define a range of qualities a published dataset should have in order to be Findable, Accessible, Interoperable and Reusable (see [Explanation of the FAIR Data Principles](#)). The SNSF as well as the EU expects researchers to share their data according to the FAIR Data Principles on publicly accessible, digital repositories. It is important to note that the FAIR Data Principles do not require researchers to share all their data without any restriction.

1.5.1. Personal and Sensitive data

Personal data is all information relating to an identified or identifiable person (Swiss FADP², article 3 a.), such as name, address, identification number, e-mail, phone number, medical records, etc.

Sensitive data, according to the Swiss FADP (article 3 c.) is data on:

- religious, ideological, political or trade union-related views or activities,
- health, the intimate sphere or the racial origin,
- social security measures,
- administrative or criminal proceedings and sanctions

In order to protect personal and sensitive data you must anonymize your datasets before publication. Anonymization has to be irreversible; data identifiers need to be removed, generalised, aggregated or distorted. Anonymous data are not considered as personal data since they do not correspond to an identified or identifiable person.

Data pseudonymization is instead a method of de-identification that replaces identifiers with pseudonyms or identifiers that are generated by the researcher. Using pseudonyms allows researchers to link de-identified data to the same individual across multiple datasets while retaining confidentiality of the individual. This means that, unlike anonymized data, pseudonymized data can be linked across datasets. Linking across datasets can make data more useful, but it can also increase the risk of re-identification.

1.5.1.1. Swiss law

In Switzerland, the processing of personal and sensitive data is subject to the [Federal Act on Data Protection](#) and to cantonal laws.

The collection of personal and sensitive data is only possible with prior and informed consent of the person concerned or when such collection is explicitly provided for in the law. Unless anonymised, the data may be processed only for the purpose indicated at the time of collection.

The Swiss law on data protection is currently under revision, also in order to better align with the European law (GDPR), and is expected to entry into force in 2021. Please note that European law (GDPR) is more demanding in terms of consent and data protection, we therefore suggest to always take into account the European legal framework when dealing with personal data (see next section).

² [Federal Act on Data Protection](#).

1.5.1.2. European law: GDPR

The General Data Protection Regulations (GDPR) regulate data protection and privacy for any individual residing within EU, as well as the communication of personal data outside EU. In the field of research, this regulation applies to all institutions and companies operating internationally that collect and process personal data of EU residents or send data from Swiss nationals abroad.

Complying with GDPR is a good practice for all research projects, regardless of the source and use of data.

Personal data may only be collected for specified, explicit and legitimate purposes, which should in principle be defined prior to processing and brought to the attention of the data subjects. GDPR requires that in case of collection and processing of personal data, consent must be expressed freely and explicitly. Therefore Consent will only constitute an appropriate legal basis if the person concerned has a real control and choice as to whether or not to accept the proposed conditions or to refuse them without prejudice.

1.5.1.3. Informed consent

Following the European law, informed consent is required for the processing of personal and sensitive data and must fulfill the following criteria:

- **Free:** the person must not feel compelled to consent and his or her consent must not be conditional on the granting of an advantage.
- **Specific:** consent must be obtained for each purpose and not for a set of purposes.
- **Informed:** the person must be informed on the identity of the controller, the purposes of the processing operation, the legal bases, etc.
- **Unambiguous:** consent must be given by a clear positive act. The controller must then keep proof of the given consent.

A template of the informed consent can be requested to the Research and Transfer Service. Further guidelines are provided by FORS.

1.5.2. Licenses

A license clarifies the terms of use of your work allowing you to share your data with the proper protection. By default, a work with no license is a copyrighted work (all right reserved).

A Creative Common (CC) license is one of several public copyright licenses that enable free distribution of an otherwise copyrighted work. It is used when an author wants to give other people the right to share, use and build upon a work that he has created. Please find below an explanation of different CC licenses, from the most open to the most restrictive:

License	Right	Description
CC0	Free	Completely free.
CC-BY	Attribution	Licensees may copy, distribute, display and perform the work and make derivative works and remixes based on it only if they give the author or licensor the credits (attribution).
CC-BY-SA	Attribution Share-alike	Attribution + Licensees may distribute derivative works only under a license identical ("not more restrictive") to the license that governs the original work.
CC-BY-NC	Attribution Non commercial	Attribution + Licensees may copy, distribute, display, and perform the work and make derivative works and remixes based on it only for non-commercial purposes.
CC-BY-NC-SA	Attribution Non commercial Share-alike	See above.

CC-BY-ND	Attribution No derivative works	Attribution + Licensees may copy, distribute, display and perform only verbatim copies of the work, not derivative works and remixes based on it.
CC-BY-NC-ND	Attribution Non commercial No derivative works	See above.

1.6. Data repositories

The public funding bodies require grantees to publish their data in non-commercial data repositories, which allow publishing of FAIR data. You are free to choose field-specific repositories (e.g. PubMed Central for Life Sciences or ArXiv for Engineering) or repositories that accept datasets from different research fields.

Publication of data and outputs on data repositories must comply with the respective legal framework (Intellectual Property Rights, licenses, informed consent, data ownership, etc.). Most of the scientific journals set an embargo period after which you are entitled to publish a specified version of your article in a FAIR repository. Check the journal's policy on <http://www.sherpa.ac.uk/romeo>.

