Drafting a Data Management Strategy Instructions for ETH Zurich Research Groups

Abbreviations

- DMS – Data Management Strategy
- DMP – Data Management Plan
- RDM – Research Data Management
- OA – Open Access

About this page

The purpose of this document is to instruct and support research groups at ETH Zurich with drafting a Data Management Strategy (DMS) for their own research group, lab, or professorship. These instructions are in compliance with the ETH Zurich Guidelines on scientific integrity[1] and the Guidelines for Research Data Management at ETH Zurich[2]. More specifically, the latter aim at supporting the establishment of best practices for working with data in line with Community Standards' required in Art. 4 of the Research Data Management Guidelines.

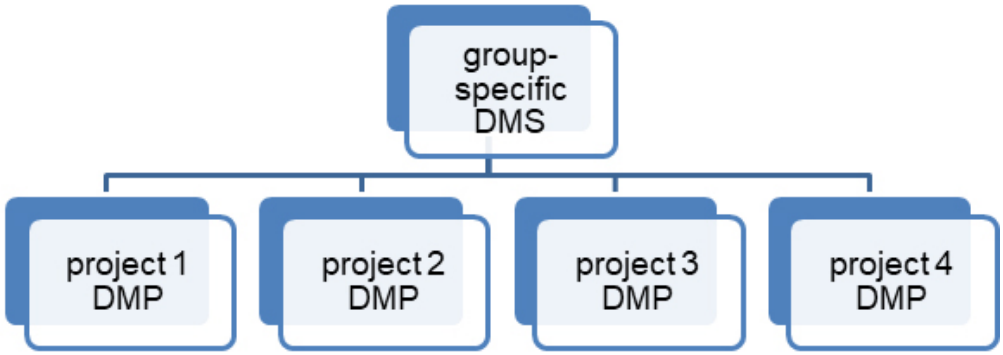
This instruction page will first define what a Research Data Management Strategy (DMS) is and how a DMS differs from a Data Management Plan (DMP). The included implementation steps provide suggestions and links to additional information that support you in (a) taking decisions on how to manage research data in your lab or research group and (b) documenting these decisions in written form. The DMS should reflect your research group's decisions and practices in data management. Changes can and should be made while your projects and your research group evolve. All our suggestions made in this guide are optional and should be adapted to your individual work situation. To refer to this guide, please use the following: "ETH Library (2022), Data Management Strategy Instructions for ETH Zurich Research Groups".

What is a Research Data Management Strategy?

In some disciplines, specific guidelines for handling research data ensure structured and uniform data management. Defined rules facilitate the planning of new projects within a research group and offer new members a helpful guide to the established standards. Beyond discipline-specific conventions and guidelines, good practice principles in Research Data Management (RDM) suggest that documenting strategies and practices in managing research data is beneficial to the success of research projects as well as to collective data management in teams.

A Data Management Strategy (DMS) documents data management practices on a level beyond individual research projects by coordinating research data management across one (or several) research group(s), lab(s), professorship(s), or institute(s). The DMS should reflect your decisions and practices in data management, although the strategy can and must be updated regularly to reflect current changes in procedures or to meet actual needs. It mirrors a project-specific Data Management Plan (DMP) insofar as it provides answers to DMP questions on the level of a research group, lab, professorship or institute. Figure 1 provides a schematic illustration of the relationship between a DMS of a research group and several project-specific DMPs for research projects conducted in that group.

Figure 1. Relationship of DMS and DMP



Can I use this guide for preparing a Data Management Plan (DMP)?

This guide addresses ETH Zurich members who want to establish a data management strategy. If, instead of a DMS, you have to prepare a Data Management Plan (DMP), for instance for a project application to the Swiss National Science Foundation (SNSF), you can find practical tips in the step-by-step guide created by the ETH Library in collaboration with the EPFL Library and other project partners. Moreover, for projects without explicit funder requirements, we provide a general DMP template and corresponding instructions that takes into account the specific conditions at ETH Zurich.

Implementation steps

The following instructions should assist you in defining rules and guidelines for the coordination of daily and long-term data management activities in your research group. Such guidelines can be documented in your DMS, which every group member should approve and adhere to.

