DATA STEWARDS AT THE UNIVERSITY OF LAUSANNE

Which challenges and how to address them

Carmen Jambé
University of Lausanne
Switzerland
carmen.jambe@unil.ch
https://orcid.org/0000-0001-5399-3948

Marielle Guirlet
University of Lausanne
Switzerland
marielle.guirlet@unil.ch
https://orcid.org/0000-0002-0184-987X

Gérard Bagnoud
University of Lausanne
Switzerland
gerard.bagnoud@unil.ch
https://orcid.org/0000-0002-4403-7919

***Abstract* – In 2021, the University of Lausanne (UNIL) allocated extraordinary funds to enhance support for research data management (RDM) and open research data (ORD). This led to the creation of a network of data stewards, supervised and coordinated by the "Research data" sector within the Information Resources and Archives Service (UNIRIS) of the University.**

The establishment of this network has raised several interconnected challenges: 1) professionalisation; 2) recognition; 3) stabilisation; 4) reinforcement; and finally, 5) ORD strategy and resources.

In this article, we discuss how UNIRIS addresses these challenges, notably by presenting the UNIL Certificate of Advanced Studies (CAS) in Data Stewardship launched in October 2024.

***Keywords* – research data management, open research data, data steward, data stewardship, data stewardship professionalization.**

I. INTRODUCTION

Since 2015, the University of Lausanne (UNIL) has actively engaged in the development and promotion of research data management (RDM) and open research data (ORD). This effort was initiated following a large pioneering survey that assessed the RDM and ORD expectations, needs, and practices of its research community [1]. The findings underscored the necessity of implementing and developing a multifaceted approach encompassing strategies, policies, infrastructure, best practices, training, and dedicated RDM support services. In response, UNIL established the "Research Data" sector within its UNIRIS service, tasked with leading the institution's RDM and ORD initiatives.

By 2019, UNIL intensified its commitment to open science (OS) and, more specifically, to RDM and ORD. The creation of the Division of Computing and Research Support (DCSR), attached to UNIL IT Service, marked a pivotal development. This division was tasked with providing computational and storage resources, as well as cross-disciplinary expertise across faculties and institutes. Its mission includes high-performance computing support, web development, and database management. Further solidifying this trajectory, UNIL adopted Directive 4.5 on research data processing and management in June 2019. This directive established regulations for the processing, storage, archiving, and preservation of research data by defining management rules, delineating stakeholder roles and responsibilities, and outlining pricing principles for the use of DCSR resources. In December 2019, the university also approved an OS strategy and its accompanying action plan, structured around two key pillars: Open Access and ORD[1].

Establishing the UNIL data stewards network

In 2021, UNIL allocated extraordinary temporary funds (for 2022–2024) to strengthen support for RDM and ORD. This funding enabled the establishment of a data stewards' network – supervised and coordinated by the "Research Data" sector (UNIRIS) – and a research engineers' network – managed by the DCSR. The allocation of resources included two full-time equivalents (FTEs) for data stewards — approximately 0.3 FTE per faculty across UNIL's seven faculties — and seven FTEs for research engineers. These two

[1] The update of this OS strategy, with the inclusion of two additional pillars – citizen science and source codes/software – was validated by UNIL Rectorate in December 2023.

SRDD 2024: Swiss Research Data Day, Chur (Switzerland).

DOI: 10.nnnn/nnnnnnn.nnnnnnn

networks were designed to complement one another, offering specialized expertise and distributed support. While their primary focus is on local faculty-level support, a portion of their workload is dedicated to institution-wide initiatives.

Despite these advances, the establishment of the data stewards' network has introduced several interconnected challenges. These include ensuring adequate and equitable resource distribution between the faculties, addressing the limitations imposed by temporary funding, fostering collaboration between decentralized and centralized positions, and developing sustainable mechanisms to support the long-term integration of RDM and ORD principles and practices into the university's research ecosystem.

These challenges and potential strategies for addressing them are discussed in the following sections.

II. CHALLENGES IN DETAILS

A. Professionalisation

In the absence of national training in data stewardship, UNIL had to define roles, identify skills required for these positions and train its data stewards, with an impact on human resources.

