UNIL’s generic model of Data Management Plan

Data Management Plan section table

[Instructions 1](#_Toc58489245)

[Projet’s information : 2](#_Toc58489246)

[Data description 3](#_Toc58489247)

[Methodology and data processing 4](#_Toc58489248)

[Organization and file naming 5](#_Toc58489249)

[Documentation and metadata 5](#_Toc58489250)

[Storage 7](#_Toc58489251)

[Backup 8](#_Toc58489252)

[Ethical and legal issues 8](#_Toc58489253)

[Data security 10](#_Toc58489254)

[Intellectual property 11](#_Toc58489255)

[Long-term preservation 12](#_Toc58489256)

[Data sharing and repository 13](#_Toc58489257)

December 2020

# Instructions

Answer only the questions in this Data Management Plan (DMP) that are relevant to your situation. Questions that you do not necessarily have to answer begin with "if" and depend on your previous answers.

If you have any questions or comments about this DMP template, you can write to us at: researchdata@unil.ch.

# Projet’s information :

**Project title**:

**Funder**:

**Project abstract**:

**Principal Investigator**

Last name, first name:

ORCID:

Email:

Phone:

Data contact person:

**UNIL’s generic model of Data Management Plan**

*To help you answer this Data Management Plan (DMP), guidance blocks are available on the right side of the questions. They contain definitions or resources that may be useful to you.*

# Data description

*In this section you must describe the data from your research project (if it uses or generates data). The description contains the source, type, volume, content, and format of those data.*

[This section relates to question 1.1 of the mySNF DMP].

**1. Data will be collected, studied, produced, or reused during your project:**

[ ] Yes

[ ] No

Guidance : According to the commonly accepted [OECD definition](https://www.oecd.org/sti/inno/38500813.pdf) (2007), research data are factual records used as primary sources for scientific research. They are generally recognized by the scientific community as necessary to validate research results. They can take many forms (experimental data, observational data, operational data, third-party data, public sector data, etc.). For more information, visit [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/les-donnees-de-recherche.html) of the UNIL Open Science website.

**2. If you will not collect, study, produce or reuse data during your project, explain why:**

The writing of your DMP stops here if you are not going to collect, study, generate or reuse research data for this project. For other situations, please continue:

**3. Does your project involve the reuse of existing data (produced by you or third parties)?**

[ ] Yes

[ ] No

Guidance : Existing data may, for example, refer to work on archives or historical sources, or to data produced in other research contexts, by you or third parties, that will be reused in your project.

**4. If your project involves the reuse of existing data, where are they coming from, what do they contain, and what are their format?**

Guidance : The data you reuse may come from a variety of sources, for example, directly from the author, from other public or private institutions, from other research projects, from archival or library collections, from online databases, etc. Here is an example of a possible answer to this question: within the framework of this project, various existing data will be reused, in particular phenological observation data provided by the Federal Office of Meteorology and Climatology MétéoSuisse in .csv format and available at <https://opendata.swiss/fr/dataset/phanologische-beobachtungen>.

**5. Indicate the type of data produced during the project (new data):**

Multiple choice answer:

[ ] Material and physical (non-digital) data

[ ] Texts

[ ] Spreadsheets

[ ] Databases

[ ] Audio

[ ] Video

[ ] Images

[ ] Geospatial

[ ] 3D, models, visualizations

[ ] Source codes, software, computer programs

[ ] Other (please specify):

**6. Briefly describe the content and the format of the data produced during your project (new data):**

Guidance : For example, the sound recordings of the interviews conducted will be in .mp3 format. The transcript of the interviews will be in a .docx file.

# Methodology and data processing

*In this section, the aim is to describe:*

*- The methodology related to the project and how the data are collected and/or produced.*

*- Data processing, i.e., what manipulations are performed on the data and how they are studied or analyzed to obtain results.*

[This section relates to question 1.2 of the mySNF DMP].

