ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ETH Library
Research Support Services

ETH Zurich
Dr. Julian Dederke
Research Data Management and Digital Curation
Weinbergstrasse 74
8092 Zurich
Switzerland

Telefon +41 44 633 86 32
julian.dederke@library.ethz.ch /
www.library.ethz.ch

Action Plan Data Stewardship ETH Zurich (2023-2024) – DSETHZ

Submission to Action Line B5.2 of swissuniversities' ORD Action Plan

Date   1 October 2022
Version   1.0
Author(s) Julian Dederke, Matthias Töwe, Fabian Schmid, Caterina Barillari, Henry Lütcke

Table of Contents

1. Introduction

Measure B5.2 of swissuniversities' national ORD action plan focuses on promoting and incentiv-ising data stewardship and hiring of data stewards and other ORD specialists within Swiss higher education institutions. This document defines the respective action plan for ETH Zurich based on the goals of swissuniversities' national ORD strategy.

2. Current situation at ETH Zurich

2.1. Institutional Setting

Research groups at ETH Zurich enjoy a high level of autonomy which is considered one of the strengths of the university. Typically, ETH Zurich's Executive Board only sets the framework conditions which must then be implemented on the level of departments, institutes or by individual research groups. This approach allows to find suitable solutions for specific research communities. However, leaving the implementation to research groups can lead to a multitude of different solutions and this is also the case for Research Data Management and Data Stewardship. At the current stage, we do not consider this as a disadvantage, but rather we plan to turn this diversity into an opportunity to explore the practical implementation of a number of different approaches to Data Stewardship within the university.

2.2. Existing Services and Identified Gaps

Services for Research Data Management (RDM) support were introduced stepwise at ETH Zurich in response to growing demand. While most services existed prior to the SNSF's announcement of mandatory Data Management Plan (DMP) submission in 2016, the SNSF's initiative increased researchers' awareness for issues of RDM. The central units ETH Library and Scientific IT Services expanded their consulting, training and service offers accordingly with additional resources.

The central service offers mainly address general issues of RDM and ORD and help early career scientists, in particular, to gain a comprehensive overview of the related questions and issues. However, implementing concepts in practice as part of a lab's or research group's routine workflows requires direct involvement of researchers and decentral RDM/ORD professionals as well as more (discipline-)specific support. This concerns data management planning and day-to-day activities in a research group, for example data acquisition, annotation and backup, as well as other issues such as data publishing or archiving. Scientific IT Services, for example, provide and support and powerful tools for active RDM. Nevertheless, it is research groups themselves that have to define how to implement and use the available tools in the context of their own research subjects and methods. The required know-how is usually available in research groups at a given point in time. However, if post-docs or doctoral students take over roles in data management for their group, frequent loss of know-how and experience must be expected due to inter-institutional and international mobility of early-career researchers. The flexibility of this approach comes at the cost of sustainability.
From the point of view of sustainability, having senior scientists or technical personnel in charge of RDM and ORD activities is a useful measure, but in reality, this will often compete with their research, teaching and other commitments, unless their role in RDM is sufficiently well defined and accounted for.

Therefore, we see a need to better define roles in RDM within departments and on the level of research groups. This is a prerequisite for further establishing RDM best practices sustainably in the university while respecting discipline and subject-specific practices. ETH Zurich

departments, institutes, and/or research groups are expected to embed these practices in their research environments. Identifying and hiring suitable candidates for a position as data steward or ORD professional at their organisational unit would strongly support this. The flow of information between central services and persons responsible for RDM in departments is a core element of this approach and it should be enhanced by peer-to-peer communication between those persons.

2.3. Policy Level and Strategy

2.3.1. ETH Domain and ETH Zurich Strategy

The ETH Domain has outlined its vision for Open Research Data[1] which also puts ORD in the context of its strategic focus area on Data Science. More concretely, the ETH Board together with the institutions of the ETH Domain has launched an ORD Program for 2021 to 2024. The institutions have committed to contributing half of the funding. The ETH Domain plans on continuing this program in 2025 to 2028.

2.3.2. ETH Zurich Guidelines

ETH Zurich for a long time had laid out basic requirements on RDM in its Research Integrity Guidelines. In recent years, it became increasingly obvious that researchers would need more precise guidelines for RDM. As a result of a comprehensive effort led by the Research Office and involving the research departments and central units of the university, two guidelines were put into force in January and July 2022, respectively. The full revision of the Integrity Guidelines[2] and the development of the new RDM Guidelines[3] were synchronised to consistently address RDM issues.

