



- Skills for the European
- Open Science
- Commons

Supporting



Data Stewardship Studies: Skills- orientated Job Profile for Librarians

9. Bibliothekskongress 2025, Bremen
25.06.2025

Achim Winandi (KIT)



This presentation is released
under a CC BY 4.0 license,
except where otherwise noted.

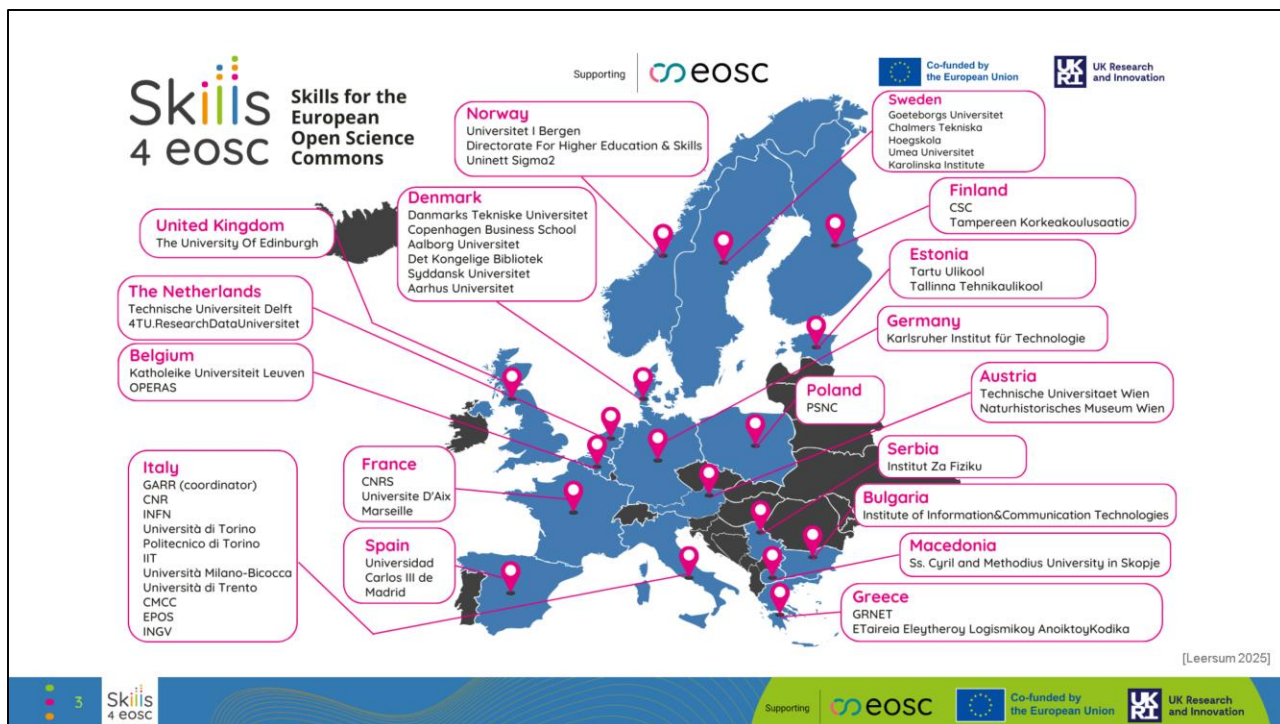


Terminology: EOSC & Skills4EOSC

- [EOSC](#): European Open Science Cloud
 - “EOSC is the European web of FAIR data and related services for research” [[EOSC-A 2023](#)]
 - “EOSC is not a cloud infrastructure!” [[EOSC-A 2023](#)]
- [Skills4EOSC](#) is one of many EU projects developing EOSC
 - „Creating an EOSC-ready European workforce“ [[Skills4EOSC](#)]

Before we start, let's clarify a few terms:

- EOSC is the European Open Science Cloud. EOSC is not a cloud infrastructure like Microsoft Azure or Amazon AWS. Instead, it aims “to federate existing data, research and e-infrastructures” [[EOSC-A 2023](#)].
- Skills4EOSC is one of many projects developing EOSC. Its goal is creating the workforce required for EOSC. Some key results from this project are presented here.



Some quick key facts about the Skills4EOSC project:

- 46 partners in 18 countries work together in the project.
- The project started in September 2022 and ends at the end of August 2025.
- The project received a total funding of about 6.5 million €.