Please find below a list of four repositories which accept datasets from different fields and fulfil the funder's requirements:

Repository	License	Metadata	Other
Zenodo	Default is CC-BY, but user can choose different license or restrict the access and set a condition.	Always publicly available (also in case of restricted access).	Allows to easily publish or cite a GitHub repository (link).
Harvard Dataverse	Default is CC0 waiver, but custom terms of use can be specified.	Always publicly available (also in case of restricted access).	
Eudat	Licenses for datasets are chosen by submitters and defined via an online form.	Always publicly available (also in case of restricted access).	
Dryad	All data submitted to Dryad are released to the public domain under CC0.	Always publicly available.	Dryad only accepts human subject data that is properly anonymized and prepared under applicable legal and ethical guidelines.

Other FAIR repositories are listed on www.re3data.org.

2. DMP for SNSF projects

2.1. Preparation of the DMP

The SNSF Data Management Plan (DMP) must be filled out in the grant application form directly on mySNF; the proposal can only be submitted once the DMP has been completed. The submitted DMP is considered a notice of intention, its content is assessed by the SNSF Administrative Offices for its plausibility and adherence to the SNSF policy on open research data, but it is not part of the scientific evaluation process.

If there are shortcomings in the submitted information, applicants will have to complete/amend specific sections of the DMP at the time of the funding decision and at the latest by the request for release of funds.

The DMP remains editable during the entire lifetime of the grant. In any case an updated version will be assessed together with the final scientific report and thereafter made publicly available on the SNSF's P3 database.

2.2. Eligible costs under SNSF grants

The costs of enabling access to research data that is collected, observed or generated under an SNSF grant should be included in the project budget. Such costs must be related to the preparation of research data in view of its archiving or to the archiving itself in a FAIR data repository.

The maximum charge per grant is CHF 10'000.

The SNSF covers also the costs of scientific open access publications, as specified below.

- **Article Processing Charges (APC):** the SNSF covers the publication charges (APC) of Gold open access publications if:
 - the article is published in a high-level peer-reviewed open access journal (gold open access);
 - the article is related to SNSF-funded research;
 - at least one of the authors is either applicant or employee of the related SNSF-funded project;
 - the article refers to the SNSF funding it resulted from (acknowledgement).
- **Book Chapters Processing Charges (BCPC):** the SNSF covers the publication charges (BCPC) of Book chapters published in an open access anthology if:
 - the chapter is published in a peer-reviewed open access anthology;
 - the chapter is related to SNSF-funded research;
 - at least one of the authors is either applicant or employee of the related SNSF-funded project;
 - the article refers to the SNSF funding it resulted from (acknowledgement).
- **Open access Books:** the SNSF finances the publication of scientific books that are freely accessible without limitations or delays (gold open access).

Publication costs must be requested via mySNF through a dedicated application form. Please contact janice.casarella@usi.ch for tailored support.

2.3. Structure of the DMP

In your grant application form on mySNF you will find the following questions in the DMP section. Short but comprehensive answers for each of them are required.

2.3.1. Data collection and documentation

1.1. What data will you collect, observe, generate or reuse?

Briefly describe the data you will collect, observe or generate. Also mention any existing data that will be (re)used. The descriptions should include the type, format and content of each dataset. Furthermore, provide an estimation of the volume of the generated data sets.

*What type, format and volume of data will you collect, observe, generate or reuse?
Which existing data (yours or third-party) will you reuse?*

See §1.2.

1.2. How will the data be collected, observed or generated?

Explain how the data will be collected, observed or generated. Describe how you plan to control and document the consistency and quality of the collected data: calibration processes, repeated measurements, data recording standards, usage of controlled vocabularies, data entry validation, data peer review, etc. Discuss how the data management will be handled during the project, mentioning for example naming conventions, version control and folder structures.

*What standards, methodologies or quality assurance processes will you use?
How will you organize your files and handle versioning?*

1.3. What documentation and metadata will you provide with the data?

Describe all types of documentation (README files, metadata, etc.) you will provide to help secondary users to understand and reuse your data. Metadata should at least include basic details allowing other users (computer or human) to find the data. This includes at least a name and a persistent identifier for each file, the name of the person who collected or contributed to the data, the date of collection and the conditions to access the data. Furthermore, the documentation may include details on the methodology used, information about the performed processing and analytical steps, variable definitions, references to vocabularies used, as well as units of measurement. Wherever possible, the documentation should follow existing community standards and guidelines. Explain how you will prepare and share this information.

*What information is required for computer or humans to read and interpret the data in the future?
How will you generate this documentation?
What community standards (if any) will be used to annotate the (meta)data?*

See §1.3.

2.3.2. Ethics, legal and security issues

2.1 How will ethical issues be addressed and handled?

Ethical issues in research projects demand for an adaptation of research data management practices, e.g. how data is stored, who can access/reuse the data and how long the data is stored. Methods to manage ethical concerns may include: anonymization of data; gain approval by ethics committees; formal consent

agreements. You should outline that all ethical issues in your project have been identified, including the corresponding measures in data management.

What is the relevant protection standard for your data? Are you bound by a confidentiality agreement?

Do you have the necessary permission to obtain, process, preserve and share the data? Have the people whose data you are using been informed or did they give their consent?

What methods will you use to ensure the protection of personal or other sensitive data?

See §1.5.1.

2.2. How will data access and security be managed?

If you work with personal or other sensitive data you should outline the security measures in order to protect the data. Please list formal standards which will be adopted in your study. Furthermore, describe the main processes or facilities for storage and processing of personal or other sensitive data.