The RDM Guidelines mandate the publication of relevant research data in FAIR repositories and encourage openness. They serve as a framework for the whole university and researchers are expected to complement them with best practices based on established and accepted standards from their specific research community.

An important feedback from researchers in the discussion of the guidelines was that the departments are not automatically seen as the appropriate organisational level for prescribing and implementing best practices because of the heterogeneity in research topics and in the methods used among groups of the same department. This led to mandating reference to the appropriate scientific communities' standards instead of delegating the task of setting common standards to

---

[1] Open Research Data - Position of the ETH Domain (2020):
https://ethrat.ch/wp-content/uploads/2021/09/ORD_Position_ETH_Domain.pdf

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

[3] Guidelines for Research Data Management at ETH Zurich (RDM Guidelines):
https://rechtssammlung.sp.ethz.ch/Dokumente/414.2en.pdf

the departmental level. It also points to the fact that researchers in practice rely on research field-specific community practice and expertise considered appropriate by their peers, even though those might not be fully developed yet with respect to ORD.

As part of the RDM Guidelines, ETH Zurich itself commits to providing the required infrastructure and services, usually via its central units which are explicitly mentioned in the annex of the guidelines. This mandate will involve translating recommendations and expectations from ETH Zurich guidelines into easily comprehensible guides in written and illustrated form as well as training formats for various target groups (doctoral students, scientific staff, data stewards, etc.).

3. Data Stewardship Implementation

The demand for expertise in topics of Data Management and Open Research Data at ETH Zurich as a data-driven research institution has grown in recent years along with increasingly specific funder requirements that demand openness, transparency and reproducibility from researchers. The increase in data volumes but also the vast number of available tools for data handling necessitate the establishment and professionalisation of support roles such as ORD specialists and data stewards in all data-driven disciplines. The ORD position of the ETH Domain formulates as a goal that "all students and researchers in the ETH Domain – from Bachelor to postgraduate level – should have access to training on research data management" (ETH domain 2020, p. 8, see footnote 1). It also acknowledges supporting services as a key element of a suitable environment for openness in research. It is clear that these goals can only be achieved by recruiting and professionalising experts in ORD support roles and/or data stewardship. This action plan focuses on initialising this process and suggests the first steps to establish a service-oriented network of employees at ETH Zurich who already fulfil tasks in data stewardship or RDM support and who want to extend these efforts or orient them more toward the openness of research data. Furthermore, a crucial step in professionalising data stewardship at ETH Zurich will include the tailoring of existing consulting services, training material, and guidance documents to the target group of data stewards and other ORD professionals. In that way, it is possible to pursue a "train-the-trainer" approach.

In addition to funder requirements and disciplinary best practice principles, in summer 2022, the [ETH Zurich Executive Board](#) has adopted new Guidelines for Research Data Management (RDM Guidelines, RSETHZ 414.2). These complement the ETH Zurich Guidelines on Scientific Integrity (Integrity Guidelines, RSETHZ 414) by specifying further aspects of research data management (RDM). Due to the impetus of the RDM guidelines and the institution-wide commitment to good practice principles in RDM, a higher demand for training and expertise is to be expected in RDM-related topics. From our perspective and based on the analysis of existing services and gaps (see section 2.2 above) such a demand is best matched by establishing and professionalising data stewardship roles at ETH Zurich. Since ETH Zurich is a very large and diverse institution with many organisational units, our approach at ETH Zurich will be a data stewardship pilot with a few data stewards involved. By contrast, a fully-fledged data stewardship programme with responsible persons at all departments (i.e., similar to the data stewardship programme established at TU Delft) is currently not feasible.

While departmental structures and departmental autonomy are important at ETH Zurich, disciplinary communities do not necessarily exist within a single department, but might instead cross-cut departmental boundaries. As an example, methods from the social sciences are employed in diverse fields such as environmental, health or architectural research, to name only a few. At the same time, ETH Zurich departments often host researchers of several different disciplines. One example is the Department of Environmental Systems Science (D-USYS) that hosts researchers from the physical, chemical and biological sciences, as well as agricultural sciences and social sciences. This is why an ETH department might not be the ideal reference point for the responsibilities of a data steward. Instead, research areas are better suited as a reference point. Therefore, we aim at implementing data stewardship at ETH Zurich by working together with pilot data

stewards who can cover and address at least several research groups (within one or several departments).

