

Project Proposal

Data Stewardship embedded: Integrating Discipline-Specific Know-How into Data Stewardship

Acronym

DSEmbedded

Participating institutions

ZHAW Zurich University of Applied Sciences, Services Research Data

Contact and project manager

Simon van Rekum – Research data specialist

simon.vanrekum@zhaw.ch

Vanessa Klaas – Research Associate R&D Unit

vanessa.klaas@zhaw.ch

Start and end dates

01.01.2023 - 31.12.2024

Remark: *In addition to the measures outlined in the present proposal, ZHAW plans to participate in the highly synergistic SwissDS-ENV project (proposal submitted by the leading house, University of Lausanne). The share of federal contributions ZHAW allocates to the SwissDS-ENV project is deducted from the overall figure of federal contributions attributed to ZHAW and thus not listed in the attached budget form concerning the present proposal. However, ZHAW's activities in the SwissDS-ENV project are enlarged upon in the sections on synergies and work packages.*

Abbreviations

EE	Embedded Expert
FAIR	Findable, Accessible, Interoperable, Reusable
ORD	Open Research Data
RDM	Research Data Management
ZHAW	Zürcher Hochschule für Angewandte Wissenschaften (Zurich University of Applied Sciences)
ZSF	ZHAW Services Forschungsdaten (Services Research Data)

Summary

Currently, RDM support along with the hosting and curation of critical RDM-related applications are run by ZHAW Services Research Data — a central unit at the interface of ICT, the university library, and the Research & Development Unit at the President's Office. However, the excellence of RDM practices and the compliance with the FAIR-principles as promoted in the context of ORD relies on more than researchers' access to a centralized support center. It is a paramount requisite of any support model that domain-specific methodologies and data characteristics are adequately considered.

To meet these demands, ZHAW proposes to integrate domain-specific know-how into its data stewardship by designing a role of Embedded Experts. This function is to be assumed by researchers with proven RDM skills and high motivation to cater for the growing demand for ORD. Retaining their core occupation as researchers, the Embedded Experts will dedicate a workload of 0,2 FTE (20%) to RDM and ORD related activities. Their assigned tasks involve RDM support for researchers at their respective department and the communication and promotion of ORD. Thereby, good ORD behavior and RDM practices gain momentum within a research environment and trigger the formation of a community of practice.

The main goal of the DSEmbedded is to implement these measures during the project period. To that end, ZHAW Services Research Data and the appointed Embedded Experts will draw up a dynamic scheme of collaboration that fits the individual structures of each involved department and ZHAW's strategic alignment.

1. Current state of Data Stewardship at ZHAW

1.1 Background

ZHAW took up activities towards Open Science on a strategy level in 2015, issuing an Open Access policy and revising its Research & Development policy (F&E-Policy). In a next step, while participating in the DLCM 2.0 project, ZHAW has carried out pilot studies to explore the challenges, needs, and processes linked to publishing research data as Open Research Data (ORD). The project revealed that ORD increases the demand for best practices of Research Data Management (RDM) in general as well as discipline-specific data processing workflows.¹ Three levels where action regarding RDM is crucial were identified:

- normative level

¹ See Fürholz, Andreas, Jaekel, Martin, 2021. Data life cycle management pilot projects and implications for research data management at universities of applied sciences. In: 3rd Swiss Research Data Day, online, 22 October 2020. Carouge: RESSI.S. 32-46. Available online at: <https://doi.org/10.21256/zhaw-23073>

- tools & infrastructure level
- support level

Against the backdrop of these findings, the ZHAW Management Board requested to set up a cross-organizational unit, *ZHAW Services Research Data* (ZSF), consisting of members affiliated to the ZHAW university library, ICT department, and the Research & Development Unit at the President's Office. The ZSF is to provide RDM services along the data lifecycle as pictured in Figure 1.