What are the main concerns regarding data security, what are the levels of risk and what measures are in place to handle security risks?

How will you regulate data access rights/permissions to ensure the security of the data?

How will personal or other sensitive data be handled to ensure safe data storage and -transfer?

STANDARD: "Project data will be stored on file servers managed by USI IT service in folders with limited and managed access permissions. The principal investigator (applicant) will be in charge of deciding which researchers have access to the folders containing the data."

2.3. How will you handle copyright and Intellectual Property Rights issues?

Outline the owners of the copyright and Intellectual Property Right (IPR) of all data that will be collected and generated, including the licence(s). For consortia, an IPR ownership agreement might be necessary. You should comply with relevant funder, institutional, departmental or group policies on copyright or IPR. Furthermore, clarify what permissions are required should third-party data be reused.

Who will be the owner of the data?

STANDARD: USI is the owner of the data.

Which licenses will be applied to the data?

What restrictions apply to the reuse of third-party data?

See §1.5.2.

2.3.3. Data storage and preservation

3.1. How will your data be stored and backed-up during the research?

Please mention what the needs are in terms of data storage and where the data will be stored. Please consider that data storage on laptops or hard drives for example, is

risky. Storage through IT teams is safer. If external services are asked for, it is important that this does not conflict with the policy of each entity involved in the project, especially concerning the issue of sensitive data. Please specify your back-up procedure.

*What are your storage capacity and where will the data be stored?
What are the back-up procedures?*

STANDARD: "The project data will be stored and processed on file servers managed by the IT service of USI, which are protected by backup copies of data (backup) made every night of the working days."

3.2. **What is your data preservation plan?**

Please specify which data will be retained, shared and archived after the completion of the project and the corresponding data selection procedure (e.g. long-term value, potential value for reuse, obligations to destroy some data, etc.). Please outline a long-term preservation plan for the datasets beyond the life-time of the project. In particular, comment on the choice of file formats and the use of community standards.

*What procedures would be used to select data to be preserved?
What file formats will be used for preservation?*

See §1.4.

2.3.4. **Data sharing and reuse**

4.1. **How and where will the data be shared?**

Consider how and on which repository the data will be made available. The methods applied to data sharing will depend on several factors such as the type, size, complexity and sensitivity of data. Please also consider how the reuse of your data will be valued and acknowledged by other researchers.

*On which repository do you plan to share your data?
How will potential users find out about your data?*

See §1.6.

4.2. **Are there any necessary limitations to protect sensitive data?**

Data have to be shared as soon as possible, but at the latest at the time of publication of the respective scientific output. Restrictions may be only due to legal, ethical, copyright, confidentiality or other clauses. Consider whether a non-disclosure agreement would give sufficient protection for confidential data.

Under which conditions will the data be made available (timing of data release, reason for delay if applicable)?

See §1.5.1.

4.3. **All digital repositories I will choose are conform to the FAIR Data Principles**

The SNSF requires that repositories are conform to the FAIR Data Principles. If there are no repositories complying with these requirements in your research field, please deposit a copy of your data on a generic platform.

Yes / No

4.4. **I will choose digital repositories maintained by a non-profit organization**

Yes / No

2.4. Examples

2.4.1. DMP – ARC

1 Data collection and documentation

1.1 What data will you collect, observe, generate or reuse?

Based on historical documents and images, this qualitative research will produce georeferenced data sheets of a group of disappeared historical buildings. According to the documents, data will include information about the construction site, the building, the architects involved and the collections inside the buildings. Bibliographic research will allow to compare and include data produced by previous research. The data produced from this research will be:

1. bibliographical files;
2. transcriptions of archival documents;
3. historical research records;
4. images, photographs (about 1200);
5. images catalogues;
6. images lists.

Data in categories 1, 2, 3, 5, 6 will be documented in Word, and also then Pdf, file format. Data in category 4 will be documented in Tiff, and also then Pdf, file format. The data produced in category 4 will amount to approximately to 80 GB.

1.2 How will the data be collected, observed or generated?

Following the approach of history of architecture, data sheets will provide descriptions and data organized according to standard systems such as Bildindex (<https://www.bildindex.de/>). Documents will be organised using the university internal system of file sharing; a selection of documents and images will be made available to a broader public on the project website. The images will be purchased as already digitized by their rights holders.

1.3 What documentation and metadata will you provide with the data?

We will provide all data sheets, bibliographical files, transcriptions of archival documents and images catalogue and list. To allow a broader audience to access the data generated by the project, the research will contribute to Wikidata, the open, free knowledge and structured repository feeding also the Wikimedia projects. Wikidata allows to share content with metadata and it has standard description systems for buildings. Further content will be shared on the project website to contextualise the research, to provide a selection of reproductions of documents and significant images.

2 Ethics, legal and security issues

2.1 How will ethical issues be addressed and handled?

This historical research in the field of architecture does not present challenges related to ethical issues.

2.2 How will data access and security be managed?

This historical research in the field of architecture does not present challenges related to personal or other sensitive data. Data produced by the project can be released under CC0.

2.3 How will you handle copyright and Intellectual Property Rights issues?

Data produced by the project is released under CC0 and on Wikidata. Content produced by the project is in CC BY-SA license and articles are published under gold or green open access. There might be restrictions in publishing historical documents and images due to the involvement of third parties; the discussion about how to release sources will be addressed during the project and it is possible that documents might be available only for publication on the project's website.