Skills4EOSC: Motivation

The Skills4EOSC project aims to close the 3 gaps identified in the [Strategic Research and Innovation Agenda](#) (SRIA) of the European Open Science Cloud (EOSC):

1. Lack of **Open Science and data expertise**
2. Lack of a clear definition of **data professional profiles** and corresponding **career paths**
3. Fragmentation in **training resources**



The EOSC Strategic Research and Innovation Agenda (SRIA) outlines the steps to achieve the overall goal of creating EOSC. This agenda identifies three major gaps that the Skills4EOSC project tries to close:

- 1. Lack of Open Science and data expertise:** For data stewards, this knowledge is essential to support researchers in creating FAIR data. This underlines the importance of a more thorough training of different roles in the research process, including data stewards. To close this gap, the project creates different training courses and training curricula.
- 2. Divergence in data professional profiles throughout Europe:** Definition and skills required for different roles, e.g. data stewards, vary widely across Europe, different universities, or even between different people inside an institution. The Minimum Viable Skillsets (MVS) for different professional roles are created, to establish a common understanding of tasks and requirements for different roles.
- 3. Fragmentation in training resources:** Training materials differ between different countries in Europe, between different universities in one country, and even between different institutes within universities. Instead of creating a single set of training materials for all researchers in Europe, this project creates a Competence Centre (CC) network. CCs act as hubs for Open Science and EOSC activities, and harmonize training materials on a pan-European level, while each CC adopts the materials to their specific communities' needs. Also, a comprehensive training curriculum for entry-level data stewards is developed during the project.

What's in it for the libraries?

- Methodology for creating open training materials/OER (**FAIR-by-Design methodology**)
- Job profiles for Open Science actors (**Minimum Viable Skillsets**)
- Reusable **training curriculum for entry-level data stewards** (licensed under CC BY)
- Other free training courses (see the [project website](#), e.g. “Learning Path for (data) librarians: Technical skills are the bridge to reproducible research.”)
- And much more (see [project website](#) or [Zenodo community](#))



Libraries can benefit from the Skills4EOSC project in the following ways :

- The **FAIR-by-Design methodology** (shown later in this presentation) can be used to create training materials following best practices. E.g., it ensures that the materials can be published as Open Educational Resources (OER) and easily reused.
- The different job profiles which have been described in the form of **Minimum Viable Skillsets** (MVS, shown later), can be used to verify that the trainings offered by the library cover all important aspects for different roles. Moreover, it can be used to assess whether library staff cover all required competencies and identify needs for additional training or personnel to support Open Science.
- The **training curriculum for entry-level data stewards** (shown later) can be used as a basis for designing training materials.
- Additionally, free training courses are available for enrolment and/or reuse for creating own trainings (see the overview on the project website <https://www.skills4eosc.eu/participate/skills4eosc-training-courses>).
- Many more project outputs can be found on the project website (<https://www.skills4eosc.eu/>) or in the Zenodo community (<https://zenodo.org/communities/skills4eosc/>).

FAIR-by-Design Methodology

Backward Instructional Design Empowered with FAIR principles

focusing on both learners and peer instructors



The FAIR-by-Design methodology developed in the Skills4EOSC project, incorporates best practices and lessons learned from other EOSC projects. The methodology

- encourages reuse of existing training materials, provided they are licensed accordingly,
- ensures that newly produced training materials adhere to FAIR principles, allowing them to be published as Open Educational Resources (OER), and
- embraces continuous improvement through feedback collection.

A detailed description of the methodology can be found in the

- FAIR-by-Design microlearning unit
<https://fair-by-design-methodology.github.io/microlearning/latest/>,
- FAIR-by-Design book
https://fair-by-design-methodology.github.io/FAIR-by-Design_Book/, and
- FAIR-by-Design Training of Trainers
https://fair-by-design-methodology.github.io/FAIR-by-Design_ToT/latest/.

Minimum Viable Product Skillset (MVS)

“the minimum viable product is that version of a new ~~product~~ training course that allows a team to collect the maximum amount of validated learning about customers target group needs with the least effort.”

[Eric Ries, 2009, [Minimum Viable Product: a guide](#), modified for training courses by (White 2025)]

- MVS are based on review of relevant sources e.g. competence frameworks, policy statements
- Use case: inform training design



[Whyte 2025], new adaption of multiple slides

The Minimum Viable Product (MVP) known from software development refers to a product delivered with the fewest features while retaining all important ones. The goal is to collect feedback from customers in an early stage to steer development in the right direction and avoid unnecessary development times for a product or feature that nobody uses.

