SwissDS-ENV

Definition and profile of Data professions

Author: Jean-Blaise Claivaz, Claire Wuillemin (UNIGE), Gérard Bagnoud, Marielle Guirlet, Carmen Jambé, Sacha Auderset, Guillaume Beausire, Christelle Cocco, Zhargalma Dandarova, Auriane Marmier, Ellina Mourtazina (UNIL), Mathilde Panes (HEP Vaud), Vassilios Ioannidis, Patricia Palagi, Grégoire Rossier (SIB), Eloïse Morisod, René Schneider (HEG, HES-SO Genève), Martin Jaekel, Felix Moerman (ZHAW)

SwissDS-ENV project, *Definition and profile of Data professions*, University of Lausanne, 2023.

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
Version 2 - 26.10.2023

Version	Date	Notes
v1	28.03.2023	Contents section 1
v2	05.2023	Text modifications
v2-1	24.05.2023	Selection of professions and radar representations
v3-1	31.05.2023	Life cycle and job cards
v4	07.07.2023	Addition of figures 1, 2, 3; page layout
v5	28.07.2023	Taking into account the latest comments and a new version of the radar displays
VF_2023	02.08.2023 15.09.2023	Cover page Legend figure 3 (life cycle adapted from FORS 2023)
VF_2023_2	26.10.2023	One sentence added in the context section Conclusion added English version: all terms in English (figure 3 and figures in job cards)

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
Version 2 - 26.10.2023

Table of contents

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
Version 2 - 26.10.2023

List of figures

List of tables

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
Version 2 - 26.10.2023

1. Context

The results presented below have been produced in the frame of the project Swiss Data Stewardship Environment (SwissDS-ENV) funded by swissuniversities. The aim of this project is mainly to develop a certificate course in Data Stewardship.

New professions are emerging around the increasingly important issue of research data management, while older professions are evolving and reshaping to meet the demands of different stakeholders. Given the wide range of fields in which these professions operate, their titles, definitions and scopes of action overlap, and more generally their characteristics are intimately linked to the contexts in which they operate (environment, institutions, teams, generalist versus specialist approach, etc.).

In order to clarify the situation and use a common language, we felt it necessary to specify the various terms and concepts that will be used in the SwissDS-ENV project. To this end, we used the six-stage research data life cycle of FORS as a reference (2023, p.9) (Figure 1), and we selected six professions that are representative and necessary to cover the spectrum of needs of different research situations and institutions.

Figure 1: Data life cycle (adapted from FORS 2023)

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

1

2. Presentation of data professions

2.1. Selected professions

A review of literature and existing projects was conducted to identify and describe different data management professions that act at different stages of the data life cycle. It is based on the landscape of data management professions Wuillemin (2022, pages 10 to 13). The profile of *Data Coordinator/Data Manager Officer* differs from the others in that it is linked to the governance and supervision of research data management at the institutional level, while the other five positions are more directly concerned with direct support for researchers in the creation, use, analysis, storage and preservation of research data. Other terminologies exist to describe the various positions, which we have included in the descriptions of each of these professions.

List of data management professions
1. *Data Coordinator/Data Manager Officer*
2. *Data Steward*
3. *Data Librarian*
4. *Data Scientist*
5. *Data Curator*
6. *Data Archivist*

2.2. Contributions to data life cycle stages

The interaction between the different stages of the data life cycle and the different professions illustrates, in a very schematic way, the main differences and the scope of action of the selected professions. For example, in the graph below (Figure 2) we see that the *Data Steward* is involved at every stage of the life cycle, while the *Data Archivist* is only involved in the second stage of the life cycle. What is not shown in this representation is the intensity with which a profession is involved in each given stage. This information may be found in the job description sheet, and more specifically in the radar representation (see section 3). For the Data Steward, this representation is completed by another figure (Figure 3) detailing the tasks performed at each stage of the life cycle. Two levels of contributions are distinguished here: Advising (lower contribution) and Supporting (higher contribution). This level of contribution may vary slightly between different contexts, depending on the institutional environment and resources available.