3 Data storage and preservation

3.1 How will your data be stored and backed-up during the research?

The project data will be stored and processed on file servers managed by the IT service of USI, which are protected by backup copies of data (backup) made every night.

3.2 What is your data preservation plan?

The data preservation plan of the research is based on sharing data and content on existing platforms. The project does not create a new database but it uses existing systems; this approach increases the potentiality of reuse of the project's data and it relies on multiple open repositories with extensive communities. Content related to the project will be also shared on an internal university website which is maintained directly by the university as a subdomain.

To guarantee the durability in time, where possible, we will store files also in open archival formats, e.g. Word and Tiff files converted to PDF.

4 Data sharing and reuse

4.1 How and where will the data be shared?

Due to its humanistic nature, data of this research will be shared on repositories specialised in art history or addressing a broader audience. Wikidata – even if it is a peculiar repository and not a typically academic one – complies with the FAIR Data Principles and it is non-commercial. Eventually data can also be shared on Zenodo. The research content will also be published on an internal website managed by the university with texts under CC BY-SA.

4.2 Are there any necessary limitations to protect sensitive data?

This historical research in the field of architecture does not present challenges related to sensitive data. Images under copyright will be available only by agreement with the rights holders. The discussion about how to release images will be addressed during the project and it is possible that these documents might be available only for publication on the project website.

4.3 All digital repositories I will choose are conform to the FAIR Data Principles.

Yes

4.4 I will choose digital repositories maintained by a non-profit organisation.

Yes

2.4.2. DMP – COM

1 Data collection and documentation

1.1 What data will you collect, observe, generate or reuse?

1. Data that will be reused:

- In agreement with CERN, text and visual sources inside their archive (period 1989-1993) will be digitized (PDF format). We will need at least 10GB of storage space.

- Sources retrieved in newspaper and private archives. The materials will be digitized, when available only in paper format. These archives are in part public and in part privately owned. Access will be required for private or semi-public archives. We plan to digitize around 2GB of data.

- Primary-sourced histories from the founders available in specialized books and websites. It will be both text files and audio files. The text files will be digitized, when available only in paper format, or downloaded, if already available in digital format. The audio files will be downloaded when possible. Audio and digitized texts will be stored in the project folder on USI servers. All these materials are publicly accessible.

Audio: 50MB per each file (around 350MB in total).

2. Data that will be generated during the project:

- Digitally recorded interviews with the Web pioneers about the birth and early years of the Web: duration of each interview around 1 hour; 25-30 interviews to be saved in .Wav format. Around 200MB per file (6GB in total)

- Transcriptions of interviews: 25-30 PDF.

1.2 How will the data be collected, observed or generated?

Materials collected in the different archives will be digitized through scans of text and images (output format: PDF. Files) The files will be named as follow: ArchiveName_FolderName_Year_Month_Day_FileNumber.

The files will be saved in folders pointing to the reference year within the original archive. For each archive, information regarding the files will be collected in an Excel file containing general information (Year, File name, Type of document, details of content, comments).

The interviews will be downloaded and transcribed. Both files will be named as follows: NameOfInterviewee_Inviewer_Date_FileNumber.

An excel list of the interviews containing general information (Name of interviewee, Date, Description Comments) will be provided.

Generated data will be transferred as soon as possible to the USI servers to avoid data loss.

1.3 What documentation and metadata will you provide with the data?

For each archive, information regarding the files will be collected in an Excel file containing general information (Year, File name, Type of document, details of content, comments).

An excel list of the interviews containing general information (Name of interviewee, Date, Description Comments) will be provided.

Content analysis will produce a catalogue of PDF files pertaining each document collected within the archives and containing technical data (metadata about the document) and the results of the thematic analysis.

Data from the content analysis will be later on processed through SPSS and an excel sheet with the main results will be provided as well.

2 Ethics, legal and security issues

2.1 How will ethical issues be addressed and handled?

We have a special permit to access, process and preserve classified materials at the CERN archive. Collected data won't be made public without CERN agreement. Data collected remain property of CERN.

Permission to access semi-public newspaper archives and private archives will be asked in due course. Data collected within these archives won't be made public without the agreement of the archives owners and managers. Data remain property of their respective owners.

Informed consent will be signed by the interviewee. Interviews will be made available for research purpose. Investigated themes are not sensitive, therefore we do not provide anonymization of data. Anonymization will be provided if explicitly required by the interviewee.

Interviews included in books or websites are already publicly available data and will be treated as such.

2.2 How will data access and security be managed?

Data collected within the different archives and digitized will be accessible only to the research team and will be stored on USI servers.

There are no specific security needs concerning the collected data. Project data will be stored on file servers managed by USI IT service in folders with limited and managed access permissions. The project manager will be in charge of deciding which researchers have access to the folders containing the data.

2.3 How will you handle copyright and Intellectual Property Rights issues?

Data collected within the different archives remain property of their respective owners. They will not be shared without the owners' permission. Data generated by the project (interviews and content analysis results) are property of the Università della Svizzera italiana.

3 Data storage and preservation

3.1 How will your data be stored and backed-up during the research?

The project data will be stored and processed on file servers managed by the IT service of USI, which are protected by backup copies of data (backup) made every night. Cloud services are not intended to be used. Researchers will be trained to transfer data from the recording devices to the file servers as soon as possible to avoid risks of data loss.