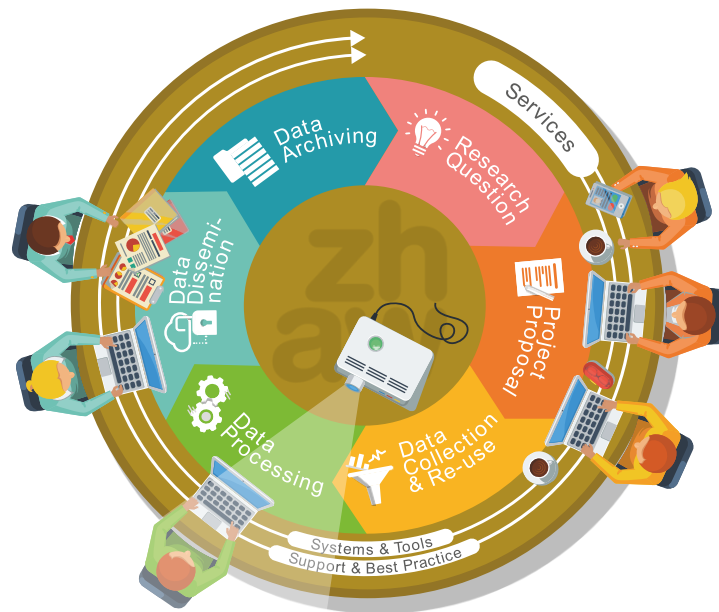


Figure 1 Centralized RDM support group creating synergies

The rationale behind the installation of a centralized RDM support team was to concentrate resources and expertise and generate synergy effects.

1.2 Current situation

The official kick-off of ZSF took place on October 1st 2020 with a budget of 2.8 full-time equivalents (FTE). ZSF is an autonomous team and reports to a project control board consisting of senior staff from the university library, ICT, President's Office, management board, and central management. Furthermore, the needs of researchers and discipline-specific focus topics are discussed in bi-annual meetings with the established advisory board consisting of one member per ZHAW department. ZSF publishes its services through the self-

service portal of ZHAW and existing communication channels such as newsletters and departmental meetings.

The three levels of action identified during the DLCM 2.0 project were implemented as follows:

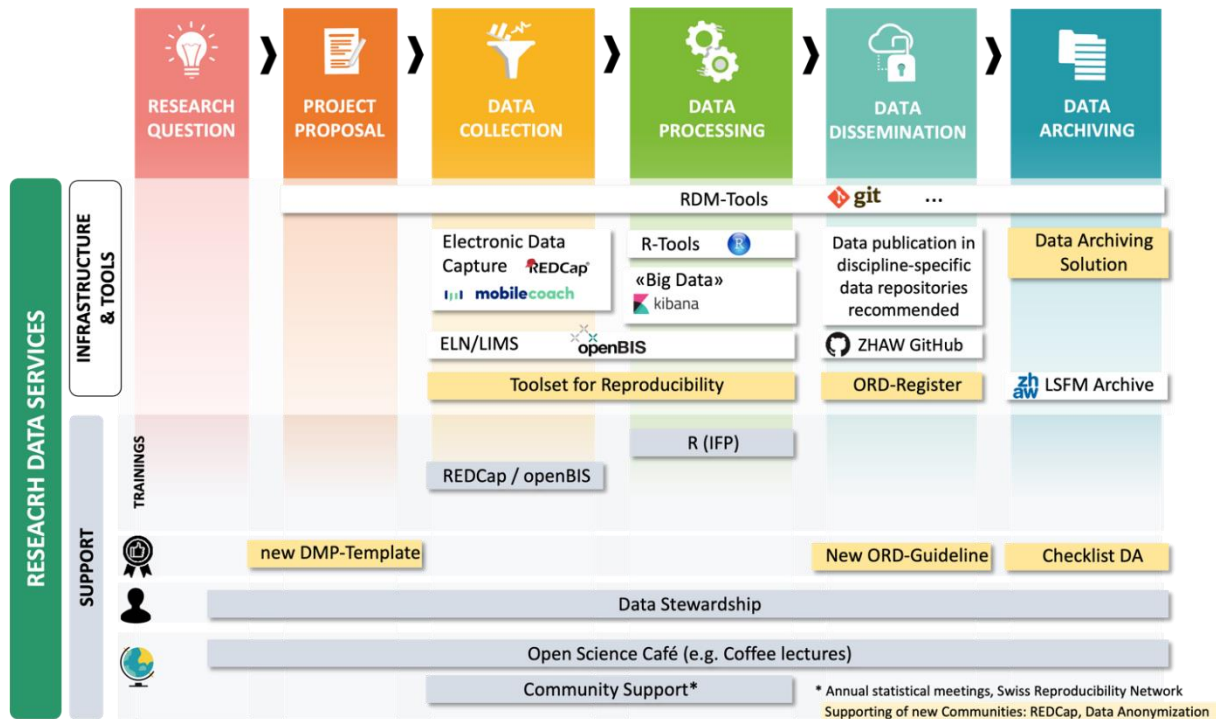


Figure 2 Service Portfolio of ZHAW Services Research Data - yellow denotes services that are currently implemented

