ETH zürich

Data Management Plans for SNSF Grants

Dr. Ana Sesartic & Dr. Matthias Töwe (ETH Library Digital Curation Office)
Dr. Henry Lütcke (ETH Scientific IT Services)

www.ethz.ch/researchdata

Disclaimer

- We are not representing SNSF
- This event is based on the results of discussions with SNSF staff
- At any time, information on the SNSF's website, in its regulations and in mySNF is binding
- SNF considers this as a pilot phase
- **SNF needs feedback from your research practice!**
- Please contact the SNSF via ord@snf.ch for questions and comments

SNSF reasoning behind DMP

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
  (1 Jan 2016, http://www.snf.ch/SiteCollectionDocuments/allg_reglement_16_e.pdf):
  "[…] 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 […]"
- A data management plan is just one of the tools to reach this goal
- Please also be aware of SNSF's updated Open Access Policy for Publications and changes to the General implementation regulations for the Funding Regulations!
  http://www.snf.ch/en/theSNSF/research-policies/open-access/

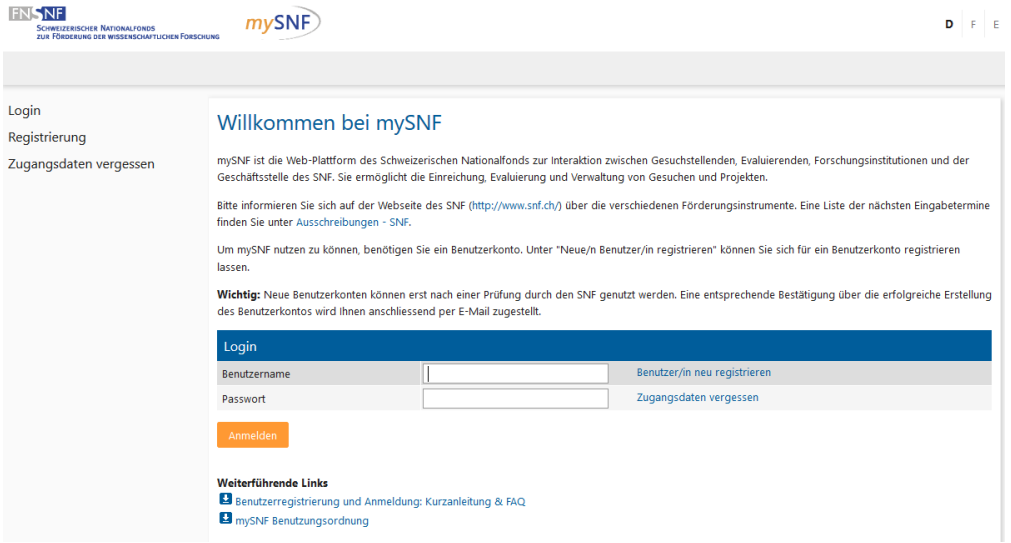
Aim of the DMP according to SNSF

- **Planning the life cycle** of data
- **Updating the plan** as the project progresses
- Offering a **long-term perspective** by outlining how the data will be:
  - Generated
  - Collected
  - Documented
  - Shared / Published
  - Preserved
- 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

DMP in practice

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**
- You cannot upload a DMP created outside of mySNF – except in Lead Agency process



https://www.mysnf.ch

- Contents of DMP: http://www.snf.ch/SiteCollectionDocuments/DMP_content_mySNF-form_en.pdf

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
  http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/data-management-plan-dmp-guidelines-for-researchers.aspx

Lifetime Management

- 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
- Final DMP version will be made available on the SNSF's P3 database (http://p3.snf.ch/)

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 (15 Jan 2018, http://www.snf.ch/SiteCollectionDocuments/snsf-general-implementation-regulations-for-the-funding-regulations-e.pdf)

DMP content of the mySNF form

Essence of Research Data Management

«…tracking back to what you did (several) years ago
and recovering it (…) immediately
in a reusable manner.»

Henry Rzepa, Professor of Computational Chemistry, Imperial College London

Data collection and documentation

- **What** data will you collect, observe, generate or reuse?
  - Data origin, formats, estimated data volume
- **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?
- What **documentation** and **metadata** will you provide with the data?
  - E.g. metadata standard, software version, etc.

"Piled Higher and Deeper" by Jorge Cham
www.phdcomics.com

Data Management - Metadata

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

Questions you might want to consider:

- What information is required for users (computer or human) 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?

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. (This relates to the *FAIR Data Principles* I1, I2, I3, R1, R1.2 & R1.3)

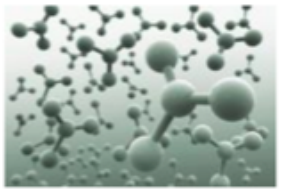
- Develop your own metadata schema or use existing standard (preferred)
  - See http://www.dcc.ac.uk/resources/metadata-standards
- Apply metadata as early as possible in data life cycle (i.e. during acquisition)

Search by Discipline

Biology

Earth Science

General Research Data



Physical Science

Social Science & Humanities

Ethics, legal and security issues

- How will **ethical issues** be addressed and handled?
- 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)
- How will you handle **copyright and intellectual property** rights issues?
  - Consider under what licenses data will be made available and when

ETH Guidelines

- 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)

- 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/133en.pdf

Data storage and preservation

- 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
- 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?

Data Storage & Backup

- 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)
  - See https://www.ethz.ch/services/en/it-services/catalogue/storage.html
- 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

Data preservation

- 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/)

Data sharing and reuse

- How and where will which data be shared?
- Are there any necessary limitations to protect sensitive data?
- I will choose digital repositories that are conform to the FAIR Data Principles
- 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 to support you

- Collection of **SNSF information** on Open Research Data including **FAQ**:
  http://www.snf.ch/en/theSNSF/research-policies/open_research_data/
- SNSF's explanation of the **DMP expected content**:
  http://www.snf.ch/SiteCollectionDocuments/DMP_content_mySNF-form_en.pdf
- **Guidance for ETH researchers** on filling out SNSF Data Management Plans:
  https://documentation.library.ethz.ch/pages/viewpage.action?pageId=14680081
  PDF: http://www.library.ethz.ch/de/Media/Files/DLCM-template-for-the-SNSF-Data-Management-Plan
  - 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**?
  - SNSF offers **financial support for community workshops** via the funding scheme **«Scientific Exchanges»** (http://www.snf.ch/en/funding/science-communication/scientific-exchanges/)
- If you encounter difficulties or have comments, suggestions, questions:
  Please contact the SNSF via ord@snf.ch

ETH zürich

Services at ETH Zurich

Services at 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 with a range of publishers
- **DOI 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)

IT Services

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)

Research Data Management

- Research data management support *www.sis.id.ethz.ch/researchdatamanagement*
- openBIS **L**aboratory **I**nformation **M**anagement **S**ystem & **E**lectronic **L**ab **N**otebook for quantitative sciences
- Developed and supported by ETH SIS
- Specific DMP templates available for openBIS users
- See *https://openbis-eln-lims.ethz.ch*
- Development of custom data management solutions

Other Services

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

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

- 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

Contact information

Research Data Website: www.ethz.ch/researchdata

Research Data Mailbox: researchdata@ethz.ch

Ana Sesartic
ETH Zurich
ETH Library
ana.sesartic@library.ethz.ch

Matthias Töwe
ETH Zurich
ETH Library
matthias.toewe@library.ethz.ch

Henry Lütcke
ETH Zurich
Scientific IT Services
hluetcke@ethz.ch

www.library.ethz.ch/Digital-Curation

www.sis.id.ethz.ch