Overview

1. Responsibilities

1.1 List of RDM tasks and responsibilities in your research group

Define tasks for data management that must be fulfilled in your research group. We recommend assembling a **list of research data management responsibilities** for every active member or alternatively for the functional levels that exist within the group. We suggest placing such a list in a location with shared access and for read-only purposes so that every group member can easily access it whenever required. When new members join the group or when group members are leaving, the list should be updated accordingly.

The list can include RDM responsibilities that are not specific to a certain role and thus apply equally to all members of the group.

If the amount of work spent on RDM and related tasks is extensive among group members it might be an option to assign the role of a data manager[3] or data steward[4] to a specific group member. This person could take care of the group's RDM activities to a certain extent. Is there somebody assigned to spend part of their work on data management duties (e.g., data manager or data steward) for the entire group?

Make explicit where to draw the line between the duties of the data steward on the one hand and RDM duties that every member of the group is responsible for on the other hand.

The list describing specific responsibilities for data management may include but is not limited to the following:

- group member's name (optional, since responsibilities can also be defined for persons depending on their role – see below)
- role / functional level (e.g., data steward, postdoc, doctoral student, scientific staff)
- data management tasks (e.g., data upload to repository, data steward)
- estimated time to be spent on data management duties
- group-specific considerations (e.g., part-time employment)
- discipline-specific considerations (e.g., regarding special technical equipment)
- define which tasks or decisions require approval by or at least communication with the research group leader.ack to overview

[back to overview]

1.2 Confirmation of group members

Obtain a confirmation from each group member to comply with the guidelines set out in this data management strategy. This might also be relevant for temporary staff like research assistants, who are involved in, for example, data collection or documentation.

[back to overview]

2. Data collection and documentation

2.1 List of scientific data and metadata:

We recommend compiling a list about data collection and data documentation procedures which shows frequently employed scientific methods together with the data types that are created by them. We suggest placing such a list in a location with shared access and for read-only purposes so that it is available to every group member whenever required.

Your list might contain but is not limited to the following:

- date (e.g., YYYYMMDD format)
- responsible person (e.g, name or initials of the person who gathers the (meta)data)
- name of scientific method for (meta)data collection
- method ID (i.e., a group-specific identification number associated with the method)
- data type and corresponding file format associated with the input and output of the scientific method
- if applicable: additional open file format (recommended for the purpose of long-term reusability)
- short description of the method (e.g., standard protocol used)
- scientific metadata (i.e., defined types of metadata required)
- if available, disciplinary metadata standards should be followed wherever possible (i.e., community standards for different experiments and corresponding file types). More information on such community standards is available in section 1.3 of our Data Management Plan Instructions for ETH Zurich Researchers
- quality control (i.e., measures to verify the quality of the results)

Some of the scientific metadata that you are gathering can also be documented in a “Readme” file that you can, for example, upload to a data repository together with your dataset. Such “Readme” files are standard text files that are intended to ensure reusability of the dataset by yourself or other researchers at a later date. A guide for writing “Readme” files is available on our Wiki-page.

[back to overview]

2.2 Research data ownership in contract agreements with third parties

Be aware that research data are considered as project results from a legal point of view. Therefore, in cooperation contracts with project partners, one should explicitly stipulate ownership and usage rights of the research data that will be created during the collaboration. Since exploitation of research data and research results might differ, an explicit agreement on the research data themselves can ensure that the data can still be reused in subsequent projects. For example, some research data might qualify for more open licenses and public sharing, while other project results might not.

[back to overview]

2.3 File types and formats

Think about the various file types and corresponding file formats that will arise during research projects conducted in your group. Consider all data related to your research project.

For the sake of long-term reusability, we recommend converting proprietary data formats into open and properly documented standard formats, if applicable. For further details on open file formats of the most popular file types as well as on how to convert files into these formats, the Research Data Management and Digital Curation team at ETH Library offers a Wiki-page on recommended standard file formats.

We suggest creating a list of common file types and corresponding proprietary and open file formats that exist in the group including preferences for the most appropriate formats for your group. File types and formats can also be included in the data collection and documentation list suggested in paragraph 2.1 above.

