

Action plan for the swissuniversities call for projects “Action line B5.2: Data Stewardship”

Project title: Data Stewardship at USI

Abstract

With the publication, in 2021, of the Swiss National Open Research Data Strategy and in accordance with a growing attention dedicated to the theme of Research Data Management, USI is moving towards the formulation and consolidation of principles and practices aimed at a correct management of the databases that are generated and used by its researchers at all levels and in all areas of study. The overarching goal of this effort is becoming an active actor in the ORD movement in the Swiss academic landscape.

As detailed by the Swiss ORD Strategy and the corresponding action plan, an integral contribution for adopting FAIR practices and reaching ORD-related goals comes from the promotion of the career of Data Stewards, the consolidation of their proficiencies and roles, and the establishment of competences centers to foster the dialogue and cooperation among Swiss HEIs.

This project encompasses the present situation of USI in terms of Research Data Management and Open Research Data, the gaps that are intended to be overcome and the prospective efforts put in place to close said gaps. The project presents the context in which the new Data Steward will operate, the synergies that will likely be shaped together with existing USI services and departments and the goals expected to be reached in the medium-long run.

Current situation of RDM practices at USI and institutional vision on ORD

USI community firmly believes in Open Science and Open Research Data, recognising the strategic importance of these two concepts and striving to have an active role in the novel landscape that has been taking shape in Switzerland and internationally in the last years. Notwithstanding, ORD is still poorly assessed by USI academic community and best RDM practices are far from being effectively and efficiently implemented. A recent survey spread among the academic community of USI revealed that only 15% of researchers are “very familiar” with ORD and knowledge about FAIR principles is also moderate among researchers as only 8% have “extensive knowledge”. Familiarity with ORD and FAIR principles is reflected on data sharing practices as only 27% of researchers share “always” their data once a project is over and 30% do it “only if requested”. In general, several criticalities were identified in the (lack of) adoption of good research data management

practices. While there is a strong heterogeneity among faculties and among research roles regarding ORD competences and skills, one can observe that ORD-related needs are relatively homogeneous among USI research personnel considering that support is sought by the great majority of researchers on data sharing, best practices about data storage and the search for appropriate data repositories.

It must be recognized that the necessity to provide a DMP as requested by the SNSF in order to obtain research funds has helped raising researchers' awareness on the topic and is contributing to the realization that RDM has become a crucial element in the administration of research projects and revolves around concepts and practices that must be accounted for during the entire data lifecycle.

Another important point lies in the absence of a structured training program dedicated to researchers at all level of seniority in which ORD and RDM themes are tackled. Granting appropriate skills through training is of utmost importance for the implementation of best RDM practices and an effective adoption of FAIR principles. The provision to students, researchers and research support staff with the necessary theoretical competences and practical tools needed for an effective handling of research data and for the creation of a dedicated mindset is an inevitable task for a modern university.

All these considerations have led the management of USI to tackle the issue and reflect on the necessary future actions to keep the pace with Swiss educational institutions showing far more advanced profiles in this sense.

In a landscape where more and more institutions are operating toward the establishment of a Data Stewardship service in order to implement effective and sustainable research processes and disseminate RDM practices among researchers, USI has taken on the task of promoting long-term initiatives on ORD and RDM aimed at the professional growth and specialization of its researchers, and as part of the strategic development of the institution. The ultimate tasks are to comply with good practices in scientific research recognized at national and international level, foster an active participation in the ORD movement, support and advise researchers on research data management practices, implement the Swiss National ORD Strategy and participate to national and international ORD initiatives. Thus, USI management is strongly committed to develop appropriate measures in terms of ORD, and the creation of a Data Stewardship service is the most important of these.

USI firmly intends to support the creation and the development of the right framework in which the Data Stewardship service will be established and given the opportunity to grow, granting all the necessary attention and resources to be in line with the times and to keep the pace with the major Swiss HEIs.

Data Stewardship is seen as a strategic asset for the growth of USI and support will be granted in the future, beyond the lifespan of the financial aid of swissuniversities. USI support will be of financial, technical and organizational nature and will permit a strategic development of the service for the coming years.

Main goals of Data Stewardship service at USI

The principal objective of this project is the professionalization of a dedicated person whose tasks will be setting the development of a USI competence center for RDM and ORD and fostering the cooperation with other Swiss institutions through the participation to existing and future networks and initiatives.