This principle was adapted to the training of Open Science actors as the Minimum Viable Skillset (MVS). Essential skills for different roles were identified through reviewing existing competence frameworks and policy statements. Once finished, the MVS are used to gather community feedback and improve them. By users, MVS can be used to create training programs or check if their team has all essential skills for practicing Open Science.

Each MVS contains the following sections:

- **Mission:** what this role does in relation to Open Science, both broadly defined and within the EOSC context.
- **Outcomes:** further statements describing results that should be achieved through practicing Open Science.
- **Activities:** actions performed to accomplish the outcomes.
- **Essential skills and competences:** those required to perform the actions.

Data Steward MVS

- 2 roles: Coordinator and Embedded (2 ends of a spectrum)
 - The tasks for the both roles differ
 - Essential skills and competences are the same



Booklet (PDF) on [Zenodo](#)



Webpage on [GitHub](#)

The MVS for data stewards is split into two roles representing both ends of a spectrum of roles:

- **Coordinator:** provides support across different research domains/units.
- **Embedded:** works closely with research teams, requiring domain-specific knowledge.

During the creation of the MVS, it has been found that while the specific tasks differ between roles, both require the same essential skills. During creation of the MVS, extensive feedback from active data stewards was collected.

The full description of the data steward role can be found in the following two resources:

- Data Steward MVS on GitHub (here the materials can be easily reused and feedback can be given with issues):
[https://fair-by-design-methodology.github.io/MVS/latest/MVS Profiles/Data Steward/data_steward/](https://fair-by-design-methodology.github.io/MVS/latest/MVS%20Profiles/Data%20Steward/data_steward/)
- Data Steward MVS as booklet on Zenodo:
<https://zenodo.org/records/14865959>

Skills4EOSC Training Curriculum for Data Stewards



Minimum Viable Skills

Skills and competencies for Data Stewards



Target Audience

Entry level data stewards (broad enough curriculum to covering a broad definition of data stewardship but not too advanced or specialised).



Landscaping

Reviewed existing material and conducted gap analysis of existing curriculum e.g. RDNL, MANTRA, CODATA-RDA etc.



Curriculum Draft

Training curriculum with 8 sections is complete (first draft)



Consultations

Continuous consultations with Data Stewards and experts from across Europe

Development

[Leersum 2025], updated heading

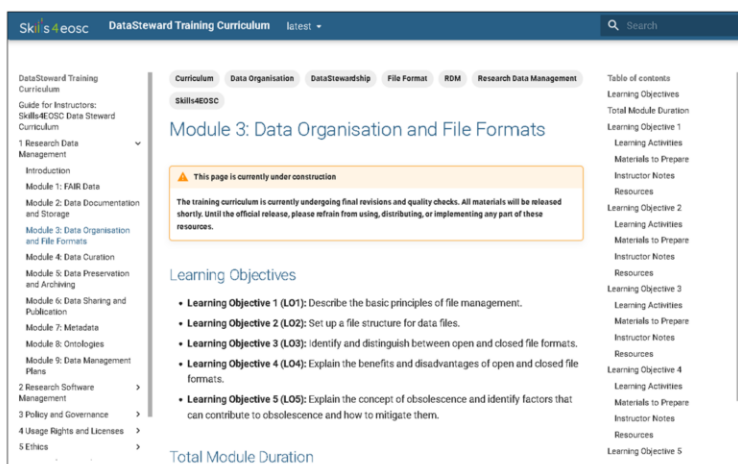
Sections of Training Curriculum

1	Training Skills	Train data stewards to develop, deliver and facilitate training to bridge current gaps in researchers and or support staff's knowledge in Open Science, FAIR and other data management related topics.
2	Research Data Management	Introduces individuals with the essential skills and knowledge required for proficient data management throughout the research lifecycle.
3	Research Software Management	Basics about research software and how to support researchers in research software management.
4	Policy and Governance	Equip learners with the knowledge and skills to take into account and implement different policies, regulations and requirements related to data and software management
5	Usage Rights and Licenses	Learn about possible usage rights and licensing issues and how to mitigate these within the research project
6	Ethics	Equip the data steward with skills to identify ethical concerns related to research projects and mitigate the risks within the scope of data handling practice.
7	Personal Data and GDPR	Make the data steward capable of assisting research projects to being compliant in relation to handling personal data
8	Transversal/Soft skills	Develop management skills and soft skills that will form the basis of all mediation and support work.

[Leersum 2025], updated heading and data

This slide provides an overview of the curriculum sections; more detailed descriptions can be found in the curriculum itself (see next slide). The landscaping and review process revealed that the curriculum should emphasize Transversal/Soft Skills, Training Skills, and Research Software Management, as these topics are currently neglected.