As shown in Figure 2, the support level plays a central role from the beginning. Support includes training, guidelines, best practices and foremost individual and hands-on support for researchers on topics along the data lifecycle (ZHAW Data Stewardship). Recurring requests from researchers concern mostly questions around data privacy and data publication, data management plans, and tool-specific questions (e.g., REDCap for data collection, MobileCoach for chatbot-based intervention studies) as well as coding support for data processing and analysis (e.g., R).

The team members of ZSF have backgrounds in a broad scope of disciplines ranging from engineering, computer science, and psychology, to art history – with a common interest in supporting researchers, Open Science, and RDM. Many of the team members hold a PhD degree and have proven research experience, which qualifies them to accurately analyze the current stage of development in the field of RDM and identify gaps. These unique sets of RDM competences by its members makes ZSF an excellent service provider.

Together with the Teams Open Access and Open Educational Resources at ZHAW university library, ZSF runs an Open Science Café – a periodical, virtual meeting that offers inputs on Open Science related topics and room for discussion, questions, and debates. However, format and contents need to be revisited and geared towards a less generic and more problem-oriented, discipline-driven approach.

In addition to providing services, ZSF members are actively involved in the national RDM community (RDM-Support CH), the German-speaking research data information portal (forschungsdaten.info) and represent the ZHAW as a local node in the Swiss Reproducibility Network (SwissRN).

1.3 Gap analysis

While the centralized team provides high quality support in general RDM topics as described above, depth, efficiency, and customization of the support vary depending on the background of the ZSF members. From the viewpoint of disciplines not covered by ZSF's competence (e.g., social work, physiotherapy, aviation, etc.), the skill sets of ZSF members as of today correspond more to the skill sets of generalists than of specialists. To deal with intricate, discipline-specific requests, the team is often bound to either refer to domain experts or to perform time-consuming research to deliver recommendations based on theoretical knowledge instead of hands-on experience. In short, the knowledge of discipline-specific standards and data processing workflows on part of the ZSF team does not equally cover the demands of different disciplines.

Moreover, owing to the increasing popularity of the services, the number of requests is on a continuous rise. This high demand for RDM support entails the challenge of organizing Data Stewardship in an efficient manner that proves sustainable and scalable in the long run. At the same time, it confirms the ZSF's potential and impact and underpins ZHAW's endeavor to further develop the RDM support services on a strategic and conceptual level.

After two years of operation, the ZSF team is repeatedly confronted with the following questions:

- How to build up a sustainable and scalable individual support?
- How to fill gaps in discipline-specific knowledge about standards and data processing workflows?

- How to improve support quality while facing an increasing number of requests?
- How to intensify the communication with researchers and get a better understanding of their needs, so we can tailor our services to increase our impact?
- How to build ORD-communities inside and across the departments?
How to promote the reuse of ORD through raised awareness and discipline-specific preparation of the research data prior to publication?

The solution to these challenges presents itself in an extended data stewardship model that partly rests on the centralized ZSF and partly is embedded in the departments. We thus anticipate a chance to further advance the support by integrating domain-specific experience concerning standards, best practices, and data processing workflows. To that end, ZHAW intends to recruit so-called “Embedded Experts” (EEs) within the research communities of the departments (their role will be described in more detail in the following section). By thus adding resources of the amount of more than one FTE, ZHAW satisfies the evident demand on behalf of the researchers – and confirms the opinion about the pertinence of allocating sufficient resources to data stewardship in relation to research effort.²

