

DATA MANAGEMENT CAMPUS

Open learning modules for researchers

Fabian Felder

*Lib4RI - Library for the
Research Institutes within the
ETH Domain: Eawag, Empa, PSI
& WSL
Switzerland
fabian.felder@lib4ri.ch
<https://orcid.org/0000-0002-5076-1487>*

Henry Lütcke

*Scientific IT Services, ETH
Zurich
Switzerland
henry.lutcke@id.ethz.ch
<https://orcid.org/0000-0001-7225-4374>*

**Angela Montano
Garcia**

*EPFL Library, EPFL
Switzerland
angela.montanogarcia@epfl.ch
<https://orcid.org/0009-0000-8079-336X>*

Moushumi Ulrich-Nath

*Lib4RI - Library for the
Research Institutes within the
ETH Domain: Eawag, Empa, PSI
& WSL
Switzerland
moushumi.ulrich-
nath@lib4ri.ch
<https://orcid.org/0000-0002-2682-760X>*

Francesco Varrato

*EPFL Library, EPFL
Switzerland
francesco.varrato@epfl.ch
<https://orcid.org/0000-0002-0983-0831>*

Abstract – Open Research Data (ORD) and effective Research Data Management (RDM) are crucial components of Open Science, promoting scientific transparency, accessibility, and reproducibility. However, researchers often encounter barriers to adopting ORD and RDM practices, including unclear benefits, fear of data misuse, or lack of accessible training. The Data Management Campus, an initiative of the ETH Domain (Swiss Federal Institutes of Technology), addresses these challenges through learning modules tailored to researchers' needs and an Open Educational Resources (OER) approach that caters to trainers.

Grounded in comprehensive surveys, stakeholder consultations, and expert input, the modules are to be delivered as centralized, accurate, and practical online learning sequences (modules) on the Swiss MOOC Service, complemented by the underlying OER material via an associated Zenodo community. This paper outlines the Campus's development process, including needs assessment, content creation, platforms selection, and challenges in researcher engagement and long-term sustainability.

Keywords – open research data, eLearning, researchers, research data management, open educational resources.

I. INTRODUCTION

Funding agencies, governments, and research institutions are increasingly adopting Open Science policies to promote transparency, accessibility, and collaboration in research. At the core of these practices lies Open Research Data (ORD), enabled by Research Data Management (RDM) - the process of managing storage, access and preservation of data and code created or collected in the course of research [1]. Despite their importance, the adoption of ORD and RDM good practices faces several barriers, including an overwhelming array of data sharing tools, unclear perceived benefits of data dissemination, concerns about data misuse, and fear of exposing methodological errors [2].

To educate researchers about these barriers and how to deal with them, we have developed the Data Management Campus, a funded initiative that converges a learning



management system with an Open Educational Resources (OER) community dedicated to learning material on ORD and RDM. This initiative aims to provide a unified access point for both learners and educators, highlighting the benefits of ORD and RDM, and offer tailored, accurate information for ETH Domain researchers. The Data Management Campus consists of a portfolio of modules hosted on Swiss MOOC service [3], allowing efficient use and updating of educational resources through the ETH Domain OER for Data Management Zenodo community [4].

II. IDENTIFYING RESEARCHER NEEDS

An important aspect of the design of the Data Management Campus is the integration of researchers' input throughout its development process, ensuring that the outcomes are both engaging and valuable to them.

Topics for the learning modules were initially determined based on responses from researchers to a Swiss 2023 survey on active data management practices, as well as consultations with researchers and expert RDM specialists. The researchers included MSc and PhD students, postdocs, and group leaders [5].

The 2023 survey assessed current ORD and RDM practices of researchers at various institutes, including ETH Zurich, EPFL, and Eawag. Out of the 1'436 responses received from researchers in ETH Domain institutes, 767 responses (53%) were complete and addressed all questions.

To supplement insights from this survey, ETH Zurich and EPFL also consulted researchers through sounding boards (31 participants) and specific workshops (6 participants) to better understand challenges and needs. Finally, 14 RDM experts from Scientific IT, libraries, and other RDM-related services were interviewed to further identify needs and knowledge gaps.

This approach enabled the identification researchers' needs categorized as follow:

- Perceived needs – *I know what I want or need to know.*
- Unperceived needs – *I don't know what I don't know.*
- Misperceived – *I think I know something I don't.*
- Emergent needs – *With new information, I realize I want or need to learn something else.*

The results suggest a mismatch between the demand for ORD and RDM training and the current level of knowledge in data management among researchers, indicating a need for further skill development. 64% of survey respondents felt no need to consult any services on data management topics. Perceived needs included topics such as writing Data Management Plans and identifying data sharing platforms (other than commercial cloud storage options like Google Drive or Dropbox) that comply with institutional data protection rules. Unperceived and misperceived needs included understanding metadata

standards and open file formats. For example, 33% of respondents did not know what metadata standards were. In addition, 48% of researchers claimed to use open file formats for data, yet most researchers continue to publish in proprietary or non-open formats (e.g., XLSX, PDF, ZIP) in popular repositories such as Zenodo or Dryad. Emergent needs also included information on different data repositories.