**7. Describe how the data will be collected and/or produced as part of your project:**

Guidance : For example, describe the research methodology and processes used, the methods of data collection, how the data will be produced, the instruments, tools or software used and the manipulations performed, the calibration, sampling, parameterization, etc.

**8. Describe how the data will be processed, studied, and analyzed in your project:**

Guidance : After the collection and/or the production of the data, the processing and analysis steps take place. Detail here the processing (automated or not) and analyses you will perform on the data once you have collected or produced them all. For example: data cleaning, transcription, organization, adaptation, quality control, dissemination, instruments used, manipulations carried out, standards implemented, etc.

# Organization and file naming

*In this section, the aim is to describe:*

*- How data are organized and managed throughout the project.*

*- The file naming adopted.*

[This section relates to question 1.2 of the mySNF DMP].

**9. Describe the organization of folders and data files adopted (classification tree):**

Guidance : Example of elements to be detailed: classification tree used (e.g. folder and file structure), management of different file versions, computer code revision management systems (c4science or Github), etc. To learn more about data organization, see [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/organisation--description.html).

**10. Describe the file naming (nomenclature) adopted:**

Guidance : You can consult the recommended naming rules available [on the Open Science website](https://www.unil.ch/openscience/files/live/sites/openscience/files/Donnees_de_recherche/Files/GC_Organiser_ses_donnees_Regles_Nommage_2019.pdf) (in French).

# Documentation and metadata

*This section details the documentation that accompanies the data so that it can be reused and reinterpreted in the future. This includes a description of the metadata used.*

[This section relates to question 1.3 of the mySNF DMP].

**11. Will you be writing documentation to accompany the data?**

[ ] Yes

[ ] No

**12. If you will not write documentation to accompany the data, explain why:**

**13. If you will write documentation to accompany the data, in what form do you plan to do it?**

Multiple choice answer:

[ ] A readme.txt file

[ ] A codebook

[ ] An Electronic Laboratory Notebook (ELN)

[ ] A field notebook

[ ] A text file (.docx, .odt, .pdf)

[ ] The documentation will be directly integrated into the data file.

[ ] Other (please specify):

Guidance  : To learn more about the organization and description of the data, visit [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/organisation--description.html).

**14. If you will write documentation to accompany the data, what information will it provide?**

Multiple choice answer:

[ ] The context of the project (researcher, institution, period in which the research is being carried out, etc.)

[ ] The research questions

[ ] The research hypotheses

[ ] The methodology adopted

[ ] How the data are collected

[ ] Data analysis procedures

[ ] The instruments used and the manipulations carried out

[ ] The software used (mentioning which version, as well as the parameters set up)

[ ] The data formats and types

[ ] The standards and variables used (names, questions, descriptions, algorithms, syntaxes, etc.).

[ ] The coding table

[ ] The units of measurement used

[ ] Data reuse license

[ ] Data access rights

[ ] Other (please specify):

Guidance : The documentation contains basic information to understand the research project and the context of data creation and collection. To learn more about the organization and description of the data, visit [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/organisation--description.html).

**15. Indicate what metadata or metadata standards will accompany the data:**

Multiple choice answer:

[ ] DataCite

[ ] Dublin Core

[ ] DDI

[ ] EAD

[ ] ISO 19115

[ ] RDF

[ ] Other (please specify):

Guidance : Metadata is information about the data. It indicates, for example, who created them, when, for what purpose, etc. Each set of data collected or created should be accompanied by comprehensive metadata that complies with the standards, rules, and conventions of a discipline. To learn more about metadata, visit [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/organisation--description.html).

# Storage

*In this section, the aim is to provide an estimate of the volume required to store the data during the project and to detail which storage infrastructures will be used.*

[This section relates to question 3.1 of the mySNF DMP].

