Research Data Management - Novel Requirements according to the SNSF

Author(s):
Ziehmer, Malin Michelle; Petrus, Ana

Publication Date:
2018-07-09

Permanent Link:
https://doi.org/10.3929/ethz-b-000296561

Rights / License:
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International
Research Data Management

Novel requirements according to the SNSF

Ana Sesartic & Malin Ziehmer

ETH Zurich, ETH Library, Research Data Management and Digital Curation
Disclaimer

- We are **NOT** representing SNSF (Swiss National Science Foundation)
- At any time, **information** on the **SNSF’s website**, in its **regulations** and in **mySNF** is **binding**
- SNSF considers this as a pilot phase
- **SNSF needs feedback from your research practice!**
- Please contact the SNSF via **ord@snf.ch** for questions and comments
So – who are we?

• sharing a scientific background ourselves

• here to discuss data management as part of your research

• to learn more about your needs in the process

• and to motivate you to think critically about the chances and limitations of data management and re-use
SNSF motives for implementing obligatory DMPs

Why?
SNSF Policy on Open Research Data

Goal of the SNSF:

Research data should be freely accessible to everyone – for scientists as well as for the general public.

Article 47 of the Funding Regulations

“[…] the data collected with the aid of an SNSF grant must also be made available to other researchers for further research and integrated into recognised scientific data pools […]”
The SNSF therefore expects all its funded researchers

- to **store the research data** they have worked on and produced during the course of their research work
- to **share these data with other researchers**, unless they are bound by legal, ethical, copyright, confidentiality or other clauses
- to **deposit their data and metadata** onto existing **public repositories** in formats that anyone can find, access and reuse without restriction
SNSF Policy on Open Research Data

- A **data management plan (DMP)** is just one of the tools to reach this goal
- Since **October 2017**, researchers have to **include a DMP** in their **funding application**
- SNSF expects that data generated by funded projects are **publicly accessible** in digital databases provided there are **no legal, ethical, copyright or other issues**
- Research **data sharing** as a **fundamental contribution** to the impact, transparency and reproducibility of research data
- Research data should be shared **as openly as possible**

http://www.snf.ch/en/theSNSF/research-policies/open-access/
What is a Data Management Plan (DMP)?

A brief plan written at the start of a project and updated during its course to define:

- What data will be collected or created?
- How will the data be documented and described?
- Where will the data be stored?
- Who will be responsible for data security and backup?
- Which data will be shared and/or preserved?
- How will the data be shared and with whom?
Aim of DMPs according to SNSF

- Planning and documenting the life cycle of data
- In the ideal case, you only need to document your current practice / best practice in your field
- Making data FAIR:
  - Findable
  - Accessible
  - Interoperable
  - Re-usable

Offering a long-term perspective by outlining how the data will be:
- Generated
- Collected
- Documented
- Shared / Published
- Preserved

Updating the plan as the project progresses
DMPs in practice

How?
How to submit a DMP

- A proposal can only be submitted if a DMP was created
- A DMP for SNSF must be created online in mySNF
- Final version of the DMP will be moved to P³ grants database
- You cannot upload a DMP created outside of mySNF – except in Lead Agency process
- Contents of DMP:
  

https://www.mysnf.ch
Assessment of the DMP

- The DMP is assessed by SNSF staff for plausibility and compliance with its Open Research Data policy.
- It is not sent to external reviewers.
- Applicants can be assigned «tasks» for enhancing their DMP as part of the funding decision.
- DMP Guidelines for researchers
DMP lifetime management

Final DMP version available on the SNSF’s P3 database

(http://p3.snf.ch/)

- The DMP is a living document
- Editable during the entire lifetime of the grant
- Its contents can and should be adapted as the project evolves
- Researchers will be prompted to update their DMP at the end of the grant, which will be assessed together with the final scientific report
Eligible costs