[back to overview]

2.4 File naming convention

Consistency in file naming helps to easily identify relevant data files and thereby facilitates interoperability for instance in collaborative work among group members or even more so in work together with people outside of the group. It avoids duplication of files and potential data loss as a consequence of overwriting files with distinct content but identical names. A descriptive part in the name of a file can also simplify identification of its content without the need for opening it. Therefore, consistent file naming can save a lot of time during your daily research activities and is key for an efficient research process. You will probably use abbreviations or codes to limit file name lengths. Make sure that those abbreviations are known to all members and used consistently. Obviously, keys to abbreviations and code must be documented and available to everyone.

There is some general advice on how to name your files:

- Use unique file names that reflect the content of a file to some extent
- Use ASCII characters (avoid e.g.: ~ ! @ # $ % ^ & * ( ) ` , ; < > ? [ ] { } ' " |)
- Do not distinguish files by upper- and lower-case letters
- Avoid spaces and use points only before file extension
- Dates should be written in YYYYMMDD or YYYY-MM-DD (chronological order when at the beginning of the file name)
- Example: 2022-04-15_JohnDoe_myfile_v02

[back to overview]

2.5 Defined vocabularies

Sticking to a list of defined terms that are commonly used in your research group increases consistency in research results and research data. Furthermore, defined vocabularies that exist among research communities and beyond make research data interoperable, machine-readable, and reusable for other researchers.

Describe the intended use for every term that you are using during your daily work as a researcher, and which might be prone to misinterpretation. You might want to collect these terms in a list. Stay in line with community or international standards. More information on disciplinary and international standards is available in section 1.3 of our Data Management Plan Instructions for ETH Zurich Researchers. You can further consider the list of scientific conventions on names and keywords for some disciplines in the guide for writing “Readme” files on our Wiki-page. We suggest taking into account your list of defined terms when defining your file naming convention. You could also aim at going a step further by making your vocabulary fully FAIR (see here[5]).

[back to overview]

2.6 File folder organization and code version control

Defined allocation of certain data to specific folders and subfolders provides additional benefit regarding reuse of data for yourself and for other researchers who will have access to your file storage in the long run. Think about the different file folders for distinction and storage of your various file types, implement a folder hierarchy and specify the file types that you will allocate to your given file folders. You might want to define a guide for a common file folder structure (e.g., graphical illustration) that can be adopted by your group members. Do not forget to define the type of data that every single folder should comprise. You might want to define a meaningful and systematic connection between your file naming convention and your defined folder structure, including folder content specifications.

There are some general aspects on file folder organisation one might consider:

- Path names should not exceed 255 characters
- Do not use more than 5 levels in your folder hierarchy
- Use unique names at least within the same folder hierarchy and avoid overlapping parts in folder names (e.g., “Data” vs. “Presentation-Data”)
- Distinguish between different kinds of content (e.g., work and private, own and others’ work, research and administrative, raw and processed data, experiments and analysis, experimental runs)
- Relocate files you are no longer working with to a steady storage location (backed up)

Groups with significant computational work should adopt a version control system and agree on a remote code repository (e.g., ETH GitLab, https://gitlab.ethz.ch).

[back to overview]

3. Data Security

Different protection levels are necessary when handling and storing research data depending on the risk posed to ETH and the information owner[6] when unauthorised individuals gain access to the information in question. Additional legal provisions apply when working with sensitive personal data[7] and affected persons must be informed in advance and need to give their consent to the use of their data. Further information on data security issues is provided on the Webpage about Data Security of the ETH IT Services.

3.1 Storage and Backup

We recommend using the professionally managed storage services of ETH IT Services or those provided by your departmental IT Support Group, which ensure daily, automatic backup. The data is retained for 90 days meaning that older versions of your data can be restored within that time frame.

If you decide to manage your backups yourself, please consider the following: backup frequency (daily is recommended), retention period for restoring older versions, spatial separation of copies of data (geo-redundancy of at least 2 distinct locations which are far away from each other, as a minimum: not in the same building), safe location for backup storage (e.g., cooled room with restricted access).