The tasks of the ETH Library and the Scientific IT Services within the proposed project will be:

- Communicate the goals of the swissuniversities ORD strategy to data stewards and other employees with RDM-related tasks,

- Communicate the requirements of ETH Zurich's RDM Guidelines and (where needed) RDM requirements of research funding agencies to data stewards and other employees with RDM-related tasks,

- Facilitate networking and exchange among data stewards and other employees with RDM-related tasks,

- Provide consulting and advice as required and in line with the specific expertise available,

- Identify (together with the selected data stewards) RDM-related workflows and practices in research groups, institutes or departments that could be oriented more strongly towards the openness and reproducibility of research data,

- Support data stewards with the identification and adoption of discipline-specific metadata standards and semantic ontologies.

Based on these considerations of the institutional preconditions at ETH Zurich, the goals and implementation steps of the project will be addressed within the two following overarching "tracks". Track A will address the communication and exchange among data stewards and employees with RDM responsibilities as well as with the central service providers (mainly the ETH Library and the Scientific IT Services). Track B will address the implementation of ORD practices within research communities.

3.1. Track A: Communication and exchange among data stewards and employees with RDM responsibilities

3.1.1. Objective A1: Establishing the Data Stewardship Network (DSN) at ETH Zurich

The ETH Library will launch a Data Stewardship Network (DSN) at ETH Zurich. This network will aim at connecting data stewards, ORD professionals, and other ETH employees with RDM responsibilities. It will provide a platform for exchange about ORD practices, tools and resources. The network will include (1) ORD professionals and RDM specialists who are already employed at ETH Zurich specifically for those tasks (including, e.g., RDM specialists from the Scientific IT Services), and (2) ETH Zurich researchers (e.g., postdocs or senior scientists) with additional

RDM-related responsibilities. This approach is important since so far only few university employees have official job titles such as 'Open Science officer' or 'data steward'. By connecting both groups in the DSN, we expect to leverage the ORD professionals'/ data stewards' expertise and allow a network effect beyond scheduled trainings and central consulting. The network is intended as a bottom-up initiative that can be shaped by the data stewards and involved people with RDM responsibilities based on the needs of their supported research groups and other organisational units. ETH Library employees at the team for Research Data Management and Digital Curation together with colleagues from the Scientific IT Services will maintain the DSN, inform interested ETH Zurich members, provide material, consulting and training for the DSN members.

The network will be enhanced by additionally including project partners who already have RDM responsibilities in research groups at ETH Zurich. These people, who are currently working in ETH Zurich research groups, institutes, or departments, will be onboarded via a part-time commitment as data stewards in a pilot programme that is part of Track B of this project (see Objective B1 of this action plan further below). Additional expert knowledge can be integrated by including staff of the ETH Library (e.g. E-Publishing unit, Knowledge Management group, and library subject specialists) or the Scientific IT Services at ETH Zurich for contributions to discussions, consultations and trainings.

Bi-annual network meetings will be established, of which at least one per year is planned to take place on-site at ETH Zurich, while the second one will be offered in a more flexible online format. ETH Library and Scientific IT Services staff will initiate and coordinate those meetings, but will mainly build on contributions from the side of DSN members.

Wherever possible, the network will encourage an open exchange among its members. Since in practice, levels of knowledge and of experience both in research and in a supporting role vary widely among the group of staff addressed with this plan, peer-to-peer mentoring between experienced colleagues and more junior colleagues can be advantageous. We want to point out this option by encouraging DSN members to form "tandems" of persons with related interests, but different backgrounds and experience. This will facilitate a more profound exchange on difficult issues which are not suitable for discussion in an open network – either due to a more confidential nature or because they are just too specific for a wider group. The project management will encourage such tandems, but will not actively manage those, since the network is intended as a bottom-up-initiative.

The main tangible deliverable supporting this objective will be an exchange platform for data stewards. The ETH Library and the Scientific IT Services will facilitate networking and exchange among the involved data stewards on this platform (i.e., a Microsoft Teams Team designated as "Data Stewardship Network at ETH Zurich"), in group meetings, and at training or network events. Project partners from the central service units will also encourage Data Stewards to use the DSN for discussing and compiling discipline-specific instructions together with peers from related fields.