Definition and profile of Data professions
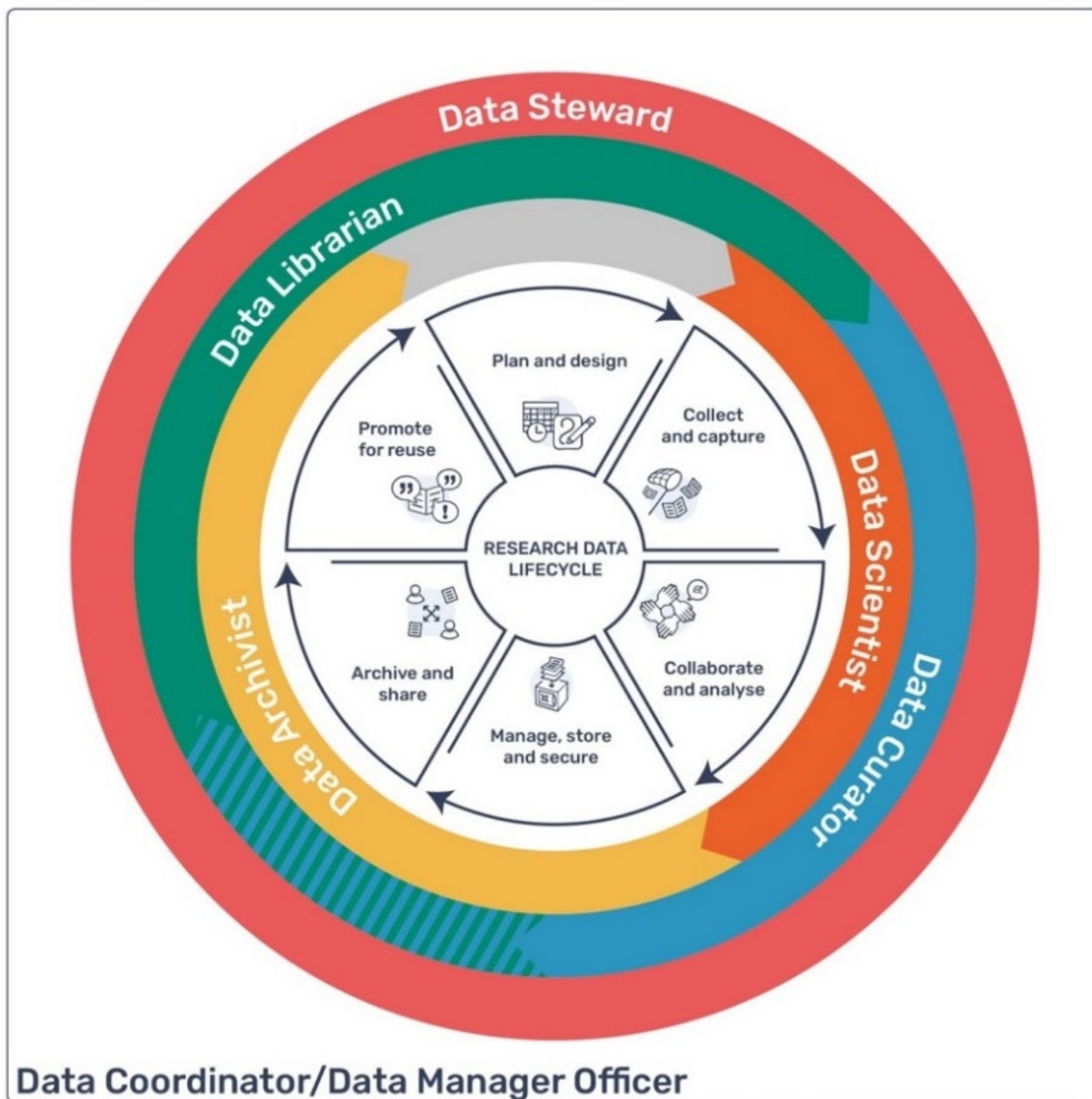SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

2

Figure 2: Contribution of Data professions in the various stages of the data life cycle (life cycle adapted from FORS 2023)

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

3

3. Detailed description of data professions - in the form of job cards

3.1. Data Coordinator/Data Manager Officer

Other name(s) / Overlap with	*(Chief) Data Manager, (Chief) Data Officer*
Keywords (tags)	Governance, strategies, policies, compliance, best practices, infrastructure, steering, supervision, coordination, support, internal and inter-institutional relations

Definition	
The Data Coordinator/Data Manager Officer occupies a strategic and political position within the institution, acting at both institutional and cross-functional levels. Responsible for the governance of research data throughout its life cycle, he/she contributes to the establishment of strategies and policies in line with the various requirements - legal, regulatory, national and international - for the management (MDM) and openness (ORD) of research data. Similarly, the Data Coordinator/Data Manager Officer participates in the development of infrastructures and tools; and in the establishment of best practices, recommendations and training for researchers. In addition, he/she coordinates the various professions involved in supporting research data management.	