3.2 What is your data preservation plan?

Data collected within the CERN archives and digitized by the research team will be preserved until the CERN archive will be opened to the public. The excel list of the materials available within the archives will be preserved for future research.

Data collected within the semi-public or private archives and digitized by the research team will be preserved until these archives will be opened to the public. The excel list of the materials available within these archives will be preserved for future research.

PDF files produced by the content analysis will be preserved until the final results of the research will be published. The excel list with the summary results of the content analysis will be preserved for future research.

Interviews digitally recorded in .wav format will be preserved together with PDF transcriptions.

10 years after the end of the project data will be re-evaluated in order to determine what to preserve.

4 Data sharing and reuse

4.1 How and where will the data be shared?

Subject to agreement with the researched archives, the list of the documents collected will be made available on Zenodo for future research purpose. Excel summary of the results produced by the content analysis will be made available on Zenodo for future research purpose. Detailed PDF files and SPSS documents produced by the content analysis will be made available upon request and only for research purposes. A written agreement will be signed in order to ensure the proper usage of the data made available to other interested researchers.

Zenodo assigns all publicly available uploads a Digital Object Identifier (DOI) to make the upload easily and uniquely citable. We will assign relevant titles and keywords to the datasets, so that researchers who are interested can easily find them.

4.2 Are there any necessary limitations to protect sensitive data?

Data collected within the different archives remain property of their respective owners. They will not be shared without the owners permission.

Data generated by the project (interviews and content analysis results) are property of the Università della Svizzera italiana and will be made available for research purposes only.

4.3 All digital repositories I will choose are conform to the FAIR Data Principles.

Yes

4.4 I will choose digital repositories maintained by a non-profit organisation.

Yes

2.4.3. DMP – COM

1 Data collection and documentation

1.1 What data will you collect, observe, generate or reuse?

In order to allow for a comparative analysis as foreseen in this project, we will have four different, interrelated and (probably) multilingual datasets: (a) 40-50 interviews to women in the Swiss academic system and their partners, corpus I in the research plan, (b) 40-50 individual interviews to women in the Swiss academic system, corpus II in the research plan; (c) 15-20 interviews to key institutional actors in the field of equal opportunities and (d) written public institutional documents about equal opportunities; (c) and (d) together make up corpus III (institutional discourse) in the research plan.

Datasets a, b, c (interviews) will be collected during the project. Based on previous similar experience, we anticipate that each interview will last about 45-60 minutes, which amounts to a total recording time (sum of datasets a, b, c) between 71.25 and 120 hours of recording. This requires between around 43GB and 75GB of storage for MP3 audio files. The data collection of datasets a, b will be preceded by a short pilot study that will enable us to refine the interview questionnaire and also to make more precise calculations about the length of interviews and the required data storage.

In a first phase, interviews will be first transcribed in a non-anonymized form, only available to the researchers as Word files (and deleted at the end of the project). The interviews will then be anonymized. Anonymized interviews will be stored in two versions. First, Word/PDF versions will be created for the purposes of the project (publications, research) and for future sharing of data (see also 4.1). Second, XML versions will be created and annotated using the freely available software UAM Corpus Tool (CT, <http://corpustool.com>, see O'Donnell 2008), version 3.3. CT has already proven useful for the purposes of an argumentative analysis and has been already used by applicant and post-doc in previous projects. CT requires transcriptions to be in XML files.

At the moment, around 80 public institutional documents in English, Italian, French and German (as PDF files) have been collected and stored on USI server, in a folder dedicated to this project. Presently, the public institutional documents require about 58MB of storage; the collection of documents, however, will be regularly updated during the project and we estimate an increase of at least 15-20 documents over the four years.

1.2 How will the data be collected, observed or generated?

Audio recordings will be made using digital recorders that are made available for researchers at USI by IT services, which allow to do recordings of the quality required for this project. Interviews will be recorded, respectively, by the two PhD students (datasets a, b) and by the post-doc researcher (dataset c). The post-doc researcher also has the task to work on the collection of public institutional documents (dataset d). During the pilot study, and during the process of data collection, the applicant and the post-doc will work with the PhD students to refine the methodology and give feedback on possible difficulties emerging during the data collection; during regular weekly project meetings, any urgent issue can be addressed by the project team.

Files of the datasets a, b, c, d will be stored separately on three USI IT folders (see 2.2).

New versions of interviews/documents will be clearly labelled; the post-doc will have the task of managing the process of updating versions.

Transcriptions of interviews will be helped by two student assistants. This will allow double review of data, which enable us to enhance the quality of transcriptions.

1.3 What documentation and metadata will you provide with the data?

For datasets of interviews (a, b, c) we will create metadata to allow the project team and other users to understand and reuse the files. We will create an identifying name (dataset letter + number of interview) for each file and corresponding transcription. For datasets a, b, in a separate Excel file, we will insert metadata: time/place of interview, language(s), name of interviewer, plus basic sociological data on interviewees, such as age, education, number of children, nationality, time spent in the Swiss academic system, level of academic career,

year of completion of the PhD. In case of a couple's interview, sociological data will also include: family status of interviewee, relationship to the interviewed person (e.g. partner/friend/relative).