- Data **storage** and **processing** costs

- People are **eligible for the costs** of enabling access (generally up to 10'000 CHF)
  - as long as those data are **open**…
  - …and in a repository which SNSF considers as «**non-commercial**»
  - For deposit in a commercial repository, only data preparation costs will be covered

- Section 2.13 of the General implementation regulations for the Funding Regulations
### Data Management Plan – content of the mySNF form

<table>
<thead>
<tr>
<th>Question</th>
<th>Help text</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 Data collection and documentation</strong></td>
<td></td>
</tr>
<tr>
<td><strong>1.1 What data will you collect, observe, generate or re-use?</strong></td>
<td>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. (This relates to the FAIR Data Principles F2, I3, R1 &amp; R1.2)</td>
</tr>
<tr>
<td>Questions you might want to consider:</td>
<td></td>
</tr>
<tr>
<td>- What type, format and volume of data will you collect, observe, generate or re-use?</td>
<td></td>
</tr>
<tr>
<td>- Which existing data (yours or third-party) will you reuse?</td>
<td></td>
</tr>
<tr>
<td><strong>1.2 How will the data be collected, observed or generated?</strong></td>
<td>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. (This relates to the FAIR Data Principle R1)</td>
</tr>
<tr>
<td>Questions you might want to consider:</td>
<td></td>
</tr>
<tr>
<td>- What standards, methodologies or quality assurance processes will you use?</td>
<td></td>
</tr>
<tr>
<td>- How will you organize your files and handle versioning?</td>
<td></td>
</tr>
</tbody>
</table>
Part I: Data collection and documentation

1.1 What data will you collect, observe, generate or reuse?
- Data origin, formats, estimated data volume

1.2 How will the data be collected, observed or generated?
- 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?
- E.g. metadata standard, software version, etc.
Part I: Data collection and documentation

Data management - metadata

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

E.g. metadata standard, software version, etc.

- Develop your own metadata scheme or use existing standards (preferred)
- Apply metadata as early as possible in data life cycle (i.e. during acquisition)

«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.»
Part II: Ethics, legal and security issues

2.1 How will **ethical issues** be addressed and handled?

2.2 How will **data access and security** be managed?

- Consider e.g. if password protection is sufficient or if you need to encrypt data, especially on mobile devices
- Consider relevant regulatory frameworks (Swiss data protection law, EU GDPR)

2.3 How will you handle **copyright and intellectual property** rights issues?

- Consider under what licenses data will be made available and when
Part II: Ethics, legal and security issues

ETH Guidelines

- ETH Compliance Guide
  - Primary data needs to be carefully archived
  - Personal data need to be preserved according to Swiss data protection law
    - https://rechtssammlung.sp.ethz.ch/Dokumente/133_en.pdf

- ETH Guidelines for Research Integrity
  - All steps must be documented to ensure the reproducibility
  - The project management is responsible for data management
    - https://doi.org/10.3929/ethz-b-000179298 (Version 2011)
Part III: Data storage and preservation

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

- Routine operation while the project is running and data is being worked on

3.2 What is your data preservation plan?

- Must be implemented as part of your project, but points far beyond its end
- Which data must be preserved – and for how long?
- How do you ensure that data remains accessible which is required to justify and reproduce your results?
Part III: Data storage and preservation

Data storage & backup (short-term)

- **Local storage on laptops, external disks etc. is risky**

- **Professional storage provisioning via your IT Support Group**
  - Several options available from central IT services: network attached storage (NAS), long-term storage (LTS), cost defined storage (CDS)

- **Always make backups**
  - Good practice: two copies, one located off-site
  - Backup procedures should be automated (i.e. hourly or daily)
  - Windows Backup tool, OS X Time Machine, Linux FreeFileSync
  - See [https://itsecurity.ethz.ch/en/#/prevent_data_loss](https://itsecurity.ethz.ch/en/#/prevent_data_loss)
Part III: Data storage and preservation

Data preservation (long-term)

- How will data be preserved in the longer run?
- Is the **repository** you choose **sustainable**?
- Will data be available in an **open and stable format**?
- Is your data **citable** via persistent identifiers, e.g. DOI?