The overarching goals that USI intends to pursue with this project can be listed as follows:

- Definition of Data Stewardship tasks tailored to the specific needs that USI research landscape demonstrates, accounting for the diversity in the disciplinary landscape of the university.
- Professionalization of the Data Steward, ideally through the attendance to a Data Stewardship training program.
- Implementation of measures aimed at monitoring and curating research databases at USI in strict collaboration with faculties' representatives. In particular, granting support to researchers in the negotiation phase for the purchase of research data.
- Creation of a register of available databases and data providers.
- Sharing of the (newly) acquired databases to the generality of institutional users (students, PhDs and professors) according to legal and contractual framework.
- Drafting of an institutional Research Data Management and Open Research Data overarching strategy, accompanied by ORD and RDM policies intended as concrete measures that implement the strategy and aimed at researchers and research-support staff.
- Creation of an informative event dedicated to researchers with the presentation of the new Data Stewardship service.
- Creation of dedicated web pages to be integrated in the USI website and intended to be the online showcase of the service.
- Inclusion of the Data Steward in the existing ORD-related networks through the participation of national and international initiatives (meetings, workshops, conferences).
- Design of a training program dedicated to all the levels of research career, with special attention to the sets of themes that ORD and RDM encompass (e.g. data sharing, data security, data storage, DMP writing) and to the disciplinary differences.

These objectives were defined, and will be pursued, according to the results of the above mentioned survey in which competences and needs of USI researchers were assessed. Important information was retrieved, especially for what concerns the heterogeneity among the disciplines and among research groups in the same faculty. The feel is that different research communities find themselves at different

places concerning the necessities related to data management, the realization of the importance of such themes, and the participation to the expanding national and international ORD landscape. Thus, it appears of utmost importance that adequate data stewardship mechanisms are developed to tackle a large variety of visions and situations. The definition of Data Steward's competences and career growth path goes hand in hand with the definition of the services and infrastructures that USI aims at providing to the faculties. In this sense, USI intends to develop a research data management service that encompasses the different needs expressed by the academic staff. Also, the definition of Data Steward's skillset and tasks has to point to and be coherent with an effective integration into the national and international Open Research Data landscape that has been shaping in the last years, with the adoption of existing best-practices and standards and the contribution to the formulation of new ones.

Development of Data Stewardship operational tasks at USI

Research Data Management is a multifaceted set of processes involving the collection, treatment, annotation, preservation, and diffusion of research data. The Data Steward will be the first point of contact for researchers who need to handle research data and rationalize their life-cycle. The Data Steward will develop more advanced sets of skills while operating side-by-side with researchers and through the collaboration with other services and will facilitate data handling solutions and procedures, eventually setting new standards and structures which will be flexible enough to be adapted to the heterogeneous needs of the different research communities in USI. The Data Steward at USI will have the overarching task of promoting the application of FAIR principles and educating about good RDM practices. In the long-term it is expected that within the Data Stewardship service subject-matter expertise will be developed and made available to researcher personnel.

The main tasks the Data Stewardship service will be given in the medium-long term can be listed as follows:

1. Collaborate on the definition of an institutional RDM and ORD strategy. USI aims at delineating an institutional Research Data Management and Open Research Data strategy in accordance to the Swiss National strategy. In this sense, the main objectives for the Data Stewardship service are:
 - collaborating with USI management in the drafting and implementation of the strategy, coherently with the measures outlined in the Swiss National ORD strategy;
 - collaborating with USI management in defining institutional policies for RDM and ORD, as is already the case for several Swiss and international institutions. Said policies will contain formal rules to actively implement the long-term strategy, and will be adapted

according to the present regulations governing USI. These policies will govern the necessary research infrastructures accompanying academic activity towards the adoption of RDM and ORD best practices.

Ideally, the strategy will aim at making USI an active member of the ORD movement and will clearly position USI in the ORD landscape. For this, the strategy will prioritize the following points:

- alignment to Swiss standards and procedures;
- making sure that research data are collected, organized, and maintained according to FAIR principles;
- formulation and adoption of best practices with a cross-disciplinary approach;
- support the participation of USI researchers to national and international ORD programs and calls for projects;
- participation to national and international ORD initiatives developed by the main actors involved in the theme.

The prospective term for developing a fully-rounded institutional strategy on RDM and ORD together with operational policies is the **end of 2023**.