For dataset c, in a separate Excel file, we will insert metadata: time/place of interview, language(s), name of interviewer, plus basic sociological data on interviewees, such as age, education, number of children, role within equal opportunities in Switzerland (e.g. head of service/administrator/other). At the moment, we have labelled and sorted the documents with a first classification but the metadata will be refined during the project.

We use Excel files for our metadata because, although the software UAM Corpus tool permits to insert meta-data, it is likely that we will start the annotation after all interviews have been transcribed; we want to compile the Excel files while progressively storing and transcribing the interviews.

Also concerning datasets a, b, c (interviews) in a separate Word file called JOURNAL, the PhD students and post-doc will write up a short comment (in the form of a journal) immediately after each interview, noting possible ambiguities and everything that was not clear to them during the interview. From previous experience, we know that this additional information will be useful when analysing data. This file will remain available only to the project team because it will include confidential information.

PhD students will be trained about metadata from the applicant and post-doc and they will be continuously monitored during the project. Two student assistants will help with the transcriptions but providing precise metadata is the PhDs' responsibility, under the general monitoring of post-doc and applicant. Note that the metadata described above are for strict use of the project team. A public version of the Excel file of metadata, in which only non-sensitive information will be included, will be created to make transcriptions available to other researchers (see section 2.1 below).

For dataset d (public institutional documents), we will create metadata to allow other users to reuse the files. We will create an identifying name (dataset letter + number of file) for each file. In a separate Excel file, we will insert metadata: name of institution that issued the document, website from which the document was downloaded, time when the document was downloaded from the website language, available translations.

2 Ethics, legal and security issues

2.1 How will ethical issues be addressed and handled?

Institutional documents (dataset d) have been downloaded from public websites of Swiss higher education institutions.

For individual interviews (datasets a, b, c), including the pilot studies, all participants will sign an informed consent, compliant with GDPR. They will be informed that only transcriptions of the data in an anonymous form will be published and possibly shared with other researchers, only for scientific reasons.

The applicant will get support from the USI ethical committee to define the legal and ethical issues for this informed consent, as she did with previous projects. Given the focus of this project, all participants will be adults and therefore will have the right to sign their own informed consent. The participants will be informed that they can withdraw their permission to use the data. They will be given the applicant's and researcher's email addresses, and know that they can always contact the project team, if they want to have further information.

As said in 1.1., we will produce as soon as possible an anonymized version of the transcription, removing names and all identifying information (Word/PDF/XML). For this anonymized version, only a selection of metadata (language, age of the interviewee, number of children) will be made public. Excerpts of interviews can be published only for scientific reasons. Audio files and other metadata will only be available to the project team in order to assure the anonymity of the participants. The "JOURNAL" Word file will remain available only to the project team because it will include confidential information.

Confidentiality is important because the interviews, especially to women academics, may include sensible topics and, in order for us to access the data we need, we need participants to feel they can express themselves freely.

Other possible legal and ethical issues emerging during the project will be discussed by the project team under the supervision of the applicant and, if needed, with USI ethical committee.

2.2 How will data access and security be managed?

Project data will be stored on file servers managed by USI IT service in folders with limited and managed access permissions. The principal investigator (applicant) will be in charge of deciding which researchers have access to the folders containing the data (see also 2.3) during the project (project team) and after, especially if some researchers will change their institutions.

2.3 How will you handle copyright and Intellectual Property Rights issues?

USI - Università della Svizzera italiana will be the owner of the data. Audio files and sensible metadata will only be available to the project team in order to assure the anonymity of the participants. Only anonymized versions of the transcriptions (PDF) will be made public to other researchers (limited by a confidentiality agreement) 5 years after the end of the project.

3 Data storage and preservation

3.1 How will your data be stored and backed-up during the research?

The project data will be stored and processed on file servers managed by the IT service of USI, which are protected by backup copies of data (backup) made every night. We do not intend to use cloud. Researchers will be trained to transfer data from the recording devices to the file servers as soon as possible to avoid risks of data loss.

3.2 What is your data preservation plan?

At the end of the project, the non-anonymized transcriptions of interviews will be immediately deleted. All other data will be kept in the USI folders, for further research, maintaining the standard of confidentiality used throughout the project. We will keep the audio files because they could be useful for further publications of the project team. After 7 years, the applicant will have the task of reconsidering what to do with the audio files – whether to keep them for another 5 years or longer, or to delete them, depending on data value.

The annotation made in UAM Corpus Tool will also be a profitable basis for future research on argumentation; therefore, we will keep the XML files. PDF files of anonymized interviews (not annotated) will be shared more broadly (see 4.1) five years after the end of the project, while annotated anonymized versions will be shared seven years after the end of the project (see 4.1).

4 Data sharing and reuse

4.1 How and where will the data be shared?

As soon as they are ready, anonymized metadata will be made publicly available on the repository Zenodo (<https://zenodo.org>), which is a non-commercial repository. Our datasets will likely be multilingual so we will specify that on Zenodo.

The audio data stored at USI will only remain available to the applicant and members of this project; while copies of annotated interviews (XML files annotated with UAM Corpus Tool) could be made available during the project, if needed for research purposes, to other researchers who sign a confidentiality agreement – it is the applicant's responsibility to make this decision. Five years after the end of the project, anonymized transcriptions of the interviews in PDF format will be made available on Zenodo only for research purposes in the field of argumentation and discourse analysis, to researchers who sign a confidentiality agreement. Seven years after the end of the project, annotated interviews (also anonymized) will also be made available on Zenodo only for research purposes in the field of argumentation and discourse analysis, to researchers who sign a confidentiality agreement. Dataset (d) is composed of written public documents taken from websites of universities. While its metadata will be shared, the documents per se will not be shared on Zenodo.