Mission(s)	• Oversee the governance of research data at the institutional level, and support data management and openness throughout the data life cycle • Ensure compliance with national and international requirements, norms and standards for managing and opening up research data
Examples of tasks	• Develop strategies and institutional policies for managing and opening up research data • Establish a framework of best practices for managing and opening up research data • Develop infrastructures/tools for managing and opening up research data • Develop communication and training materials for researchers • Coordinate the various stakeholders involved in supporting research data management, including *Data Stewards* • Represent the institution in various national and international bodies
Life cycle stages • Plan and design • Collect and capture • Collaborate and analyse • Manage, store and secure • Archive and share • Promote for reuse	**The Data Coordinator/Manager Officer is involved in all stages of the life cycle in a cross-functional manner**

Table 1: Data Coordinator/Data Manager Officer job description

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

4

3.2.Data Steward

Other name(s)	*Data Consultant, (Research) Data Manager, Research Data (Management) Specialist*
Keywords (tags)	Support for researchers, provide best practices in research data management and openness, ensure compliance with requirements

Definition	

Through their consultative expertise, *Data Stewards* support and facilitate researchers for the management of research data (and metadata) throughout their research life cycle. They are often trained and experienced in research.

Data Stewards can be either *embedded* or *generic* (Verheul et al. 2017, p.7). *Embedded Data Stewards* are often attached to a specific research unit/department and directly involved in the research being carried out. They have a solid knowledge of the working methods specific to their unit, which enables them to provide operational support to researchers, for example in the creation of code or scripts for data analysis, and also to translate institutional policies into operational measures in line with the realities of the field.

Generic Data Stewards are located within a helpdesk or at faculty level, possessing general knowledge of research data management as well as an overview of the institutional situation, enabling them to answer all kinds of questions and redirect researchers to other competent institutional services if necessary.

Mission(s)	• Promote best practices in research data management within the institution and among researchers; provide advice, support and expertise in data management practices to researchers • Participate in the development of institutional projects regarding research data management and openness
Examples of tasks	• Promote strategies, institutional policies and best practices for managing and opening up research data • Support researchers in managing research data throughout the research life cycle • Inform researchers about existing infrastructures/tools for managing and opening up research data • Co-develop training materials for research data management skills • Relay researchers to other experts (technical, DPO, ethical, legal, ...) as needed
Life cycle stages • Plan and design • Collect and capture • Collaborate and analyse • Manage, store and secure • Archive and share • Promote for reuse	