**Ideal case:**

- You **use open, documented file formats** and document others comprehensively
- You **choose a sustainable repository** which ensures preservation and registers persistent identifiers automatically, e.g. international subject repositories or [Research Collection at ETH](https://www.research-collection.ethz.ch/)
Part III: Data storage and preservation

Data preservation (long-term)

**Everyday storage on NAS + backup**
- **up to 10 years**
- Immutable storage (secondary copy does not depend on primary)
- It is **up to you to keep track** of what is where
- Satisfies requirements for being able to present data on demand

**Storage on LTS**
- **10 years to permanent**
- Storage on NAS/HSM or LTS, both with replication
- Retention for 10 or 15 years or permanently
- **Metadata required, citable with DOI, format** analysis
- Satisfies requirements for data sharing, publishing, and preservation
Part IV: Data sharing and reuse

4.1 How and where will which data be shared?

4.2 Are there any necessary limitations to protect sensitive data?

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

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

**Meaning:**

- You can opt out of data publication with a cause, but **you must give reasons**
- You choose a repository meeting the criteria. It can be subject specific, generic or an institutional one like **Research Collection** at ETH (www.research-collection.ethz.ch/)
Information from SNSF to support you

- **Collection of SNSF information** on Open Research Data including FAQ:
  
  [http://www.snf.ch/en/theSNSF/research-policies/open_research_data/](http://www.snf.ch/en/theSNSF/research-policies/open_research_data/)

- SNSF’s explanation of the **DMP expected content**:
  

- **Guidance for ETH researchers** on filling out SNSF Data Management Plans:
  
  [https://documentation.library.ethz.ch/display/DD/Guidance+for+ETH+researchers+on+filling+out+SNSF+Data+Management+Plans](https://documentation.library.ethz.ch/display/DD/Guidance+for+ETH+researchers+on+filling+out+SNSF+Data+Management+Plans)

  PDF: [https://documentation.library.ethz.ch/download/attachments/13762611/DLCM_SNSF-DMP_v2.pdf](https://documentation.library.ethz.ch/download/attachments/13762611/DLCM_SNSF-DMP_v2.pdf)

  - Includes:

    explanations per question, examples from DMPs, contacts and links specific for ETH Zurich
Send feedback to SNSF!

- Consider this as a pilot phase
- SNSF is aware of limitations: not everything applies to everyone – give reasons
- SNSF needs feedback from your research practice!
- Please get involved with your colleagues:
  - What do you consider as best practice in your field?
- If you encounter difficulties or have comments, suggestions, questions:
  Please contact the SNSF via ord@snf.ch
Services @ ETH Zurich

How we can support you in managing your research data
**Services @ ETH Library**

- Share and publish research output according to SNSF guidelines for **FAIR** data:
  
  ETH Research Collection ([https://www.research-collection.ethz.ch](https://www.research-collection.ethz.ch))

- Publications, **Research Data**

- Web upload, **DOI-reservation** and registration, ORCID, Export to OpenAire…

- Long-term preservation in **ETH Data Archive** ([http://www.library.ethz.ch/Digital-Curation](http://www.library.ethz.ch/Digital-Curation))

---

FAIR image (4.9.2018) by Sangya Pundir / CC BY-SA 4.0
Services @ ETH Library – Research Collection

Registry of Publications/ University Bibliography
- Source for AAA
- Personalized publication lists
- Import from Web of Science and Scopus

Open Access Repository
- Primary and secondary publication
- Legal check/legal consulting
- Consulting on OA