[back to overview]

3.2 Data classification and data access

Four different classification levels (i.e., PUBLIC, INTERNAL, CONFIDENTIAL, STRICTLY CONFIDENTIAL) exist at ETH Zurich for the data to be protected and “information owners are responsible for classifying information that falls within their remit.”[8]

ETH Zurich research groups have to use the classification system indicated above to define the level of protection and corresponding access rights that the various types of research data generated in the group belong to. For instance, non-anonymized, sensitive personal data[9] must be classified as STRICTLY CONFIDENTIAL. More information including a classification recommendation, which is based on risk assessments, can be found in the Directive on “Information Security at ETH Zurich”[10]. A draft of a list with classified examples of research data including data access and user-specific rights is available.

In principle, ETH-internal or external Information and Communications Technology (ICT) services such as cloud services that are offered and approved by the IT operators of ETH Zurich or the persons responsible for IT at the institutes and professorships, are to be used for data processing and storage.[11] A non-exhaustive list of released external cloud services at ETH Zurich, which also takes into account the different data classification levels, can be found here.

The ETH Zurich IT Services department is responsible for the ETH Zurich-wide technical and operational review of ICT resources on behalf of the Chief Information Security Officer.[12] When in doubt about the use of external ICT resources, please contact central IT services or the IT support groups at your department.

ETH internal cloud services such as polybox with storage space available to all ETH members can be used without hesitation to store and process data that belongs to classification levels INTERNAL and CONFIDENTIAL. By contrast, data classified as STRICTLY CONFIDENTIAL shall never be stored or processed in a cloud. For processing and collaboration on data defined as STRICTLY CONFIDENTIAL the highly secure data platform Leonhard Med can be used.

[back to overview]

3.3 Software and infrastructure for active and secure research data management and transfer of research data:

Define in your DMS, which software and infrastructure for active and secure research data management and transfer of research data you are planning to use. Depending on the conditions defined for data classification and data access (section 3.2), only certain infrastructure and software might be suitable. As an example, commercial software providers that are not hosted at ETH Zurich or that do not ensure storage in Switzerland or the EU, might transfer stored data to the USA, where other safety and data protection standards apply. A list of services and IT infrastructure available at ETH Zurich can be found at ethz.ch/services/en/it-services/catalogue. This is not an exclusive list. For more specific requests please contact IT Service Desk.

[back to overview]

4. Publication and sharing

Handling research data FAIR (i.e., making data Findable, Accessible, Interoperable, and Reusable) has become good practice in science. A key aspect of making one's data FAIR is publishing them in a FAIR data repository. It is suggested as good practice and required by ETH Zurich (see RDM Guidelines) as well as by funding agencies like the SNSF to share at least the data underlying a publication at the time of publication of research results. Research publications such as journal papers can be published under Open Access conditions in line with ETH Zurich's Open Access policy and funder requirements.

4.1 FAIR data repositories

Data should be made available in compliance with the FAIR data principles[13], including assignment of a persistent identifier (e.g., DOI), and open access to the data apart from sensitive personal data and confidential data that cannot be publicly shared. If possible, deposit your data in a well-established, trustworthy, non-commercial repository that implements the FAIR principles. If available, a FAIR data repository that operates within the researchers' field(s) of research is often a better choice than a general data repository since it might better reflect metadata community standards. Alternatively, the institutional FAIR data repository of ETH Zurich, the Research Collection, can be used. Also keep in mind: "If a recognised ETH-external FAIR repository is used for depositing Research Data or Programming Code, it is recommended to register a metadata-only item in the ETH Research Collection linking to the dataset" (RDM Guidelines[14]).

Specify in the DMS, which FAIR data repository (or several repositories) is/are used in the research group/lab and how the workflow to submit research data to the repository looks like. Such a workflow could include details or recommendations on when data is ideally submitted to the repository (e.g., after acceptance of a paper) and which metadata should accompany a submitted dataset.

There are several resources that can support you:

- A wide range of data repositories are listed on re3data.org (most should be listed there) or fairsharing.org/databases.
- The SNSF has its own recommendation list of repositories.
- We provide a Step-by-Step Guide on Data Publication that supports you in preparing your datasets for publication in a FAIR data repository.
- You might also want to check our recommendations for a Data Availability Statement (for more details see section 4.3 below), which is suitable to indicate the location of your research data underlying your publication.