**16. Provide an estimate of the volume needed to store the data:**

Single choice answer :

[ ] Less than 50GB

[ ] From 51GB to 500GB

[ ] From 501GB to 1023GB

[ ] 1TB (1024GB) and more

[ ] I don't know yet

Guidance : To learn more about units of measurement, [this page](https://www.canto.com/blog/digital-storage-units/) may be useful. To estimate the volume of your data, you can, for example, refer to the volume generated by previous projects or look at the current weight of your files on your computer.

**17. What storage infrastructure will you use?**

Multiple choice answer:

[ ] Storage infrastructure from the Scientific Computing and Research Support Unit (DCSR)

[ ] Other Unil servers

[ ] Unil Computer

[ ] Personal computer

[ ] External hard disk

[ ] USB key

[ ] Switchdrive

[ ] Other cloud services (please specify):

[ ] Other (please specify):

Guidance : Keeping research data only on your computer or on external storage devices (hard drive, USB stick) is not recommended. Storage is preferably done on UNIL's servers via the Scientific Computing and Research Support Unit ([DCSR](https://www.unil.ch/ci/fr/home/menuinst/calcul--soutien-recherche/scientific-computing-and-research-support-unit.html)). For more information on the storage of research data at UNIL (infrastructure, costs, etc.), you can consult [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/stockage--securite.html) or read the [Directive 4.5 : Traitement et gestion des données de recherche](https://www.unil.ch/central/files/live/sites/central/files/textes-leg/4-rech/dir4-5-donnees-rech2.pdf) (in French), or the two following articles of the CiNN: « [Des ressources de stockage et de calcul pour les chercheurs de l’UNIL](https://wp.unil.ch/cinn/2019/08/des-ressources-de-stockage-et-de-calcul-pour-les-chercheurs-de-unil/) » et « [Le FNS exige désormais un plan de gestion des données pour accorder ses fonds](https://wp.unil.ch/cinn/2017/09/le-fns-exige-desormais-un-plan-de-gestion-des-donnees-pour-accorder-ses-fonds/) » (in French as well).

# Backup

*This section concerns data backup, i.e. the creation of additional copies to compensate for any possible loss of data. It describes what mechanisms will be put in place to avoid this risk.*

[This section relates to question 3.1 of the mySNF DMP].

**18. Describe how the data will be backed up:**

Guidance : Backup consists of creating additional copies of the data. It is essential to avoid the risk of data loss through accidental erasure, hard drive failure, theft, or damage of equipment.

To answer this question, you can mention, for example, the use of Crash-Plan (or other backup automation software), external backup media, the frequency of your backups, the number of copies made, the measures to be taken in the event of an incident (loss of backups), etc. To learn more about backup options at UNIL, you can read the following article from the CiNN: « [Le FNS exige désormais un plan de gestion des données pour accorder ses fonds](https://wp.unil.ch/cinn/2017/09/le-fns-exige-desormais-un-plan-de-gestion-des-donnees-pour-accorder-ses-fonds/) » (in French).

# Ethical and legal issues

*The use of data in a research project sometimes raises ethical issues and/or is subject to legislation. The purpose of this section is to assess the ethical issues that your project may raise and to detail the measures put in place to ensure compliance with scientific ethics and legislation.*

[This section relates to question 2.1 of the mySNF DMP].

**19. Please indicate the ethical issues raised by your project:**

[ ] Research on human beings (even non-medical) or with human participants

[ ] Collection of biological data or material

[ ] Clinical Trial

[ ] Animal experiments

[ ] Outcomes that could affect the environment, public health, or safety

[ ] Potential for military use

[ ] Other (please specify):

[ ] My project does not raise any ethical issues

**20. Does your project involve the collection or processing of personal and/or sensitive data?**

[ ] Yes

[ ] No

Guidance : According to the [Personal Data Protection Act (LPrD)](https://prestations.vd.ch/pub/blv-publication/actes/consolide/172.65?key=1543934892528&id=cf9df545-13f7-4106-a95b-9b3ab8fa8b01) of the Canton of Vaud, personal data is "any information that relates to an identified or identifiable person". Personal data can be, for example, direct identifiers (name, date of birth, address, etc.) or indirect identifiers (likely to reveal a person's identity when aggregated). Still according to the LPrD, sensitive data is defined as "any personal data relating to:

- religious, philosophical, political, or union opinions or activities, as well as ethnic origin.