Through this process – combining survey responses, expert insights, and consultations, ten topics were identified that address the learning needs of researchers in the ETH Domain. Many of these topics are already covered in different institutional RDM courses. The Data Management Campus therefore provides a centralized access point for learning sequences and related resources.

The ten learning modules' topics of the Data Management Campus are:

1. Open Science and FAIR Data Principles
2. Data Organisation and Management
3. Data Documentation and Metadata
4. Data Storage, Backup, and Versioning
5. Data Publishing and Long-Term Preservation
6. Code Management and Reproducibility
7. Sensitive Data Treatment
8. Writing Useful Data Management Plans
9. Licenses
10. Open File Formats and Open-Source Code

The first module to be released will focus on "Data Publishing and Long-Term Preservation". The order of release was determined based on perceived importance. Note that the listed modules are topics rather than titles.

By providing these comprehensive, user-friendly learning resources and fostering a related community, we aim to support researchers in managing the challenges of ORD and RDM adoption, thereby encouraging an open and collaborative research environment within the ETH Domain and beyond.

III. IDENTIFYING ONLINE PLATFORMS

In order to engage with two target groups, researchers and RDM specialist educators, two platforms were required. These platforms included a Learning Management System (LMS) to deliver educational content to learners and an Open Educational Resource (OER) platform to provide and update openly licensed educational resources for those teaching the respective topics. Platform selection had to reflect key values of Open Science, including sustainable operations, easy accessibility, user-friendliness, hosting and security, open format support, and visibility. The selection process involved library services from EPFL, ETH Zurich, and Lib4RI, as well as the Scientific IT Services of ETH Zurich, and the Learning and Educational Technologies group from ETH Zurich. Using a semi-structured selection process that combined prior knowledge of well-known platforms with information gathered from websites and



documentation materials, Swiss MOOC Service was chosen as the LMS and Zenodo as the OER platform.

Swiss MOOC Service was selected due to its no-cost access through existing institutional contracts, neutral branding (no institutional logo), ability to publish content under a creative commons license (CC-BY), self-enrolment for learners, and easy access for educators with an ETH Domain login. Moreover, it is based on EdX, a popular and well documented open-source online LMS, that operates by leveraging on open formats. Combined with its established presence within the ETH Domain, this made it a practical solution [3].

Zenodo, widely adopted by researchers across ETH Domain institutions, supports the creation of Communities. A Zenodo Community was created where curators can assess submissions of learning content related to ORD and RDM topics [4]. Anyone with a Zenodo account can contribute to a Community's works. Features such as DOI dissemination, adapted publishing metadata, and simplified licensing for open publication make it easy to manage. Moreover, its open and FAIR-compliant nature aligns perfectly with the project's criteria for Open Educational Resources (OER).

Swiss MOOC Service and Zenodo together provide complementary, yet interconnected access points for learners and educators, fostering long-term engagement with ORD and RDM practices. Beyond the cross-linking between these platforms, a dedicated showcase and unique access point exists [6]. This web portal allows both target groups to easily search for learning modules, organized as pedagogical sequences, along with related materials for enhanced accessibility.

IV. DEVELOPMENT OF MODULES

Library and Scientific IT services from EPFL, ETH Zurich, and Lib4RI have contributed their current training materials on RDM and ORD for module development. These materials were categorized by relevance to each learning module being created. Using these materials, we developed a storyboard for each module to guide Swiss MOOC module development with relevant information, resource links, and interactive exercises.

Researchers are being integrated into module development through two key processes: providing educational materials and offering feedback on modules. Surveys have been conducted to identify potential contributors of relevant materials from the research community. A recruitment campaign will gather feedback on each module through questionnaires and focus groups for in-depth discussions on improvements.

Moreover, given a lack of demand for ORD and RDM consultation by researchers (64%), as indicated by the 2023 survey, the Data Management Campus highlights various benefits and integrates tools to both motivate and facilitate application of good ORD and RDM practices by researchers.

Still, some challenges and key questions must be considered during the design process. For instance:

- How to effectively target researchers from different institutes, disciplines, and levels of prior RDM knowledge?
- How to develop content that is specific enough to be actionable for senior research scientists, yet generic enough to be inclusive for early-career researchers?
- How to deliver Open Educational Resources in a way that maximizes reuse and updates by RDM specialists?