[back to overview]

4.2 Data selection for sharing

Define, which part of the research data collected or produced in the research group/lab is supposed to be shared. Such specifications should follow discipline-specific conventions or recommendations, research funder requirements, and requirements applicable at ETH Zurich. For example, the *Swiss National Science Foundation (SNSF)* "expects its funded researchers to share at least all data underlying a publication, meaning that these data have to be directly and freely available and deposited on a FAIR data repository. Shared data must enable other researchers to reproduce the published study."[15] The RDM Guidelines at ETH Zurich follow the same principle.[16]

If there is a specific person (or several) responsible for implementing parts of the data sharing workflow, make sure this is in line with the list of responsibilities in your research group (see section 1). Define, what metadata shall be provided by individual researchers in the group (e.g., title of the dataset, names of creators, other metadata, files in a zip/tar container) so that the dataset is ready for upload to the selected repository.

If some of the research data cannot be shared due to legal, ethical, copyright, confidentiality or other constraints, mention these constraints and define how to ensure FAIR data standards that are appropriate in such situations. One should, for instance, at least provide the metadata of the dataset that cannot be made publicly available. In such a case, define for the group/lab how to provide suitable descriptive metadata of the dataset in a repository, if possible and as long as these metadata are not (strictly) confidential themselves.

Table 1 below provides data categories that are expected to be shared based on the criteria of the SNSF.

Table 1. Data Sharing – Basic criteria of the SNSF

Data category	Publishing/Sharing?
Data underlying publications	Yes, minimum requirement by SNSF*
Data related documentation / **metadata**	Yes, required by SNSF*
Code needed for processing	Yes, required by SNSF*
Other processed data – if not fully recreatable from raw data	Encouraged by SNSF
Raw data	Encouraged by SNSF,* but often not easily feasible
Lab Notebooks / Lab Logs (paper and electronic)	No

(Strictly) Confidential data [note that data from research in humans might need to be destroyed at some point]	No
Internal project documentation (correspondence etc.)	No

*see the DMP guidelines of the SNSF

[back to overview]

4.3 Data availability statements

At ETH Zurich, the following applies: "All publications of research results must contain a *Data Availability Statement*" (RDM Guidelines).[17] A Data Availability Statement. A Data Availability Statement is a short paragraph that specifies where the data that are underlying a scientific publication can be found. The statement indicates the location of the dataset by providing a persistent web link, for example a DOI link (e.g., doi.org/10.3929/ethz-b-000315707). Typically, these statements are included directly in a scientific publication. Scientific journals increasingly require authors to include data availability statements. Beyond such recommendations, providing data availability statements should be considered good practice in RDM and research integrity and must be included in publications of research results created by ETH researchers (see RDM Guidelines).

Specify in your research group's/lab's DMS, how researchers should handle data availability statements. Even in situations when no research data in a narrow sense are underlying a publication (e.g., for literature reviews or in opinion pieces), providing a data availability statement can ensure clarity about the fact that no data exist that could be linked to the article. A guide with examples for data availability statements is available at the Wiki resource page of the ETH Library team for Research Data Management and Digital Curation.

[back to overview]

4.4 Usage of licenses

In principle, the data creator is the owner (i.e., copyright holder) of the data and therefore has the right to assign a license to these data. To clarify under which condiditions your research data can be reused, you should indicate in the DMS the licenses that are recommended to be used for data publishing in your research group. Recommendations from the side of ETH Library can be found in the Research Collection handbook (for papers and research data see the page of the ETH Library on Creative Commons licenses, and for open-source software and scripts see the page of the ETH Library on licensing open source software and scripts).

However, always take into account which licenses your funding agency suggests or even prescribes.

There might also be use cases where more restrictive access options are necessary, for example when research data classified as INTERNAL, CONFIDENTIAL, or STRICTLY CONFIDENTIAL are concerned (see section 4.2 of these instructions). Such data can obviously not be licensed. In addition, anonymized, previously personal data such as, for example, personal health data or household income data require special attention. In such cases, ETH Library does not recommend using open licenses such as CC licenses even if these data can be shared.

Software (including code scripts in, for example, R, Python, Matlab) that was produced at ETH Zurich must be registered at ETH transfer under any circumstances, before licensing it (under open source licenses or more restrictive ones).