- to the intimate sphere of the person, in particular to his or her psychological, mental or physical state.

 - to individual measures and aid resulting from social legislation.

 - penal and administrative proceedings or sanctions".

Anonymized data are no longer personal and/or sensitive data and are no longer protected. Visit [this page](https://unil.ch/openscience/en/home/menuinst/open-research-data/conformite--exigences/donnees-personnelles--sensibles.html) to learn more.

**21. If your project raises ethical issues, indicate the specific measures you will put in place to address those issues:**

Multiple choice answer:

[ ] Inform participants fully about the project objectives

[ ] Inform the participants about the nature of the data collected, their treatment and the modalities of their communication and conservation

[ ] Explain to participants their rights

[ ] Gathering informed consent from participants

[ ] Collect data anonymously

[ ] Coding or pseudonymizing personal and/or sensitive data

[ ] Anonymize personal and/or sensitive data

[ ] Seek the approval of the competent authorities

[ ] Other (please specify):

Guidance : According to UNIL's Research Ethics Commission ([Commission d’éthique de la recherche](https://www.unil.ch/cerunil/home.html)), collecting data in an anonymous way implies not collecting any (directly or indirectly) identifying data. Data are considered anonymous when the link between the research subject and the data is irreversibly broken. Where a key exists, materials and data are considered coded or pseudonymized. Note that unless consent is obtained from participants, research data must be either destroyed or anonymized upon completion of the research.

**21. If your project raises ethical issues, has an ethics commission validated your project?**

Single choice answer

[ ] Yes

[ ] An application is in progress or will be filed

[ ] No

Guidance : If your research falls within the scope of the [Swiss Federal Act on Research involving Human Beings](https://www.admin.ch/opc/en/classified-compilation/20061313/index.html) (HRA), then you are obliged to submit it to the Cantonal Research Ethics Commission ([CER-VD](http://www.cer-vd.ch/)). The HRA concerns research on human diseases and the structure and functioning of the human body that is carried out with natural persons, deceased persons, embryos and foetuses, with biological material or with personal health-related data.

Submission to UNIL’s Research Ethics Commission (CER-UNIL) is for any researcher or student wishing to obtain an attestation of ethical compliance for a research project that does not fall within the scope of the HRA. He or she can then [submit an application electronically](https://cer-submission.unil.ch/en). The process is entirely voluntary. For more information, visit the [CER-UNIL website](https://www.unil.ch/cerunil/home.html).

**22. If an ethics commission validated your project, which one is concerned?**

[ ] CER-VD (Cantonal Research Ethics Commission)

[ ] CER-UNIL (UNIL’s Research Ethics Commission)

[ ] Other (please specify):

# Data security

*This section concerns the measures put in place to ensure the computer security of data throughout the project and in particular any measures taken to ensure the security of personal or sensitive data.*

[This section relates to question 2.2 of the mySNF DMP].

**24. Describe the measures put in place to ensure the security of data throughout the project:**

Guidance : For example, talk about your security policy, the standards adopted, the identification and assessment of risks and the measures put in place to manage them, the security of access, sharing, storage, backups, transfers, the physical security of the premises, etc. To learn more about data security, consult [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/stockage--securite.html) or the [UNIL's information security policy](https://www.unil.ch/ci/files/live/sites/ci/files/shared/securite-public/Politique-de-securite-de-l%27information-Ci-UNIL-v1.7.pdf) (in French).