The ETH Library is committed to maintaining the Data Stewardship Network at ETH Zurich beyond the project period with its own resources, in collaboration with central service units at ETH

Zurich as partners. Maintaining this network does not involve financial responsibility for the employment of the DSN members though.

3.1.2. Objective A2: Alignment with policies by communication and exchange

The project at ETH Zurich will also build on the momentum from the recent adoption of the RDM Guidelines and the institution-wide commitment to good practice principles in RDM (summer 2022). On the level of communication and exchange (Track A of this project) data stewards will play an important role in the concrete implementation of the guidelines and the dissemination of relevant training material as ambassadors and multipliers in close proximity to researchers. This will enhance the central service units' efforts. Additionally, library and IT service staff will continue, together with interested data stewards in the network, to identify online training material that can be used to increase competences of RDM trainers and researchers in ETH Zurich research groups.

3.2. Track B: Implementation of ORD practices within research communities

3.2.1. Objective B1: Building up know-how on ORD topics and practices

In line with swissuniversities' ORD strategy, a key objective of the project is to foster RDM know-how and ORD expertise within research communities at ETH Zurich. This can be achieved by the commitment of current employees in ETH research groups, labs or institutes to routinely engage with RDM-related tasks. Ideally, these individuals already have responsibilities for RDM and/or data stewardship that extend beyond the data they have created themselves. Moreover, such data stewards should operate across several research groups or institutes to allow implementation of good practice principles beyond individual organisational units. To achieve this goal, ETH Library will set up a data stewardship pilot programme. We have collected several letters of intent from project partners who are willing to commit work time to the goals and issues addressed in swissuniversities' ORD programme (see attachments to this project proposal). The selected project partners will be part of the Data Stewardship pilot programme. The efforts by the selected data stewards in the pilot programme will include, for example:

- Organisation of RDM-centred workshops and courses for several research groups, and potentially the entire Data Stewardship Network at ETH Zurich;

- Setting up and implementing data management and inventory management system, such as openBIS in additional research groups at ETH Zurich;

- Enrichment of unpublished datasets on department storage with suitable metadata;

- Exchange between different disciplines on data management and data publication practices;

- Extension of RDM trainings to students by selected partners, in line with the goal of the ETH domain to provide RDM training from Bachelor level onwards (ETH domain 2020, p. 8, see footnote 1);

The pilot programme will be evaluated internally towards the end of the project period. Presenting the results of this pilot can provide a model for ETH Zurich research groups, institutes or departments who might want to finance data stewards in the long run. As a main deliverable from addressing this objective we will produce an internal evaluation report of the data stewardship pilot programme in collaboration with all project partners.