[back to overview]

4.5 Open-access (OA) strategy for publications

Funders such as the SNSF require funded researchers to publish their research output under OA conditions.[18]

ETH Zurich's open access policy[19] encourages all ETH Zurich scientists to publish their research results in OA channels. It obligates ETH Zurich researchers to take the Green Road (self-archiving of scientific publications on an institutional or subject-specific repository) and encourages them to take the Gold Road (direct OA publication of an article with an OA publisher).

Specify in the DMS, how OA publishing is handled in the research group or lab, by mentioning recommended or optional repositories for the Green road. For example, ETH Zurich's institutional repository is the Research Collection, which is only one option for Green OA publishing.

Further information on OA publishing is available at library.ethz.ch/open-access-en

[back to overview]

5. Data long-term preservation

Long-term preservation of data comprises all measures (e.g., file format evaluation and potential file format conversion into a current format) that ensure reusability of data in the future. A minimum storage period of 10 years after deposition or publication has been defined by funders (e.g., SNSF) and ETH Zurich (see RDM Guidelines) as a reasonable time for the preservation of scientific data. Long-term preservation is best enabled by the usage of open-source and widely used file formats (see also section 2.3.).

Preservation of research data for 10 or more years requires careful planning and might come with additional effort and costs. Therefore, you need to define in the group the type of your data which possess long-term value for your research community and even beyond. Such data that have long-term value might include but are not limited to the following types:

- data that are unique and are not professionally stored elsewhere
- data that are unrepeatable (e.g., observational data from astronomical events, environmental events, weather events)
- data that are highly reusable (e.g., in research, for teaching, for commercial applications or policy making)
- data that are underpinning published research results

Besides data archives, one solution to keep research data for a relatively long storage period of e.g. 10 years, can be the *Long Term Storage* (LTS) at ETH Zurich. This solution is offered by the ETH Zurich departmental IT (see contact list). Be aware that in the LTS your data will simply be stored by bitstream preservation, which is clearly distinct from active, long-term preservation measures that are undertaken in professional data archives.

5.1 What happens with the data when they are not actively used anymore (e.g., at the end of a project)?

Define in the DMS how you deal with research data at the end of a project which is carried out in the group. For example, the DMS might require from each project to draft a list of the kind of data which you want or must preserve for long-term and define non-proprietary and sustainable file formats that will be used for data preservation.

[back to overview]

5.2 Data usage agreement: group leader and leaving employee

Responsible persons and members of the project team should define as early as possible to what extent research data and materials can be further used by persons outside the group or by group members who are leaving the project team or ETH Zurich (Integrity Guidelines).[20] ETH legal office offers a template (document: "Agreement on Employees leaving research group") to establish individual data sharing agreements.

[back to overview]

5.3 Professors or research group leaders who are leaving ETH Zurich or retire

Professors or research group leaders should determine with the responsible department the possible future use and exploitation of research data and materials that remain at ETH Zurich after they have left ETH (Integrity Guidelines[21]) and define processes in order to prevent the loss of valuable data and to make sure that archiving periods are observed (RDM Guidelines[22]).

[back to overview]

6. Determining Costs of Research Data Management

RDM can involve additional costs. Ideally, make explicit which costs are directly related to RDM activities in your project(s) or research group and how they are covered. This information can be included in a supplementary document with access for instance restricted to those team members with budgetary responsibilities.

A guide for calculating RDM costs from UK Data Service is available at the Wiki resource page of the ETH Library team for Research Data Management and Digital Curation.

RDM costs can potentially be covered by your research funder or by additional funding lines. The SNSF covers costs related to the preparation of data for archiving and for the archiving itself in a non-commercial FAIR data repository that incur within the scope of an approved project.[23]

In contrast, similar costs incurring for data uploaded to a commercial FAIR data repository will not be covered.