In 2021, when UNIL decided to recruit "data stewards", the title and associated responsibilities of this "new" emerging profession were still poorly discussed and studied – let alone harmonised, standardised, or defined – both regionally and globally. Unlike established "data" roles (e.g. data librarian, data scientist, etc.) and "IT" roles (e.g. research engineer, full-stack developer, etc.), the term "data steward" lacked consensus in its interpretation, with notable variations between continents [2], and even within the European context [3]. In French-speaking regions, further ambiguity arose from debates over adopting the English term or a French translation, which remains unsettled[2]. At UNIL, where French is the official language, the term "data steward" was perceived unfavourably by some, prompting the adoption of the title "Spécialiste données de recherche (*Data Steward*)" to ensure alignment with the institutional culture.

The lack of an established profession also meant the absence of predefined career trajectories. For instance, the role of data steward does not appear in Switzerland's official career orientation platforms. Consequently, UNIL had to develop the position entirely from scratch. This process involved crafting job descriptions, specifying missions and tasks, outlining the required training and competencies, and identifying potential career prospects. These activities demanded significant investments of time and human resources, none of which had been budgeted or allocated specifically for this purpose.

The institutional context further shaped the implementation of the data stewardship network. UNIL's decentralized structure, with seven faculties of varying sizes and research profiles, required a tailored approach to recruitment and support. For example, faculties in social and political sciences prioritized expertise in managing personal or sensitive data, whereas humanities faculties had different requirements. To maintain institutional coherence while addressing these diverse needs, UNIL established a dual-level support structure:

1. Local/Faculty-Level Support: Tailored to the needs and expectations of individual research communities within faculties.

2. Institutional Support: Centralized efforts aimed at institution-wide projects and initiatives.

The specific tasks and responsibilities associated with faculty-level support were developed collaboratively with the faculties to reflect their unique contexts. Similarly, the desired profiles and competencies of data stewards were defined to align with disciplinary and faculty-specific requirements, including education level, professional experience, and specific knowledge and skills.

Recruitment posed its own set of challenges. The lack of specialized training programs in Switzerland necessitated internal training for the newly hired data stewards. The staggered recruitment process, which spanned approximately eight months, precluded a standardized onboarding protocol. Moreover, the diversity of recruits' backgrounds resulted in different levels of expertise in RDM, FAIR[3] principles, and ORD practices. The limited allocation of 0.3 FTEs per faculty further also constrained the time available for onboarding activities.

B. Recognition

The lack of formal recognition (e.g. through certification) of these positions sometimes makes it difficult to promote this expertise within the institution and the research community.

At the institutional level, a significant challenge precedes even the formal recognition of data stewardship roles: the wider recognition of the efforts, practices, and acquisition of skills — often even at a basic level — by researchers in the areas of RDM, FAIR and ORD. Within Swiss academic institutions, despite the widespread adoption of RDM/FAIR/ORD strategies and policies, as well as support for research

[2] For instance, in France, data stewards are called "ambassadeurs", "correspondants", or "référents" des données [de la recherche] [4].

[3] Findable, Accessible, Interoperable, Reusable. See [5].

evaluation reforms[4], these commitments often remain aspirational. Institutions frequently fail to translate intentions into actionable frameworks or measurable incentives.

Consequently, individuals supporting the research community in developing these skills and implementing best practices—namely, data stewards—are similarly underacknowledged. Their expertise and contributions are rarely recognized, leaving data stewardship marginalized within the broader landscape of research support. This lack of institutional recognition is regrettable, as data stewards hold a critical role in driving cultural transformation and fostering a transition toward research evaluation systems that value sound data management and FAIR/ORD principles.

The invisibility of data stewardship as a profession has tangible consequences. Without formal acknowledgment, this role remains underappreciated and is not fully integrated into the broader network of research support services. To address this gap, academic institutions must go beyond declarations of intent, embedding recognition and support for both researchers' efforts in RDM/FAIR/ORD and the vital contributions of data stewards. Formalizing these roles as essential components of research support infrastructure will be crucial for ensuring their visibility, sustainability, and effectiveness in fostering open and responsible research practices.

C. Stabilisation

The lack of permanent funding creates an unstable context, hindering recruitment.