Example from Training Curriculum



The current **draft** version of the curriculum is already available [on GitHub](https://skills4eosc-dscurriculum.github.io/DataSteward-Training-Curriculum/latest/).

The training curriculum is published on GitHub (<https://skills4eosc-dscurriculum.github.io/DataSteward-Training-Curriculum/latest/>) for easy reuse and feedback through GitHub issues. Currently, the curriculum is still undergoing some final changes following an internal peer review. The final version should be published in the next few weeks.

The curriculum starts with a "Guide for Instructors" providing a high-level overview of the development process and the rationale behind the curriculum.

The curriculum is structured into sections, which are itself divided into modules. This modular structure allows an easy partial reuse of the curriculum. Each module contains multiple learning objectives (LO) using Bloom's taxonomy. Each learning objective has

- learning activities,
- materials to prepare
- instructor notes, and
- resources (e.g. links to already available training materials that can be reused).

The "Guide for Instructors" (<https://skills4eosc-dscurriculum.github.io/DataSteward-Training-Curriculum/latest/>) features a more detailed description of the contents of the different parts.

Data Steward Models at universities



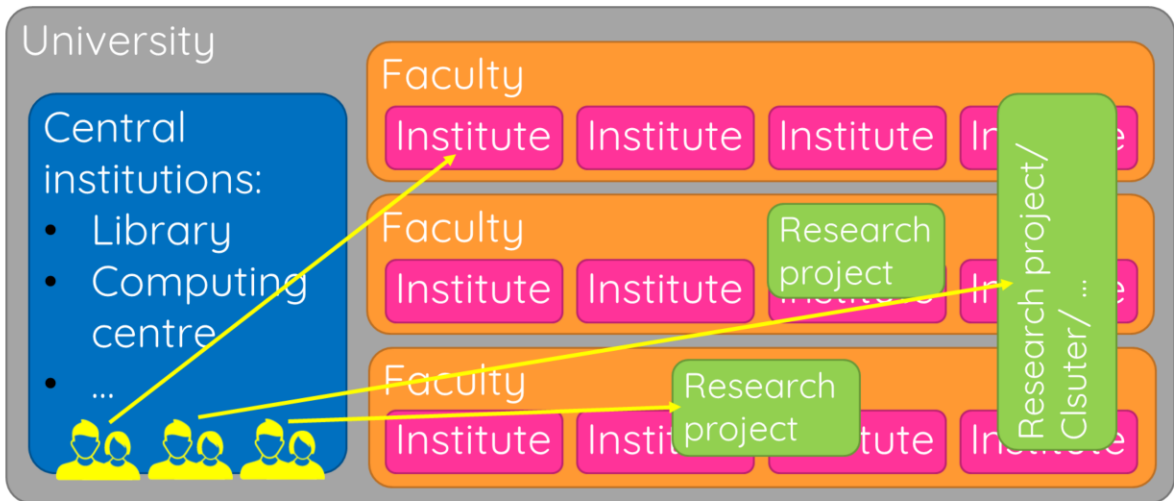
To better understand the role of libraries in data stewardship, we will review several current data stewardship models. The following list is not exhaustive, many more variants exist.

Most universities consist of:

- central institutions (e.g. library, computing centre, etc.),
- multiple faculties covering different subject areas,
- institutes within each faculty, and
- research projects, which can be conducted by a single institute or spanning multiple institutes (or even universities!), especially larger projects like SFBs („Sonderforschungsbereiche“).

The next slides quickly introduce various models showing where data stewards could be located.