2. Access to and management of USI research databases. Good data management practices entails taking care of research databases. The Data Steward will monitor and administer databases owned by USI, encouraging their dissemination and use among USI scholars. In this sense, it will be crucial to facilitate opportunities of data reuse between USI researchers in a multi-disciplinary framework, fostering future research and collaborations. The aim is to grant the sharing of the acquired databases for the greatest possible number of users of the University and in the least binding way possible. A potential outcome will be the creation of an institutional register of available databases that can be shared among USI researchers. The Data Steward also supports the research community in the process of obtaining, organizing and managing research data in the context of research projects and in a long-term perspective.

The Data Steward will get in touch with the representatives of all the faculties in order to comprehend the needs and expectations of the researchers on the management (acquisition, storage, long-term preservation, sharing, anonymisation) of databases. This way it will also be possible to start the monitoring of existing research databases and the dissemination of all the necessary information to the whole research community in order to foster potential inter- and intra-disciplinary collaborations.

The prospective term for generating the necessary network with researchers and creating an infrastructure in which research databases are listed and detailed is the **end of 2023**.

3. Creation of a structured website for Data Stewardship. The online presence of the new Data Stewardship service will be granted through the design of

dedicated web pages. These will report all relevant information concerning USI Data Stewardship-related activities and initiatives, the list of institutional research databases, a guide to support the identification of the most appropriate data repository as advised by SNSF, ORD and DMP guidelines and policies, relevant calls for projects, dedicated national and international events, the program of training sessions and Q&A.

The prospective term for having fully operational webpages dedicated to Data Stewardship service at USI is the **end of 2023**.

4. Data Management Plan (DMP) support. The Swiss National Science Foundation (SNSF) and other funding agencies are embracing Open Research Data principles and practices more and more. In the context of Data Stewardship, USI will organize assistance and advice for researchers that aim at applying for a grant and are requested to present a DMP. In particular, guidelines and examples will be published in order to help researchers in this sense, with a particular attention to SNSF requirements and DMP format.
This task can be immediately and continuously performed according to SNSF deadlines.
5. Researchers training and support on ORD and RDM. One of the main tasks of USI data stewardship program will be to facilitate the transition towards a FAIR-data mentality among academics of all level and disciplines. We feel that the most effective way to do so is organizing a well-structured and tailor-made set of training cycles primarily aimed at young researchers who have to be considered the next generations of scholars operating within and outside academia. Young researchers will be confronted with ORD issues more systematically and this will happen across all scientific fields. USI is firmly convinced that RDM and ORD competences are an integral component of soft skills needed by PhD graduates (similar to ethics, grant acquisition, academic teaching) and will work to grant such competences. Part of USI strategy in this sense is the definition of the expertise, the needs and the expectations in terms of ORD coming from the academic body. Ideally, the training cycles will have to be sensitive about differences on the ground considering the strong heterogeneities among USI faculties in terms of disciplinary differences as well as acquired competences, capabilities, and technical skills on RDM and ORD. It will then be crucial to create the right balance between general and discipline-specific skills. In this sense the Data Manager will work alongside professors and group leaders who are already familiar with ORD and RDM and will grant the right approach to design the training.
The first step in the definition of a training program will be to design a potential offer of generalist and discipline-specific courses covering the set of arguments in the RDM and ORD landscape. Faculty representatives and the persons in charge for doctoral schools will be ideally involved in the definition

of the program considering its potential integration in the curriculum of PhD students.

The second step will be a gradual activation of the training program which may take the form of one-day or multiple-day workshops, dedicated classes in summer schools such as the Summer School in Social Science Methods and collaborations with third-parties such as FORS.

The prospective term to design a general offer of training courses is **mid-2024** with the expectation to launch the first training sessions at the beginning of the 2024-2025 academic year.

Internal synergies and relationships with service providers

The USI Data Steward will be the primary point of connection between researchers and the large plethora of services and infrastructures providers (inside and outside the organization) that allow the formers to successfully obtain, manage and elaborate their data and, as a consequence, to effectively and efficiently pursue data-driven research. In particular, the Data Steward will act as the trait d'union between researchers and 1) USI services which can be related, directly or indirectly, to the context of Open Science and more specifically Open Research Data, 2) data providers, and 3) services and infrastructures providers related to data management and data processing.