Closely linked to the issues of recognition and visibility is the pressing challenge of stabilizing data stewardship positions. This issue is critical not only for academic institutions but also for the research community. At UNIL, the need for sustained support, guidance, and training in RDM/FAIR/ORD was reaffirmed through an internal survey[5] conducted in 2021. The survey revealed significant gaps in awareness and familiarity with key frameworks:

- Only 10% of respondents reported being "completely familiar" with the FAIR principles.
- 53% were unaware of the Swiss National Science Foundation's (SNSF) ORD requirements.
- 68% were unfamiliar with UNIL's own RDM policy (i.e. Directive 4.5).

These findings underscore the necessity of continuing and deepening efforts to raise awareness and provide training on these critical issues over the long term. However, academic institutions often struggle to recognize the value of allocating permanent resources to such roles. This reluctance translates into persistent difficulties in recruiting and retaining skilled professionals. The temporary nature of these positions, combined with their limited visibility and lack of clear career paths, has led to a phenomenon akin to a "brain drain," where talented individuals are driven away by the precariousness of these roles.

D. Reinforcement

Current part-time allocations (i.e. 20%-30%) do not guarantee optimal support – considering the importance of proximity work with researchers.

It has been estimated that, on average, 5% of overall research costs should go towards data stewardship and that data FAIRification requires about 1 professional for every 20 researchers (Mons, 2020). In Europe alone, this means about 500,000 professionals of various kinds to support researchers through experimental design and data capture, curation, storage, analytics, publication, and reuse (OECD, 2020).

The situation at UNIL is a good illustration of the significant gap between these recommendations and the current resource allocations. With only 2 FTE positions dedicated to data stewardship for a research community of approximately 2,500 researchers, UNIL falls markedly short of the resources required to ensure the quality of research data produced and their "FAIRness". Achieving these goals demands intensive fieldwork, including personalised support and tailored training programs – activities that, while time-consuming, are indispensable. This close engagement is essential not only to foster trust and collaboration, but also to raise awareness of RDM Support services, thereby facilitating interaction with researchers and mobilising them towards more open science and RDM/FAIR/ORD-oriented practices.

One approach at UNIL has been to assign data stewardship roles in some faculties to people who already hold research support positions at faculty level (e.g. research consultants or ethics committee coordinators). This "dual role" model offers both advantages and challenges. On the one hand, it leverages existing expertise and faculty and disciplinary familiarities, which can streamline processes, improve efficiency, and increase the visibility of data steward positions. On the other hand, there is a risk of confusing roles, overburdening individuals, and diluting the attention necessary for effective data management, which could undermine the quality of support provided.

E. ORD strategy and resources

The different timescales between the medium and/or long-term institutional and national commitment to ORD and the funding of these positions hinders the

[4] E.g. DORA, CoARA, Barcelona Declaration on Open Research Information.

[5] Not published yet.

efficient implementation of the strategy and the optimal planning of the required resources.

A significant challenge within academic institutions lies in their pronounced inconsistencies, particularly concerning temporal frameworks. First, institutional strategies and policies – at local/institutional and/or national levels – are often defined with a short-term perspective, typically three to five years. This approach neglects the need for long-term commitment and the complexities of higher education environments, where the sustainability of initiatives depends on continuity and sustained engagement. This question of temporality is also an issue when it comes to the relationship between the research community and support staff. Given that research data and its management may be a personal and private matter for researchers, it seems necessary to be able to cultivate and maintain long-term relationships of trust. Short-term institutional planning fails to provide the stability required to establish these essential bonds, ultimately jeopardizing the effectiveness and ethical foundations of research practices. Finally, while academic institutions frequently advocate for "sustainability" in their rhetoric, their actions often contradict this principle. A notable example is the reluctance or inability to allocate permanent positions, despite the clear need for long term human resources to sustain initiatives and build institutional memory. This inconsistency between discourse and practice undermines the very notion of sustainability and contributes to systemic inefficiencies.

III. SwissDS-ENV PROJECT

In this context, and in parallel with the setting up of the data stewards network at UNIL, UNIRIS partnered in 2022 with other Swiss Academic institutions to answer a call for projects opened by swissuniversities.