Zenodo assigns all publicly available uploads a Digital Object Identifier (DOI) to make the upload easily and uniquely citable. We will assign relevant titles and keywords to the datasets, so that researchers who are interested can easily find them.

4.2 Are there any necessary limitations to protect sensitive data?

Initially, only anonymized metadata will be made publicly available on Zenodo. Anonymized transcriptions of the interviews in PDF format (not annotated) will be made available only for research purposes in the field of argumentation and discourse analysis five years after the end of the project. Seven years after the end of the projects, the anonymized annotations will be made available only for research purposes on Zenodo. In order to further protect sensitive data, we will anyway ask to sign an agreement to those researchers who are interested to use our anonymized datasets making sure that they will only use the data for research purposes.

We will be very strict when interpreting what it means to produce "anonymized" transcriptions, checking that any information that could make the interviewee identifiable will not be shared. The applicant will monitor this aspect throughout the process, asking relevant advice from the Ethical committee at USI when necessary.

4.3 All digital repositories I will choose are conform to the FAIR Data Principles.

Yes

4.4 I will choose digital repositories maintained by a non-profit organisation.

Yes

2.4.4. DMP – ECO

1 Data collection and documentation

1.1 What data will you collect, observe, generate or reuse?

The project will combine the following datasets:

- Equity returns and other stock-market information of publicly listed companies
- Balance-sheet information of these companies
- Macroeconomic indicators

Part of the data is obtained from data providers (Datastream, Orbis) for which USI has an existing licence. Other data (such as, some macro series) are from central banks and other organizations (such as the FED or World Bank). Data is received in CSV or Excel format. We anticipate that the overall dataset will amount to approximately 20 GB.

1.2 How will the data be collected, observed or generated?

The equity data will be accessed and downloaded from Datastream and Obis. Other data will be downloaded from the website of economic organizations (such as the FED or World Bank). To keep track of possible revisions/versions of the data, they be stored into a folder that is named based on the date of download. To guarantee the replicability of our analysis, new downloads are stored in separate folders, and previous downloads are preserved.

1.3 What documentation and metadata will you provide with the data?

The data will be accompanied with a precise description of the series that are used in the study, along with their reference code when present (e.g., Mnemonic in Datastream). The raw data will then be processed (through some dedicated software such as Stata, Matlab, or Phyton) to be converted into a format suitable for the empirical analysis. The procedure used to process the data will be described in a technical report or in a data section of the publication, including the type of software used, and made public together with the file and an indication of the software version that was used.

2 Ethics, legal and security issues

2.1 How will ethical issues be addressed and handled?

There are no ethical issues in the generation of results from this project.

2.2 How will data access and security be managed?

Data will be stored on a dedicated disk space of one of the USI servers that are managed by our IT department. The data will be accessible using virtual desktop technology only by authorized participants to the project. The list of authorized participants will be managed by the applicant. Access to the database will be logged, thus each access is traceable.

2.3 How will you handle copyright and Intellectual Property Rights issues?

Part of the data is obtained from data providers (Datastream, Orbis), and is subject to a non-disclosure agreement. Other data (such as, some macro series) are from central banks and other organizations (such as the World Bank) and are freely available. The research is not expected to lead to patents.

3 Data storage and preservation

3.1 How will your data be stored and backed-up during the research?

Data will be stored on a dedicated disk space of one of the USI servers that are managed by our IT department. Our servers have redundancy, mirroring and are monitored. The servers are backed up on a regular basis (at least once per week).

3.2 What is your data preservation plan?

We will preserve the data for at least 10 years on the university's server, and also deposit it in an appropriate data archive (such as Zenodo). Where possible, we will store files in open archival formats, such as Excel files converted to CSV. In case this is not possible, due to

the large volume of the data, we will include information on the software used and its version number.

4 Data sharing and reuse

4.1 How and where will the data be shared?

Datasets from this work which underpin a publication will be made available through the project participants' websites or common published on Zenodo, and made public at the time of publication. Data in the repository will be stored in accordance with funder's data policies. The retention schedule for data will be set to 10 years from date of deposition in the first instance, with possible extension for datasets which remain in regular use.

4.2 Are there any necessary limitations to protect sensitive data?

Datasets from this work which underpin a publication will be made available at the time of publication. Data that is purchased by third-party data providers will be published in a form that does not constitute a violation of the license agreement.

4.3 All digital repositories I will choose are conform to the FAIR Data Principles.

Yes

4.4 I will choose digital repositories maintained by a non-profit organisation.

Yes

2.4.5. DMP – INF

1 Data collection and documentation

1.1 What data will you collect, observe, generate or reuse?

Data will be collected from static and dynamic program analyses applied to well-known benchmarks and to code in public software repositories. Depending on the concrete metrics to be collected (which will be investigated in WP 1 and WP 2) and the number of open-source projects where our analyses are applicable, the total amount of data collected may reach the order of terabytes. We will maintain proper scripts to be able to reproduce the results of our evaluations.