[back to overview]

Endnotes

[1] ETH Zurich Guidelines on scientific integrity (Integrity Guidelines, RSETHZ 414, in force 01 January 2022), https://rechtssammlung.sp.ethz.ch/Dokumente/414en.pdf

[2] Guidelines for Research Data Management at ETH Zurich (RDM Guidelines, RSETHZ 414.2, in force 01 July 2022), https://rechtssammlung.sp.ethz.ch/Dokumente/414.2en.pdf

[3] “A data manager is a person responsible for the management of data objects including metadata. These people think about managing and preserving data". Jetten M. et al., 2021. Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship (1.1). Zenodo. doi.org/10.5281/zenodo.4623713

[4] “A person responsible for keeping the quality, integrity, and access arrangements of data and metadata in a manner that is consistent with applicable law, institutional policy, and individual permissions. Data stewardship implies professional and careful treatment of data throughout all stages of a research process. A data steward aims at guaranteeing that data is appropriately treated at all stages of the research cycle (i.e., design, collection, processing, analysis, preservation, data sharing and reuse).” Jetten M. et al., 2021. Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship (1.1). Zenodo. doi.org/10.5281/zenodo.4623713

[5] Cox SJD, Gonzalez-Beltran AN, Magagna B, Marinescu M-C (2021) Ten simple rules for making a vocabulary FAIR. PLoS Comput Biol 17(6): e1009041. doi.org/10.1371/journal.pcbi.1009041

[6] “As a general rule, the heads of organisation units (professors, heads of administrative departments, heads of non-departmental teaching and research facilities, heads of staff units) will be the information owners.” Directive on “Information Security at ETH Zurich” (status of 1 August 2021), Art. 7, paragraph 1, https://rechtssammlung.sp.ethz.ch/Dokumente/203.25en.pdf

[7] ETH Factsheet “Data Protection in Research Projects” https://ethz.ch/content/dam/ethz/associates/services/organisation/Schulleitung/Generalsekretariat/dokumente_rechtsdienst/Dataprotection_Research_Final.pdf

[8] Directive on “Information Security at ETH Zurich”, status as of 1 August 2021, https://rechtssammlung.sp.ethz.ch/Dokumente/203.25en.pdf

[9] See endnote 8.

[10] See endnote 8.

[11] IT Guidelines and IT Baseline Protection Rules of ETH Zurich, Edition 2021, https://rechtssammlung.sp.ethz.ch/Dokumente/203.23en.pdf

[12] ETH Zurich Acceptable Use Policy for Information and Communications Technology (“BOT”) and Appendix, partial revision as of 1 June 2021, https://rechtssammlung.sp.ethz.ch/Dokumente/203.21en.pdf

[13] Wilkinson, Mark D et al. “The FAIR Guiding Principles for scientific data management and stewardship.” *Scientific data* vol. 3 160018. 15 Mar. 2016, doi: 10.1038/sdata.2016.18

[14] Guidelines for Research Data Management at ETH Zurich, Article 6(2), https://rechtssammlung.sp.ethz.ch/Dokumente/414.2en.pdf

[15] Guidelines for Research Data Management at ETH Zurich, Article 6(1), https://rechtssammlung.sp.ethz.ch/Dokumente/414.2en.pdf

[16] see endnote 15.

[17] Page of the SNSF on Open Research Data, https://www.snf.ch/en/dMILj9t4LNk8NwyR/topic/open-research-data, accessed 21/02/2022

[18] See “Open Access (OA) rules of the SNSF – guidelines for researchers” https://www.snf.ch/SiteCollectionDocuments/Dossiers/dos_OA_policy_grafik_e.pdf (accessed 21/02/2021).

[19] ETH Zurich’s open access policy dated 17 January 2018, https://rechtssammlung.sp.ethz.ch/Dokumente/134en.pdf

[20] ETH Zurich Guidelines on scientific integrity, Article 11(2), https://rechtssammlung.sp.ethz.ch/Dokumente/414en.pdf

[21] ETH Zurich Guidelines on scientific integrity, Article 11(3), https://rechtssammlung.sp.ethz.ch/Dokumente/414en.pdf

[22] Guidelines for Research Data Management at ETH Zurich, Article 8(1)(b), https://rechtssammlung.sp.ethz.ch/Dokumente/414.2en.pdf

[23] General implementation regulations for the Funding Regulations of the SNSF (Version 1.7.2022) Art. 2.13, https://www.snf.ch/media/en/B0SWnPsrDCRTaiCx/snsf-general-implementation-regulations-for-the-funding-regulations-e.pdf