By systematically integrating researchers' feedback and material contributions throughout its development, the Data Management Campus aims to enhance good ORD and RDM practices, fostering a culture of Open Science within and beyond the ETH Domain. In accordance with this aim, we developed a Feedback Framework, an evolving plan and documentation of researcher integration into the development of the Data Management Campus, that will also be released at the end of the project.

V. RELEASE OF MODULES AND CHALLENGES

The production of the first released module is expected to refine the team's workflows and establish a benchmark for streamlining future module releases. The first module on Data Publication and Long-Term Preservation is scheduled for release by late January or early February 2025, coinciding with events such as Love Data Week, Open Data Day, and Open Education Week. The remaining modules will be released throughout 2025. Post-release, a key challenge will be motivating researchers to engage with the materials. To address this:

- Modules were designed to be concise (less than 30 minutes) with flexible topic selection based on researchers' interests and needs.
- Researcher feedback is being integrated to ensure modules are useful and engaging.

A central question that arises is why researchers would choose to learn about RDM with Data Management Campus rather than searching on the internet or asking a generative AI chat bot? The solution likely lies in our ability to:

- Build a community-centred platform with an active Zenodo presence.
- Provide up-to-date and accurate information specific to ETH Domain researchers.
- Offer actionable RDM suggestions and tools that researchers can integrate into their existing workflows (as opposed to information overload).

Another post-release challenge is sustainability. While a dedicated team is currently active in developing and releasing the learning modules, several post-project concerns require further consideration:



- Maintenance and updating of materials in Swiss MOOC.
- Continued curation of the Zenodo community.

Current plans involve libraries and IT services (EPFL, ETH Zurich, Lib4RI) continuing administrative and curation duties at reduced capacity. The potential support from emerging data stewardship networks in maintaining the Data Management Campus is to be explored.

VI. CONCLUSION

The Data Management Campus aims to facilitate the integration of good ORD and RDM practices into researchers' daily routines. To facilitate this, we have integrated researcher input throughout the process, from identifying needs to designing engaging and useful content.

The Data Management Campus serves as a centralized access point for information delivery and resources through the interlinked Swiss MOOC learning platform and Zenodo community for resource sharing. It provides a coordinated platform for both learners and educators.

While challenges persist in designing courses that integrate diverse perspectives and ensuring the long-term platform maintenance, the Data Management Campus establishes a solid, shared foundation for advancing good RDM and ORD practices and, ultimately, Open Science within the ETH Domain.

ACKNOWLEDGMENTS

We gratefully acknowledge funding for this project from the ETH Board, the Boards of the participating institutions, and from their directly involved units as part of the ORD Program of the ETH Domain, Measure 3.

CONTRIBUTIONS

Fabian Felder: conceptualization, funding acquisition, project administration, writing – review & editing

Henry Lütcke: conceptualization, investigation, writing – review & editing

Angela Montano Garcia: conceptualization, investigation, writing – review & editing

Moushumi Ulrich-Nath: project administration, writing – original draft, writing – review & editing

Francesco Varrato: conceptualization, investigation, writing – review & editing

REFERENCES

- [1] Australian Research Data Commons, "Research Data Management Terminology," [Online]. Available: <https://vocabs.ardc.edu.au/repository/api/lda/codata/codata-research-data-management-terminology/v001/resource?uri=https://terms.codata.org/rdmt/research-data-management>. [Accessed: Dec. 6, 2024].
- [2] M. G. Bertram, J. Sundin, D.G. Roche, A. Sánchez-Tójar, E.S.J. Thoré, and T. Brodin, "Open science," *Current Biology*, vol. 33, no. 15, pp. R792-R797, 2023. doi: <https://doi.org/10.1016/j.cub.2023.05.036>.
- [3] Swiss MOOC Service, "Swiss MOOC Platform," [Online]. Available: <https://www.swissmooc.ch/>. [Accessed: Dec. 6, 2024].
- [4] ETH Zurich, "ETH Domain OER and Research Data Management" Community on Zenodo, [Online]. Available: <https://zenodo.org/communities/eth-domain-oer-rdm/>. [Accessed: Dec. 6, 2024].
- [5] F. Varrato, F. Felder, F. Schmid, et al., "Quantitative Assessment of Research Data Management Practices - 2023 (Version v1) [Data set]," Zenodo, 2024. <https://doi.org/10.5281/zenodo.13836948>.
- [6] Swiss Data Science Center, "Open Research Data Portal Switzerland," [Online]. Available: <https://open-research-data-portal.ch/>. [Accessed: Dec. 6, 2024].