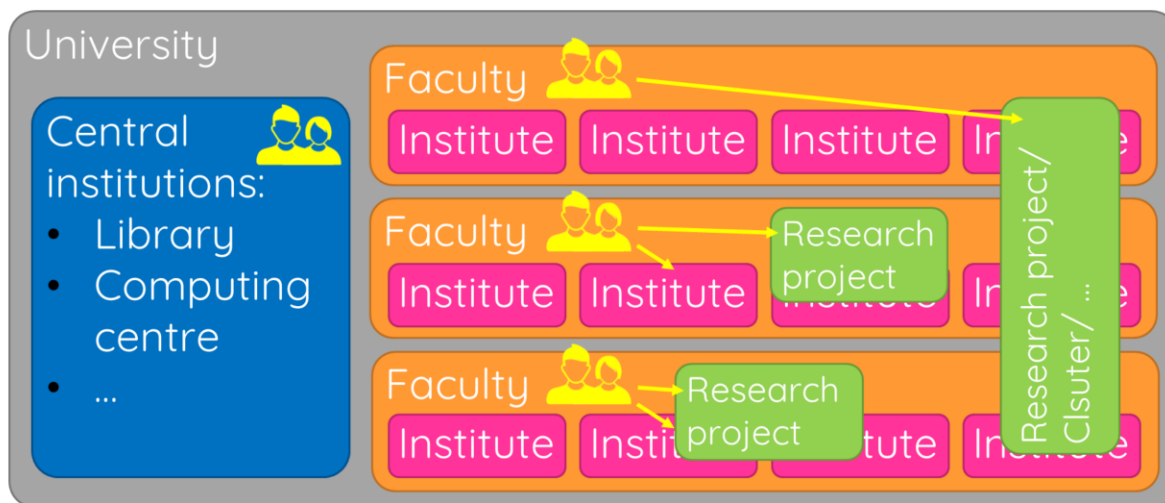
Data Steward Models at universities



One option is locating all data stewards in a central institution, such as the library, computing centre, or a joint venture of both. Data stewards can then be requested by different research projects or institutes within the university.

This model is currently used at TUM (Munich). The TUM Research Data Hub is operated by the library and the Munich Data Science Institute (MDSI). More information can be found at <https://web.tum.de/en/researchdata/request-a-data-steward/>.

Data Steward Models at universities



Another option is to have one or more data stewards per faculty, providing necessary domain-specific knowledge. Additionally, a central institution of the university can host another data steward as the coordinator.

This model is currently used at TU Delft, with the coordinator located at the library and faculties having one or two data stewards. More information is available at <https://www.tudelft.nl/en/library/data-management-and-sharing/get-support-on-data-management/data-stewardship-strategic-framework-for-data-stewardship>.

Data Steward Models at universities



The final option embeds data stewards directly into specific research projects, requiring the most domain-specific knowledge. A coordinator data steward at a central institution can add further benefit by organizing a network and fostering the exchange.

This model is used at RWTH Aachen, where data stewards are located within large research projects. The central RDM team supports their activities and organizes networking events. For more information, see <https://doi.org/10.17192/bfdm.2020.2.8278>.

What's in it for the libraries!

- “Information services focus on **supporting the entire research process**, on creating attractive learning spaces and on promoting information literacy.”
[KIT Library's mission statement (2019), accessed 24.06.2025]
- The tasks of the reference librarian are diverse and vary depending on the focus of the respective institution. [...] As a link between the library and academia, **reference librarians also support the research process.**
[Kommission für Fachreferatsarbeit des VDB, accessed through <https://web.archive.org/web/20220120130349/https://www.vdb-online.org/kommissionen/fachreferat/>, translated by author]

The models presented in the previous slides demonstrate two key approaches for libraries in data stewardship: hosting all data stewards centrally or coordinating decentralized data stewards across the institution. Two statements about university library tasks further support engagement in this field.

The KIT library's mission statement exemplifies this alignment, declaring that "information services focus on supporting the entire research process." With the central role of data in research, data stewardship naturally extends this support function.

Reference librarians are particularly well-positioned for this role. While responsibilities vary across institutions, they consistently serve as intermediaries between library resources and researchers throughout the research process. This relationship provides an ideal foundation for expanding into data stewardship services.



Skills4EOSC has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101058527 and from UK Research and Innovation (UKRI) under the UK Government's Horizon Europe funding guarantee, Grant No. 10040140

Questions? What are your experiences with libraries and data stewardship?

To cite this presentation, please use:

Winandi, A. (2025, June 25). Data Stewardship Studies: Skills-orientated Job Profile for Librarians. 9. Bibliothekskongress 2025, Bremen. Zenodo. <https://doi.org/10.5281/zenodo.15735149>



This presentation is released under a CC-BY 4.0 license

Authorship

- Some slides have been reused from existing presentations. The slides are marked on the bottom right with a short citation and quick note, if changes to design or content have been made. These slides are exempt from the CC BY license.
- Sonja Filiposka. (2023, December 4). FAIR-by-Design Methodology: How to Develop FAIR Materials. Zenodo.
<https://doi.org/10.5281/zenodo.10256852>
- Nida van Leersum, Saba Sharma. (2025, June). Data Stewards Training Curriculum. Unpublished
- Angus Whyte, Dominique Green. (2025, April 29). Minimum Viable Skillsets – a space to structure your training needs. Unpublished