This call was launched in the framework of the programme Open Science, Part B – ORD, in the context of the Swiss National Open Research Data Strategy and its Action Plan [6]. The call was open to projects that aim to contribute to Measure B5 of the Action Plan and its action line B5.2. Measure B5 addresses the professionalisation of ORD specialists and related services. Action line B5.2 addresses the establishment and strengthening of data stewardship in HEIs.

The project proposal was accepted at the end of 2022. The SwissDS-ENV project started in early January 2023 and is funded until the end of 2024. UNIL/UNIRIS is the leading house for the project. The other academic institutions involved are HEP-Vaud, HEG HES-SO, ZHAW, FORS, SIB, UNIGE, UZH.

The main motivation of SwissDS-ENV is to contribute to a better recognition of the role of the data stewards in Switzerland. This is to be done by increasing the visibility of their position, of their tasks and responsibilities, and by contributing to the professionalisation of this role. As discussed above, very often data stewards, though having acquired a high level of knowledge and of expertise in their skills, do not hold a recognized qualification in this area. The project undertakes to offer a certification training course in data stewardship in the form of a CAS, which does not yet exist in Switzerland.

The project is divided into three main objectives and three corresponding work packages. The first work package aims to specify the roles, the skills, and the capacities of data stewards. The second work package focuses on the design and the production of the certification course[6]. And the third one focuses on setting up and coordinating a national network of data stewards. This latter has been handed over to the Swiss Research Data Support Network (SRDSN) and this work package has been redesigned to establish a network of course alumni.

To better define the role of a data steward (first work package), we started by defining in detail the profile of six specific data jobs: data management officer, data librarian, data scientist, data curator, data archivist and data steward. For each of them we specified their mission and their responsibilities. We also gave examples of typical tasks and at which stage of the data lifecycle these data professionals are likely to get involved [7]. This allowed us to clarify the scope of the role and the tasks of the data stewards by reference to those of other data professions. Specifically for data stewards, we then identified their tasks at each step of the data lifecycle, and we assessed the typical rough level of contribution for each of these tasks: advice, guidance; or support (more active way) (see Figure 1). At this stage of the work, we strongly involved the data stewards of UNIL, to make the best use of their field experience. Moving a step further, and still in close collaboration with the data stewards, we specified the competences needed to perform these tasks, in as precise a formulation as possible, with the corresponding lifecycle step and the level to which the course should aim to train [8]. We then translated these competences into learning objectives of the course, which were to be addressed by the modules and the sub-modules of the course.

[6] Disclaimer: The certification course (CAS) in data stewardship is being developed in line with the directives of LEHE, of the LFCo and of the reference criteria of swissuniversities. The CAS is produced by universities and for members of universities. As such, it is thus not in competition with private offers.

Figure 1: SwissDS-ENV project: data stewards tasks and contributions at each step of the data lifecycle (lifecycle adapted from FORS). University of Lausanne, 2023. The design of the course, the specific modules, sub-modules, and their detailed contents were then discussed extensively with the project partners and the Data Stewards at UNIL.

The structure of the course is shown in Figure 2. After the transdisciplinary core curriculum, learners can do individual project work or follow a specialised module in Biological and Medical Sciences or in Social and Human Sciences. Upon successful completion of the assessments, they will receive a Certification of Advanced Studies (CAS) in data stewardship (and 12 ECTS) from the University of Lausanne and the Professional learning department (FCUE). The certificate will indicate their orientation, either generalist or specialised.

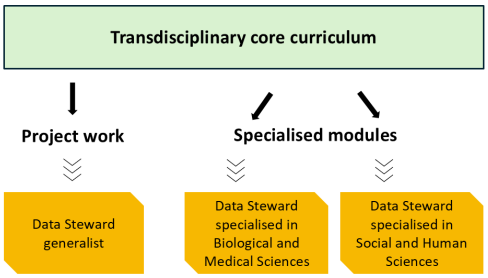

Figure 2: General structure of the CAS UNIL in Data Stewardship (2024-2025 edition). The development of the transdisciplinary core curriculum has been coordinated by HEG/HESSO-GE. The specialised module in Biological and Medical Sciences has been developed by SIB and the specialised one in Social and Human Sciences by FORS.