**25. If your project involves the collection or processing of personal and/or sensitive data, describe the additional security measures put in place to ensure the protection of this type of data:**

Guidance : For example, additional security measures can be put in place by applying confidentiality clauses, access subject to authorization or limited, or anonymization or pseudonymization of data. Also think about taking special measures for the security of storage (beware of cloud storage), transfer or sharing of this data, as well as the conditions for its retention or destruction, etc. Also remember that security measures for personal data must be put in place from the first contact with participants, and not only once the data has been collected.

# Intellectual property

*Intellectual property issues are often difficult to comprehend. This section is intended to clarify these issues in the context of your project.*

[This section relates to question 2.3 of the mySNF DMP].

**26. If you have previously indicated that your project involves the reuse of existing data (produced by you or third parties), under what conditions is this reuse possible?**

Guidance : Examples of situations that permit the re-use of third party data: you own the copyright of the data; you have requested and obtained permission for re-use from the copyright owner(s); you have an agreement, contract or license with the data provider that permits you to re-use the data; a copyright license permits you to re-use the data (e.g. the Creative Commons); the data are in the public domain; an exception in the Copyright Act (CopA) allows you to do so ([right of quotation](https://ccdigitallaw.ch/index.php/english/copyright/5-how-may-other-people-use-work/55-right-quotation) or [restriction for scientific research](https://ccdigitallaw.ch/index.php/english/copyright/5-how-may-other-people-use-work/55-restriction-for-scientific-research) (text and data mining in particular)). It is to be noted that the restriction for scientific research for text and data mining does not exceed a possible contractual commitment between the data provider and you. It is therefore advisable to check that contract first.

Mention here the terms and conditions under which such reuse is made possible (e.g. those of the license or contract, the rights granted, the restrictions provided, the distribution terms permitted, etc.).

**27. Will the data created during your project be subject to any restrictions related to patents or inventions, or to a contract?**

[ ] Yes

[ ] No

Guidance : UNIL is the owner of the data you will collect or produce ([see Directive 4.5](https://www.unil.ch/central/files/live/sites/central/files/textes-leg/4-rech/dir4-5-donnees-rech2.pdf) in French), but the right to make special agreements with external partners on ownership resulting from collaborative projects in reserved. If this is the case for you, we suggest you to contact the [PACTT](https://www.pactt.ch/) to discuss this matter.

**28. If the data created as part of your project is subject to restrictions related to patents or inventions, or to a contract, what restrictions are involved?**

Guidance : For example, mention the terms of the contract, the rights granted, the restrictions, the terms of distribution, etc. Moreover, [the Directive 4.5](https://www.unil.ch/central/files/live/sites/central/files/textes-leg/4-rech/dir4-5-donnees-rech2.pdf) stipulates in its article 15: "2In case of patent filing and upon request of the Technology Transfer Office (PACTT), the data related to UNIL's patents will be kept for the entire duration of the patent's validity (20 years), respectively of the supplementary protection certificate in case of patents on therapeutic products".

# Long-term preservation

*Long-term preservation is the retention of research data after the project is completed. It is therefore different from storage, which aims at preserving data during the project. Not all data necessarily need to be kept for the long term (more than 10 years). In this section you should therefore inform which data from your project are intended to be kept for the long term, why and how.*

[This section relates to question 3.2 of the mySNF DMP].

**29. How much of the research data from your project is expected to be preserved for the long term (more than 10 years)?**

Single choice answer:

[ ] No data

[ ] Between 1% and 25%.

[ ] Between 26% and 50%.

[ ] Between 51% and 75%.

[ ] Between 76% and 99%.

[ ] All data

**30. On which selection criteria is this estimate based?**

Guidance : For example, state the value of the data over the long term or for reuse (quality, integrity, accessibility), the value of the data itself (is it complete? unique? can it be reproduced and at what cost? etc.), the existence of legal or contractual obligations for disposal and destruction of the data, stakeholder requirements, cost, etc.

To learn more about selecting which data to preserve, visit [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/archivage--partage.html).