Research Data Repository
- Publication of Research Data
- DOI reservation and registration
- Automatic archiving

www.research-collection.ethz.ch
Services @ ETH Library

- Share and publish Research Output according to SNF guidelines for FAIR data: ETH Research Collection (https://www.research-collection.ethz.ch)
  - Publications, Research Data
  - Web upload, DOI-reservation and registration, ORCID, Export to OpenAire…
  - Long term preservation in ETH Data Archive (http://www.library.ethz.ch/Digital-Curation)
Services @ ETH Library

- Share and publish Research Output according to SNF guidelines for FAIR data: ETH Research Collection (https://www.research-collection.ethz.ch)
  - Publications, Research Data
  - Web upload, DOI-reservation and registration, ORCID, Export to OpenAire…
  - Long term preservation in ETH Data Archive (http://www.library.ethz.ch/Digital-Curation)

- Get support for Open Access (http://www.library.ethz.ch/en/Open-Access) including payment of Article Processing Charges (APCs) with a range of publishers

- DOI reservation and registration (http://www.library.ethz.ch/DOI-Desk-EN)

- ORCID (http://www.library.ethz.ch/en/ORCID - add your ORCID ID to your nethz-account)
Scientific IT Services (ID-SIS)

Data Storage & Versioning

- Storage provisioning (usually via your IT Support Group)
- Gitlab - gitlab.ethz.ch (hosted by IT services)
- SharePoint - mysite.sp.ethz.ch (free up to 1 GB)

Active Research Data Management (ARDM)

- Active research data management support [www.sis.id.ethz.ch/researchdatamanagement](http://www.sis.id.ethz.ch/researchdatamanagement)
- **openBIS** Laboratory Information Management System (LIMS) & **Electronic Lab Notebook** (ELN) for quantitative sciences [https://labnotebook.ch/](https://labnotebook.ch/)
- **Developed** and **supported** by ETH Scientific IT-Services (ID-SIS)
- Specific **DMP templates** available for openBIS users
- Development of **custom data management solutions**
- since 2018, ID-SIS has the mandate to provide active data management services to all ETH
Additional services @ ETH

ETH transfer
https://www.ethz.ch/en/the-eth-zurich/organisation/staff-units/eth-transfer.html

- Advice on Intellectual Property, Patents, Licensing of Software etc.
- Software disclosure workflow with ETH Data Archive (http://www.library.ethz.ch/Digital-Curation)

Statistical Consulting Service @ D-MATH
https://www.math.ethz.ch/sfs/consulting.html

- Consulting service and contractual data analysis
  https://www.math.ethz.ch/sfs/consulting/consulting-service.html
- Statistics and software courses
  https://www.math.ethz.ch/sfs/consulting/software-courses.html
Trainings @ ETH

- **Trainings & Courses by ETH Library** (e.g. data management, information research, reference management, scientific writing, open access, «Book a Librarian»):
  
  www.library.ethz.ch/en/Services/Training-courses-guided-tours

- **Trainings & Courses by ETH SIS** (e.g. openBIS, Python, HPC, bioinformatics): sis.id.ethz.ch/consulting

- **Research Data Management Workshops** by ETH Library and ETH SIS: www.ethz.ch/researchdata

- Courses by the ETH Information Center for Chemistry/Biology/Pharmacy: www.infozentrum.ethz.ch/en/whats-up/events/

- Further topics on demand researchdata@ethz.ch
Thank you! Questions?

Research Data Management and Digital Curation

www.library.ethz.ch/Digital-Curation
data-archive@library.ethz.ch

Dr. Ana Sesartic
Research Data Management and Digital Curation
ETH Library
Rämistrasse 101
8092 Zurich
044 632 7376
ana.sesartic@library.ethz.ch

Dr. Malin Ziehmer
Research Data Management and Digital Curation
ETH Library
Rämistrasse 101
8092 Zurich
044 633 86 32
malin.ziehmer@library.ethz.ch

Research Data

www.ethz.ch/researchdata
researchdata@ethz.ch