The first edition of the CAS UNIL in data stewardship started in October 2024 and is due to end in June 2025. It hosts 16 trainees, most of them are data stewards of UNIL and the rest are other data professionals, mainly from the project partners. The trainees of this first edition are expected to provide feedback to adapt, if necessary, the content and structure of the CAS for future editions. The next edition of the CAS will be launched in spring 2026.

By better defining the role of data stewards and their tasks in relation to the research data lifecycle, by specifying their competences, and by offering a certification course in data stewardship, the SwissDS-ENV project has achieved its main objectives. These outputs should contribute to a better recognition of the emerging role of the data stewards in Switzerland. By training and providing a certification to current and future data stewards, it will improve the professionalisation process of Data Stewardship and offer a new career opportunity in the field of scientific research. However, it is crucial that these efforts are accompanied by a consolidation of the current data stewards positions, and by the opening of more stable and long-term positions supported by rectorates of Swiss higher education institutions. If not, the CAS will not be attractive enough and this contribution to the professionalisation of Data Stewards will remain only short-term.

IV. Conclusion

In the light of these observations and lessons learned from building a network of data stewards on the one hand, and, on the other hand, from creating a certified training course in data stewardship, we hope that this exchange and feedback will pave the way for a formal recognition of the profession of data steward – with a training course, a professional pathway and associated career prospects. We think that data stewardship should become more often a career option for graduates, rather than a “default” career option when one decides or is forced to change career direction.

At least at UNIL, former academic researcher females – who did not wish or could not pursue an academic career – are over-represented in the data stewards positions. This bias makes us think that there is still a long way to go before the profession of data steward is fully recognised and valued by academic environments, until it can be considered as a first-choice potential career path by young professionals. Only then UNIL and other similar academic institutions will be aligned with the principles of Equality, Diversity, and Inclusion (EDI)[7] they defend in relation to data management.

We therefore hope that this work will contribute to the establishment, professionalisation, and recognition of this new profession of data steward, and help to make these positions part of an inclusive environment.

[7] See UNIL’s position and statement on EDI here: https://www.unil.ch/unil/en/home/menuinst/universite/egalite-diversite-inclusion.html.

REFERENCES

[1] Jambé, C. (2015). La gestion des données de recherche à l'Université de Lausanne : enjeux transdisciplinaires. SONAR|HES-SO. https://sonar.ch/hesso/documents/314397.

[2] See Ayres, B., Lehtsalu, L., Parton, G. *et al.* (2022). RDA Professionalising Data Stewardship - Current Models of Data Stewardship: Survey Report (1.0). Zenodo. https://doi.org/10.15497/RDA00075.

[3] See Molloy, L., Nordling, J., Grootveld, M. *et al*. (2020). D3.4 Recommendations on practice to support FAIR data principles (1.1). Zenodo. https://doi.org/10.5281/zenodo.5357329.

[4] See the work of Working Group 4 of the Réseau des ateliers de la donnée: https://recherche.data.gouv.fr/fr/page/reseau-des-ateliers-de-la-donnee.

[5] Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al*. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, https://doi.org/10.1038/sdata.2016.18.

[6] Swissuniversities (2021). Swiss National Open Research Data Strategy and Action Plan. Available online at: https://www.swissuniversities.ch/en/topics/open-science/open-research-data/national-strategy.

[7] SwissDS-ENV project (2023). Definition and profile of Data professions. University of Lausanne. Available online at: https://wp.unil.ch/swissds-env/action-1/.

[8] SwissDS-ENV project (2023). Data Steward Skills and Abilities Grid. University of Lausanne. Available online at: https://wp.unil.ch/swissds-env/action-1/.

ACKNOWLEDGMENTS

SwissDS-ENV is a swissuniversities project funded in the framework of the programme Open Science, Part B – ORD, Measure B5 of the Action Plan and its action line B5.2.

The authors would like to thank the partners of the SwissDS-ENV project, the entire UNIL RDM support network: UNIL data stewards and. Research engineers, DCSR members, and the faculty members involved in the RDM support network.

CONTRIBUTIONS

Contributors' roles:

Carmen Jambé: conceptualisation, formal analysis, methodology, resources, writing, review & editing

Marielle Guirlet: conceptualisation, formal analysis, methodology, resources, writing, review & editing

Gérard Bagnoud: conceptualisation, review & editing