**31. If you are going to keep some of your data, what archival formats will be used?**

Multiple choice answer

[ ] Web Archive : WARC

[ ] Containers: TAR, GZIP, ZIP

[ ] Tabular data: CSV

[ ] Videos: MOV, MPEG, AVI, MXF

[ ] Geospatial: SHP, DBF, GeoTIFF, NetCDF

[ ] Images: TIFF, JPEG 2000, PDF, PNG, GIF, BMP

[ ] Sounds: WAVE, AIFF, MP3, MXF

[ ] Statistics: ASCII, DTA, POR, SAS, SAV

[ ] Text: XML, PDF/A, HTML, ASCII, UTF-8

[ ] Other (please specify):

Guidance : Consider using open and standard formats to facilitate data sharing and reuse. To learn more about data archiving, visit [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/archivage--partage.html).

# Data sharing and repository

*At the end of the research project, at the latest at the time of publication of your results, you are expected to share the data that were exploited during the project. This sharing may be subject to restrictions. In this section, you must therefore detail the terms and conditions under which data sharing will be implemented.*

[This section relates to questions 4.1 and 4.2 of the mySNF DMP].

**32. Is the sharing of data from your research project subject to access restrictions (embargo, sharing on demand, data that cannot be shared because of legal or ethical clauses, etc.)?**

[ ] Yes

[ ] No

Guidance: The SNSF expects data to be shared as soon as possible, but no later than the scientific publication on which the data are based. Only legal, ethical, copyright, confidentiality or other clauses may impose restrictions. In this case, it is advisable to consider whether a confidentiality agreement could adequately protect the data. UNIL is in line with these recommendations.

**33. If there are any restrictions on the sharing of data from your research project, specify and justify these restrictions:**

Guidance : For example, mention here restrictions due to legal, contractual, ethical or deontological obligations and for how long they apply; the existence of contracts for sharing or non-disclosure agreements etc. Mention which data are concerned. Explain who will be able to access the data and how those restrictions will be implemented.

**34. Under which license will the data be made available?**

Single choice answer:

[ ] CC0

[ ] CC-BY

[ ] CC-BY-SA

[ ] CC-BY-NC

[ ] CC-BY-NC-SA

[ ] CC-BY-ND

[ ] CC-BY-NC-ND

[ ] I don't know yet

[ ] Other (please specify):

Guidance : To be open, data must be free to be used, modified and shared by anyone for any purpose, subject, at most, to requirements that preserve the provenance and openness of the data ([The Open Definition 2015](https://opendefinition.org/)). Thus, clauses prohibiting modification and/or commercial use are incompatible with open the data.

The license that is recommended for open data is the [CC0](https://creativecommons.org/share-your-work/public-domain/cc0/) license of Creative Commons, which places the data in the public domain and thus allows it to be freely used, modified and shared without any restriction.

For more information on licenses, please see [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/conformite--exigences/licences-dutilisation.html).

**35. In which repository will the data be deposited?**

Single choice answer :

[ ] Zenodo

[ ] Dryad

[ ] FORSbase

[ ] DaSCH

[ ] C4science (for computer code)

[ ] Github

[ ] Other (please specify):

[ ] I don't know yet

[ ] I will not deposit my data in a repository (please explain why)

Guidance : According to [the Directive 4.5](https://www.unil.ch/central/files/live/sites/central/files/textes-leg/4-rech/dir4-5-donnees-rech2.pdf), Article 15, "Data related to a publication must be deposited in a non-commercial repository, subject to other requirements formulated by the research funding agency. Non-publication-related data may be archived on a Ci [“Centre informatique”] infrastructure and/or on a non-commercial repository. The decision to deposit is the responsibility of the PI [Principal Investigator]." The SNSF further recommends that research data be archived in recognized scientific data repositories that meet FAIR principles and are non-commercial. To learn more about data archiving and sharing, please visit [this page](https://www.unil.ch/openscience/en/home/menuinst/open-research-data/gerer-ses-donnees-de-recherche/archivage--partage.html) of the Open Science UNIL website.