2. Planned measures: the appointment of Embedded Experts

2.1 Organizational integration, recruitment, and involved departments

Whereas the ZSF acts as a central, cross-organizational support unit, the EEs, being affiliated to the departments, operate in a decentralized manner. This type of organizing data stewardship in fact has gained increasing popularity and been successfully pioneered by TU Delft.³ However, unlike the data stewardship model by TU Delft, ZHAW’s proposal holds that the core function of the staff assuming the role of EE needs to remain their research activity. This way, it can be ensured that they maintain a pragmatic lens on data stewardship grounded in the daily tasks of researchers. Also, their permanent integration in research programs,

² See e.g., Mons, Barend, 2020. Invest 5% of research funds in ensuring data are reusable. In: Nature 578, 491. Available online at: <https://doi.org/10.1038/d41586-020-00505-7>

³ See Plomp, Esther, Wang, Yan, 2022. Data Stewardship at TU Delft. Available online at: <https://doi.org/10.5281/zenodo.7015283>

panels, and discussion culture ensures their access to communication channels as well as their authority in communicating best RDM practices and promoting ORD.

In the run-up to this project proposal, contacts with candidates for the role of EE and other stakeholders at the departments already have been established. These preliminary discussions have shown that, whereas in most departments, the role of EE is likely to be assigned to a single person, in others, the role will be shared between several persons. Concretely, five of the seven targeted departments favor an implementation of the role EE by one researcher, who will dedicate 0,2 FTE to the involved tasks. At two departments, the same workload will presumptively be distributed among a core group of 2-4 persons depending on their respective skill sets, interest, and expertise. These different ways of implementing the role of EEs warrants flexibility for the departments and further allows an appraisal of the advantages and disadvantages of either model towards the end of the project period.

The 1,4 FTE assigned to EEs will be deployed in equal shares to the following departments:

- School of Applied Linguistics
- School of Engineering
- School of Applied Psychology
- School of Health Sciences
- School of Life Sciences and Facility Management
- School of Management and Law
- School of Social Work

As of the moment of this proposal's submission, all of the aforementioned departments have assured their support. In addition, adequate candidates have already been identified and recruited for these mandates.

2.2 Collaboration between ZHAW Services Research Data and Embedded Experts

DSEmbedded contemplates a collaborative management of support processes. This approach seeks to bundle bottom-up and top-down efforts with the EEs acting as brokers between researchers and the ZSF. In so far as their capacity allows, EEs will function as an additional first contact point for RDM support requests (the minute regulations and communication paths in handling queries will be defined during the project period). Whenever required, ZSF will

provide additional support or pass the question to further relevant experts (for example, in delicate legal matters or in case of overlaps with Open Access related topics). The EEs and the ZSF team will establish a panel of regular (e.g., weekly, bi-weekly) meetings to exchange ideas, skills, and experiences. In this forum, ZSF and EEs will tackle the optimization of RDM workflows, by forming separate task groups (e.g., on how to deal with informed consents?) based on the needs of researchers as communicated by EEs (see below, work package 3). The aim of this bottom-up process is to develop customized RDM services on the bedrock of a larger generic framework. At the same time, EEs also serve as a channel in top-down communication. In this regard, their task is to advocate good RDM practices by informing and updating their local research communities about standards, recommendations, and critical documents.

The ultimate objective of the above-described approach is to lay a foundation for communities of practice, in which problems can be informally communicated and discussed, and knowledge and expertise are shared. The figure below illustrates this envisaged model of embedded data stewardship:

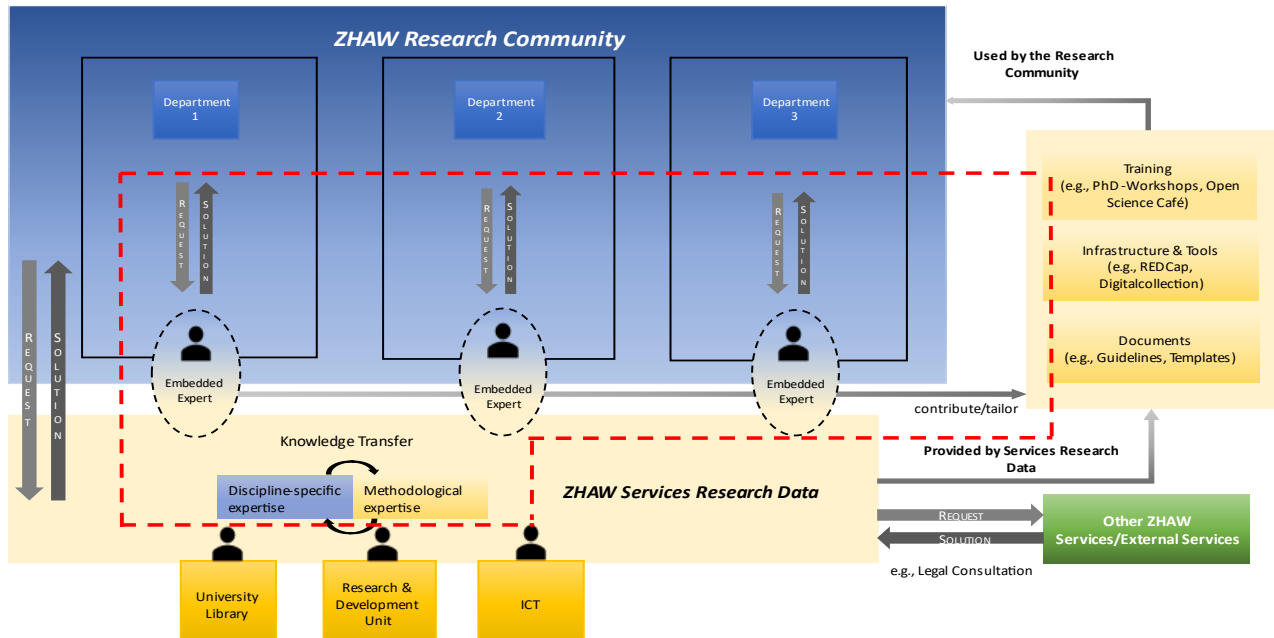


Figure 3 Model DSEmbedded – the red-dotted line confines the planned extension of data stewardship

3. Sustainability and strategic alignment

3.1 Sustainable workflows and collaboration scheme

DSEmbedded aims at establishing a two-dimensional model of data stewardship encompassing the centralized ZSF and the EEs. During the project period, the involved parties (that is the ZSF, the EEs and their respective departments) will outline a collaboration scheme, which addresses the following points:

- Definition of remits and tasks assigned to ZSF and the EEs
- Support-model: the process of handling requests
- The modes of communication between departmental and central RDM services
- The modes of communication between researchers and service providers

The collaboration scheme is to be understood as a dynamic working model, adaptable to changes at the ZSF and the departments. It will be periodically evaluated and reviewed during the project period and recorded in a living document that serves as a guideline, which can be updated even after the close of the project period. The scheme comprises a set of regulations that warrant efficient task-sharing and facilitate the onboarding of future EEs in case of personnel turnover. As a result, the model is made sustainable and not contingent upon the employment of one person. Once the scheme is drawn up and the workflows established, the initial coordination effort of the involved parties is likely to subside.

Moreover, the scheme will be tailored to the needs and internal structure of each department. It will thus ensure that the departments retain a high flexibility to design, realize, and further develop the role of EEs. Thus, the role can be formally linked to a key person, to an impersonal role (e.g., in the ever more popular organizational model of holacracy), or a task-sharing core group of persons. That scope of options enables fathoming the most fitting and sustainable solution and gives room for different funding sources and solutions of the role of EE.

3.2 Communities of Practice

A more indirect but crucial contribution to the sustainability of the model Embedded Experts lies in raising awareness of the relevance of RDM and ORD according to discipline-specific best practices. In the preliminary surveys and contact with potential candidates of EEs, the proposal applicants observed that researchers with high dedication to RDM tend to be proactive and well interconnected in the research community. Also, it could be seen that RDM and ORD related information circulate most effectively through low-threshold communication channels among researchers (bilateral talks, giving updates in research colloquia, etc.). Based hereon, the appointment of EEs promises a spill-over-effect of RDM and ORD related discourses. Best practices unfolding and spreading in a bottom-up way within a research environment are likely to prove more pervasive and resilient than any attempts to impose standards from top down.⁴ Such best practices will be pushed and promoted by the EEs in a close dialogue with ZSF and within the scope of their resources. Concretely, that could involve occasional contributions, e.g., in form of coffee lectures based on case studies reflecting domain-specific needs, the formation of communities of practices around tools (e.g., REDCap) & processes (e.g., data anonymization), or the advising of ZSF in developing RDM and ORD guidelines. Overall, the build-up of communities will have a lasting effect on the RDM and ORD behavior of affiliated researchers.