1.2 How will the data be collected, observed or generated?

We will use scripts to fully automate the collection and aggregation of measurements, to compute statistics, and to generate figures conveying the results of our experimental evaluations. We will pay special attention to the reproducibility of our results. To ensure that experiments are repeatable on specific workloads, we will save detailed provenance data, such that one can analyse a given workload version even if it is superseded by a more recent version. The maintenance of provenance data is particularly important for large-scale evaluations on open-source projects in software repositories such as GitHub. However, please note that full reproducibility of experiments is only possible for deterministic workloads.

For metrics that are prone to fluctuations, we will repeat measurements a sufficient number of times and report mean values as well as different quantiles, variance measures, and confidence intervals. In some cases, measurements need to be taken after a warmup phase of a system (e.g., when a system has reached a steady state after initial dynamic compilation); the details of the measurement procedures will be maintained in scripts to ensure consistency of measurements and reproducibility of results. Measurement results will be kept in private repositories shared by the project team. We will use free and open-source distributed version control systems such as Git.

1.3 What documentation and metadata will you provide with the data?

For each data set, we will store metadata about the input data (e.g., concrete revisions of open-source software) and on the toolchain (e.g., version numbers of used analysis tools) used to obtain the data. The details of the measurement process will be provided via scripts that allow other researchers or practitioners to repeat the measurements and reproduce the data. The scripts will be properly documented to ease their reuse.

2 Ethics, legal and security issues

2.1 How will ethical issues be addressed and handled?

Our planned research does not raise any ethical concerns.

For the evaluation of our methods and techniques, we will be working on open-source projects maintained in public code repositories. The code and data contained there are freely available and so our results do not need any particular protection, as we are not dealing with any sensitive data.

2.2 How will data access and security be managed?

As explained before, we will not be dealing with sensitive data. Hence, the results of our work do not require any special protection.

2.3 How will you handle copyright and Intellectual Property Rights issues?

The software developed in this project will be released open-source under the most appropriate license to benefit the community. Thus, the researchers funded by this project can freely exploit the results of their research in their future careers.

3 Data storage and preservation

3.1 How will your data be stored and backed-up during the research?

The project team will use backed-up, private repositories (such as those provided by the university) for sharing documents, code, data, and any other artefacts during the project. Completed and tested software will be released open-source in public repositories. Data collected in experiments will be kept in a distributed database. Since we are not dealing with sensitive data and the results from experimental evaluations can be easily recomputed, special precautions regarding data storage are not needed.

3.2 What is your data preservation plan?

As discussed before, all relevant artefacts will be kept in backed-up repositories, ensuring their long-term availability. Completed software will be released open-source and maintained in public code repositories. We also plan to submit artefacts for accepted publications, which will be made available through the publisher's digital library.

4 Data sharing and reuse

4.1 How and where will the data be shared?

During the research activities, all data and artefacts will be shared between the project team (and possibly with external collaboration partners) in backed-up, private repositories (e.g., provided by the university's IT department). Completed and tested software will be released open-source and kept in public code repositories to ensure the widest possible impact and reuse. Experimental data underlying publications will be submitted as artefacts, to be included in the publisher's digital library. We will also include scripts to reproduce all measurements and the figures included in our papers.

4.2 Are there any necessary limitations to protect sensitive data?

Data underlying a publication will be publicly released as an artefact accompanying the publication. Whenever supported by the publisher, we will submit artefacts with our papers for inclusion in the publisher's digital library. Scripts used to reproduce measurements will be part of the open-source releases of our software, to be maintained in public code repositories.

4.3 All digital repositories I will choose are conform to the FAIR Data Principles.

Yes

4.4 I will choose digital repositories maintained by a non-profit organisation.

Yes

3. DMP for H2020 projects

3.1. Preparation of the DMP

A first draft of the DMP must be submitted by month 6 of the project. It is not required to provide detailed answers to all the questions in the first version of the DMP. Rather, the DMP 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.

As a minimum, the DMP should be updated in the context of the periodic evaluation/assessment of the project. If there are no other periodic reviews envisaged within the grant agreement, an update needs to be made in time for the final review at the latest.

3.2. Structure of the DMP

The following template is a set of questions that you should answer with a level of detail appropriate to the project. You can download the template from the EC [participants portal](#).

1. Data Summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?

What types and formats of data will the project generate/collect?

Will you re-use any existing data and how?

What is the origin of the data?

What is the expected size of the data?

To whom might it be useful ('data utility')?

2. FAIR data

2.1. Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

What naming conventions do you follow?

Will search keywords be provided that optimize possibilities for re-use?

Do you provide clear version numbers?

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

2.2. Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

How will the data be made accessible (e.g. by deposition in a repository)?

What methods or software tools are needed to access the data?

Is documentation about the software needed to access the data included?

Is it possible to include the relevant software (e.g. in open source code)?

Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

Have you explored appropriate arrangements with the identified repository?

If there are restrictions on use, how will access be provided?

Is there a need for a data access committee?

Are there well described conditions for access (i.e. a machine readable license)?

How will the identity of the person accessing the data be ascertained?

2.3. Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

2.4. Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

How long is it intended that the data remains re-usable?

Are data quality assurance processes described?

3. Allocation of resources

What are the costs for making data FAIR in your project?

How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

Who will be responsible for data management in your project?

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

4. Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Is the data safely stored in certified repositories for long term preservation and curation?

5. Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

Contatti

Research and Transfer Service
Università della Svizzera italiana
Via Buffi 13
6900 Lugano
Switzerland

e-mail igor.sarman@usi.ch
web www.usi.ch

© Università della Svizzera italiana