3.2.2. Objective B2: Train-the-trainer approach

Central RDM support from the Library (together with partners, for example, the Scientific IT Services) will apply a train-the-trainer approach, transferring ORD know-how to data stewards as well as researchers interested in increasing their ORD knowledge and competences. These persons involved in data stewardship can be considered as multipliers who offer direct support in research groups and teach scientists how to implement ORD principles in their everyday work. This can be achieved by tailoring existing consulting services, training material, and guidance documents to this target group. Targeting data stewards as participants of trainings can be facilitated by keeping in constant exchange with the participating data stewards during the project period. Recruited RDM professionals who take part in the data stewardship pilot (see Objective B1 above) can be co-opted into this train-the-trainer approach as trainers. In this way, the data stewards in the pilot programme will take a crucial position between central service units (library, IT services etc.) and the researchers.
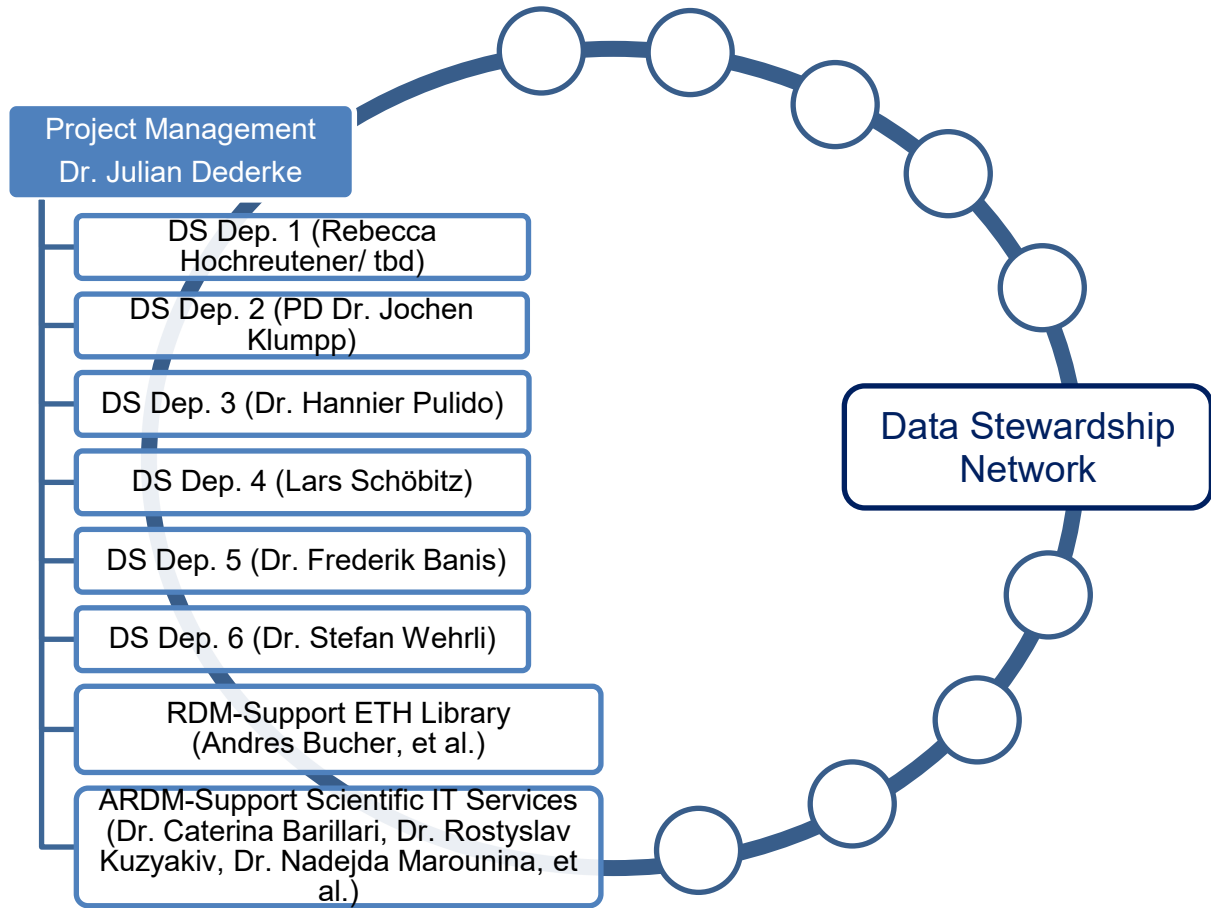
3.2.3. Objective B3: Alignment with policies by building up know-how

Just like in Track A of this project, Track B will also take up the impetus from the recent adoption in 2022 of the RDM Guidelines and the institution-wide commitment to good practice principles in RDM. The aim is to translate recommendations and expectations from ETH Zurich guidelines into easily comprehensible guides in written and illustrated form for wider distribution. Since a higher demand for training and expertise in RDM-related topics is to be expected, such written and illustrated guides can leverage the RDM expertise at the central service units (ETH Library and Scientific IT Services) for a larger audience.

4. Financial Planning

[The section on financial planning was removed from this version.]

5. Organisational Chart

Project Management
Dr. Julian Dederke

DS Dep. 1 (Rebecca Hochreutener/ tbd)

DS Dep. 2 (PD Dr. Jochen Klumpp)

DS Dep. 3 (Dr. Hannier Pulido)

DS Dep. 4 (Lars Schöbitz)

DS Dep. 5 (Dr. Frederik Banis)

DS Dep. 6 (Dr. Stefan Wehrli)

RDM-Support ETH Library (Andres Bucher, et al.)

ARDM-Support Scientific IT Services (Dr. Caterina Barillari, Dr. Rostyslav Kuzyakiv, Dr. Nadejda Marounina, et al.)

Data Stewardship Network

6. Annexes

Letters of Intent from pilot partners.