3.3 Financial sustainability and ORD Commitment on President's Office and department level

In its strategic alignment, ZHAW has placed a pronounced focus on Open Science. The President's Office's Research & Development Policy, issued in 2019, stresses the university's commitment to the Open Science and Open Innovation movements. This commitment is substantiated by the decision to mobilize the required resources to the pursuit of this goal.⁵ As a project partner in the DLCM 2.0 and the openRDM.swiss projects, ZHAW moreover has previously demonstrated its adherence to the Open Science agenda in the area of ORD.

⁴ See Bongji, Gaia, Guirlet, Marielle, et al., 2021. ORD capability of scientific communities Mandate 2 commissioned by Swissuniversities. Available online at:

https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Hochschulpolitik/ORD/ORD_Mandate2_FinalReport_Master_VF_EN.pdf?sword_list%5B0%5D=oa&sword_list%5B1%5D=easi&no_cache=1

⁵ See ZHAW, Rektorat, Ressort Forschung & Entwicklung, 2019. F&E Policy. Available online at:

https://gpmpublic.zhaw.ch/GPMDocProdZPublic/1_Management/1_04_Governance/1_04_01_Fuehrungsgrundlagen/Z_PY_F_und_E_Policy_ZHAW.pdf



On a department level, the proposed project has been unanimously met with approval. Moreover, there have been independent initiatives to assign persons in charge of the public sharing of research data in at least two departments. These endeavors and strategic alignments underscore the university's support of the proposal and warrant a cooperative implementation of its objectives as well as the long-term apportionment of resources and funds to the model of embedded data stewardship.

3.4 Synergies with collaborative project SwissDS-ENV

Alongside the implementation of DSEmbedded, ZHAW is a joint applicant of the collaborative project *Swiss Data Stewardship Environment: Profile – Training – Network (SwissDS-ENV)*.⁶ The proposed project aims at professionalizing data stewards, whereby the assumed measures revolve around the following actions:

- 1) The definition and profile of *data stewards*,
- 2) Training leading to certification,
- 3) Formalization and coordination of a *Swiss Data Stewardship Network (SDSNet)* community of practice.

ZHAW intends to participate in all three actions both in-cash and in-kind. The applicants see mutual benefits, especially in the second action dedicated to the training of data stewards. It will be of avail to ZHAW's project DSEmbedded insofar as the collected experiences in training designs and the resulting modules in various disciplines facilitate the onboarding of new EEs. In reverse, DSEmbedded provides a concrete case study regarding the translation of RDM into a research environment and the challenges it entails. Furthermore, the EEs at the departments of Applied Psychology and Social Work potentially could assume an advisory role in the development of training modules for Social Sciences and point out the challenges in Applied Sciences.

⁶ Contact and Project Manager: Gérard Bagnoud - Director/head of ORD, Information Resources and Archive Services (UNIRIS), University of Lausanne: gerard.bagnoud@unil.ch

3.5 The potential for reusability of the model embedded data stewardship

Although the notion of the two-dimensional – that is partly centralized, partly decentralized – data stewardship initially took shape by considering the specific needs and organizational culture of the ZHAW, it bears potential of being adapted to and reused by other higher education institutions. This is likely to be relevant for institutions where resources are limited and insufficient to assign full time RDM support staff. In this regard, DSEmbedded assumes a certain pilot character: the concept will be tested, evaluated, adjusted, and finally be shared through public dissemination.