In terms of internal synergies between Data Stewardship and other USI departments, collaborations will be established that will foster the provision of transversal skills for an effective effort dedicated to data management tasks. The following is expected:

- Data Stewardship will be integrated into USI Research and Transfer Service whose main objective is to provide support to researchers with regard to the acquisition and management of research grants. This integration will help the new resource to operate in a context in which relationships with researchers are already consolidated, the Data Steward will be able to benefit from the know-how and the collaborations that have consolidated over the years. In this sense, Data Steward's tasks will include support and advice in drafting a Data Management Plan as well as support and advice for the acquisition, management, sharing and reuse of project-related data.
- The Data Steward will benefit from the competences of the University Library on Open Access so to define common practices related to Open Science. In this context, cooperative training sessions will be organized and dedicated to researchers in the form of seminars, workshops or academic lectures having the overarching goal to introduce and spread OS practices so to form a mindset oriented towards Open Access, Open Research Data and the adoption of FAIR as well as Research Data Management best practices.

- For the very same nature of the tasks that a Data Steward has to pursue, a strict collaboration with the IT services will be established in order to draft the standards and policies that will be implemented in the field of data security, high performance computing, storage (including backup practices) and long-term preservation of data, data transfer and data destruction. As far as data security is concerned, USI is particularly solicited and particular attention is required from research teams dealing with sensitive data and working in the domain of life sciences and social sciences (Biomedicine, Communication and Economics).
- In order to tackle ethical, legal and contractual aspects characterizing Research Data Management, we envision the cooperation between the Data Steward and the Legal Service as well as the USI Ethical Committee. Also, the Data Steward will benefit from the Competence Center in Digital Law born from the collaboration between the USI, UNIBAS, UNINE, UNIGE, and the Conference of Swiss Libraries (CBU-KUB). This will allow the assimilation of know-how and the adoption of best-practices regarding data protection and copyright issues (e.g. creative commons).
- Lastly, there will be a constant dialogue between the Data Steward and the faculties. USI has a varied and multidisciplinary landscape of faculties and this heterogeneity of disciplines poses formidable challenges. Furthermore, the level of maturity in data management differs greatly among USI faculties and addressing these differences in terms of capabilities, practices and habits is essential. Thus, it appears crucial for the Data Steward to develop a set of skills that will help accounting for disciplinary heterogeneity and define specific data management and curation practices aimed at the different communities. This will be possible thanks to the relationship that the Data Steward will have with the various research groups, both within USI through the faculties and outside USI through the networks that collect the experiences of Swiss and international research groups (e.g. DaSCH, DARIAH and CLARIN) to which many USI professors belong.

As far as the relationship between Data Steward and data providers, it is well known that one of the central tasks of Data Stewardship is the monitoring of project-related data acquisitions as well as the management of the bulk of research data an institution holds. USI has a large number of research groups involved in data-intensive research projects, often relying on datasets provided by commercial vendors. In this sense, one example is represented by the Institute of Finance whose budget dedicated to data acquisition represents a substantial share of the USI budget. The experience accumulated in data purchase in the context of these institutes will be important to clearly define coherent standards to be adopted for the whole university. In this sense, the Data Steward will benefit from a well-established know-how and will be able to reinforce the set of competences concerning the entire process of data purchase (e.g. evaluation of the researchers' requests, and bargaining for the terms and conditions of data usage and distribution with the goal

to benefit not only the single researcher but the plethora of potentially interested institutional users).

Research data archiving and managing platforms, online data repositories as well as High Performance Computing providers represent only few examples of services and infrastructures which researchers interact with in their everyday practices of data management. For researchers, these services are pivotal to pursue high quality data-driven research and to share and promote the results of their work in the form of open data.

At the moment, USI does not have a proper, institutional infrastructure aimed at managing and operating with research data, hence researchers need to turn to third parties.

This lack of infrastructure leads to the definition of lines of action that place the Data Steward as the main figure to implement concrete solutions that grant, on the one hand, the centralization of the information relating to all the databases in use by USI researchers and, on the other, the support to researchers to identify external services, such as repositories where data can be stored in accordance with FAIR principles.

To conclude, it is important to stress that the role of Data Steward will not be merely as a passive contact person for researchers, on the contrary s/he will have a proactive role in terms of providing support and advice on operational, administrative and disciplinary level.