4. Planning

4.1 Work packages

No	Title	Responsible unit	Estimated effort in PM	Start	End
1	Project management	ZSF	6,4 PM	Jan 2023	Dec 2024
2	Onboarding EEs	ZSF/EE	ZSF: 5,4 PM EEs: 22,4 PM	Jan 2023	Dec 2024
3	Subject-specific data workflows	ZSF/EE	ZSF: 2,2 PM EEs: 8 PM	Feb 2023	Dec 2024
4	Scheme of collaboration	ZSF/EE	ZSF: 3 PM EEs: 2,4 PM	Apr 2023	Nov 2024
5	Dissemination	ZSF/EE	2,2 PM EE 0,8 PM	Sep 2023	Dec 2024
6	Collaboration with SwissDS–ENV ⁷	Separate resources	Separate resources	Jan 2023	Nov 2024

Key distribution of resources (in total): ZSF 19,2 PM (=0,8 FTE); EEs 33,6 PM (=1,4 FTE)

WP1 Project management: This WP involves coordination, communication with involved staff, organization of events, meetings, project reporting and financial reporting.

WP2 Onboarding EEs: In a kick-off event in February 2023, the members of ZSF and the EEs will get to know one another; a competence profile of each EE will be outlined to optimize task-

⁷ See separate proposal: SwissDS–ENV, submitted by the leading house University of Lausanne

sharing and enable cooperative RDM support even across the borders of departments. During an introduction phase, the EEs will be gradually introduced to the working systems, tools, and meeting structures of ZSF. Since the prospective main task of all involved personnel is RDM support, the EEs will assume first-support tasks at their respective department from the very start and throughout the whole project period. Furthermore, to facilitate mutual exchange and integration of new competences into existing service offerings, the EEs will actively participate in the further development of the Open Science Café in the following ways: they will a) sporadically conduct “coffee lectures”, b) advertise the Open Science Café at their respective departments and c) explore the possibilities of “coffee lectures” or similar types of RDM/ORD micro trainings having guest appearances in already existing academic series of events (PhD seminars, colloquia, etc.)

WP3 Subject-specific data workflows: From February 2023, a panel in form of weekly/bi-weekly meetings will be established, in which ZSF and the EEs determine subject-specific data workflows and derive best practices. In a first phase, the panel will identify and examine selected subject-specific data-workflows at different departments. Thereby, the focus will be placed on the required infrastructure and tools, already existing practices, parallels between departments, and further optimization potential. Based on these findings, the panel in a second phase will establish smaller task groups (e.g., dedicated to qualitative data analysis or REDCap), which design best practices that pay heed to reusability and actual reuse of data as well as the pooling of competences – where possible across departmental borders. In a third phase, the panel will support researchers in the formation of communities, in which these practices are to be imparted.

WP4 Scheme of collaboration: This WP involves setting up an organizational framework for the collaboration between ZSF and EEs. In a first workshop in April 2023, the ZSF and the EEs will delineate the criteria underlying the design and evaluation of the collaboration scheme. Concretely, this concerns the technical framework of the support system and the modes of communication between ZSF, EEs, and researchers (e.g., in the handling of requests: ticket-system, common vs. personalized mailbox, etc.) In a follow-up workshop towards the end of year 1, a review on the collaboration will be conducted, serving as a basis to record the collaboration model in a first document draft. A review of the second year will be take place in fall 2024; and a final version of the collaboration scheme will be outlined at the end of the project period.

WP5 Dissemination: An end-of-year review on the project’s achievements and challenges will be published in ZHAW internal newsletters in fall 2023. At the end of the project period, the

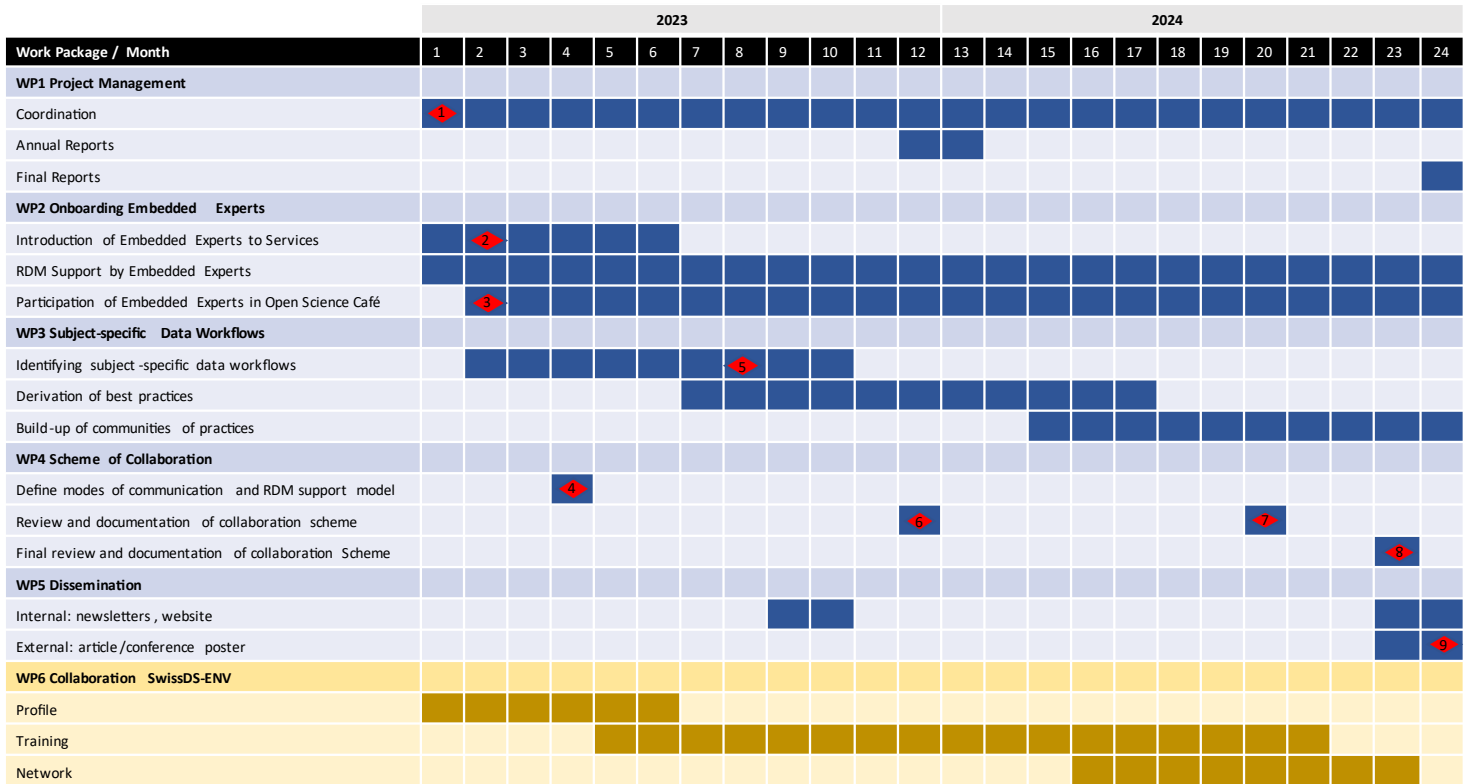
over-all experiences and results will be shared both internally and publicly, e.g., in the form of an article or a conference poster.

WP6 Collaboration with SwissDS-ENV: Corresponds to ZHAW's participation as outlined in the separate proposal, submitted by the University of Lausanne. ZHAW participates with in-cash and in-kind contributions in all three actions (see above, section "synergies"). The granular planning and assignment of tasks is yet to take place during the kick-off phase of the project.

4.2 Milestones and Deliverables

Milestones & Deliverables				
No	Date (month)	Milestone	Associated deliverable	Status deliverable
1	M1	Project kick-off	Project contract	Restricted
2	M2	First workshop/kick-off event	Competence profile of EEs	Internal
3	M2	Relaunch Open Science Café with the involvement of EEs	-	-
4	M4	Internal Workshop – discuss support models; tools of communication; expectations, demands, and evaluation criteria for collaboration	Criteria catalogue to evaluate collaboration scheme	Internal
5	M8	Thematic working groups are established	-	-
6	M12	Internal Workshop - Review of collaboration year 1	First draft documentation of collaboration scheme	Internal
			Annual reports	Restricted
7	M20	Internal Workshop – Review of collaboration year 2	Protocol of review	Internal
8	M23	Documentation of collaboration scheme is ready	Final version of documented collaboration scheme	Internal
9	M24	Final project review, results ready for dissemination	Manuscript on project results ready for submission	Public
			Final reports	Restricted

4.3 Timeline



◆ Indicates Milestone and associated deliverable

Figure 4 Gantt chart of work plan