Table 2: Data Steward job description

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
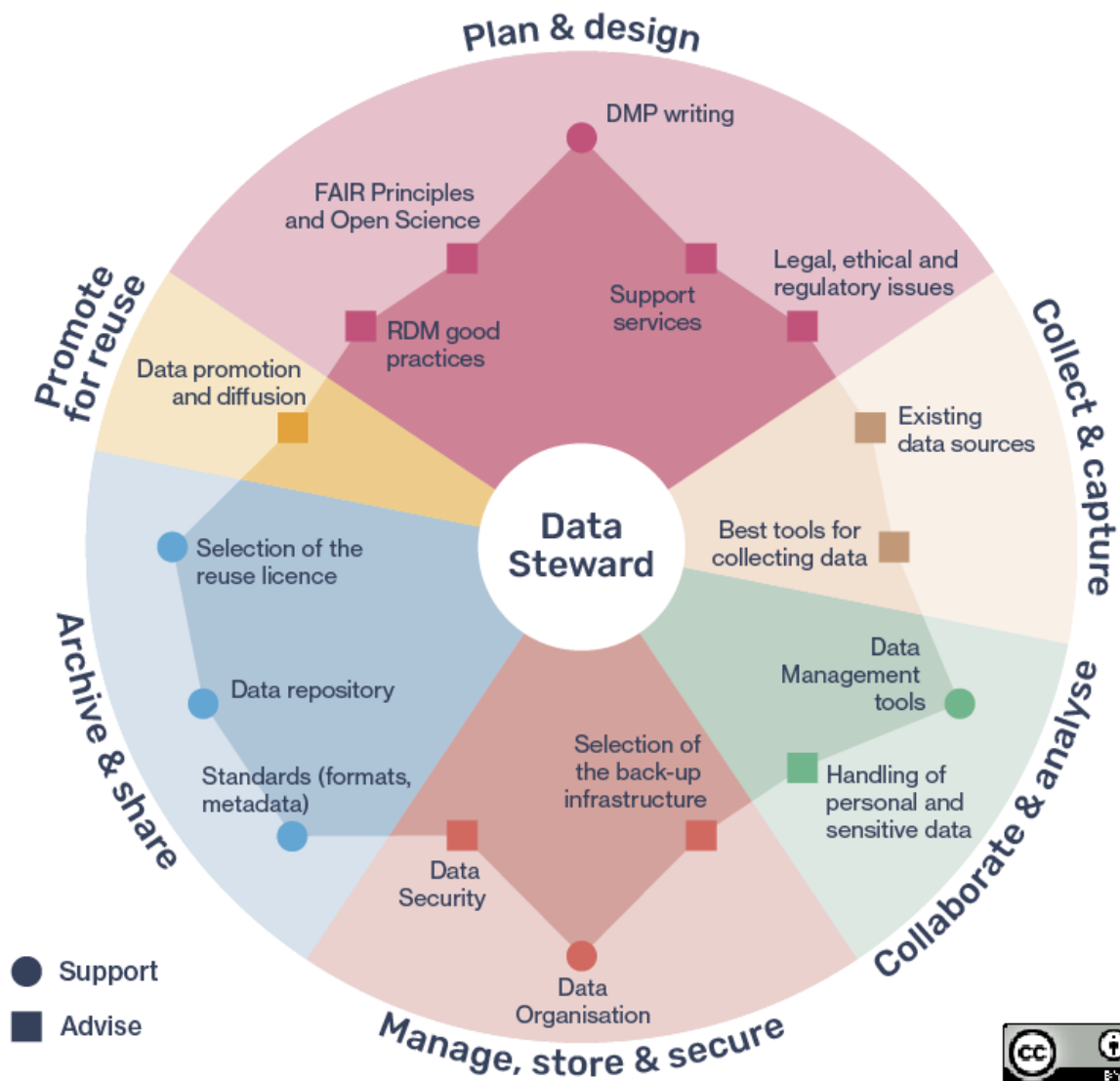02.08.2023

5

Figure 3: List of tasks performed by Data Stewards at each stage of the life cycle and level of contribution ("advise": lowest; "support": highest). Adapted from FORS life cycle (2023)

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

6

3.3. Data Librarian

Other name(s)	*Data Services Librarian, Research Librarian*
Keywords (tags)	Library and research data management specialist, researcher support
Definition	

As information professionals specializing in research management, *Data Librarians* are responsible for implementing support and training services and solutions for research data management and disseminating best practices to researchers and institutional staff.

Mission(s)	• Use librarian skills to manage research data (copyright, preservation, etc.) • Facilitate the management of research data for researchers

Examples of tasks	• Prepare and lead training sessions on research data management
	• Carry out communication, awareness-raising and community involvement initiatives
	• Support grant writing, for example as Open Science specialists (see Féret and Cros 2019)
	• Support research data management, including:
	  ○ Help draft and/or edit and evaluate DMPs
	  ○ Help manage an institutional data repository
	  ○
	  ○ Help researchers select the right license for dataset access and reuse
	  ○ Identify/locate reusable data

Life cycle stages • Plan and design • Collect and capture • Manage, store and secure • Archive and share • Promote for reuse	

Table 3: Data Librarian job description

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

7

swissuniversities

3.4. Data Scientist

Other name(s)	/
Keywords (tags)	Programming, statistics, data analysis
Definition	
Data Scientists advise and support researchers in analyzing and modeling their data, answering specific scientific questions or developing algorithms that automatically perform certain difficult tasks. They are experts in machine learning, statistics, computer vision, natural language processing, or other related disciplines and skills. *Data Scientists* can work in academia, business or industry. Their functions are highly dependent on the context in which they work.	
Mission(s)	- Advise and support researchers in the collection, processing, analysis and application of their data - Develop and implement appropriate tools
Examples of tasks	- Develop models and/or algorithms/scripts - Produce and apply suitable tools, for example: - Data collection forms - Processing scripts - Visualization and analysis tools - Quality control and formal validation tools
Life cycle stages - Collect and capture - Collaborate and analyse	

Table 4: Data Scientist job description

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

8

3.5.Data Curator

Other name(s)	/
Keywords (tags)	Data collection, selection, processing, enrichment, documentation and archiving
Definition	
Data Curators carry out data processing tasks with the perspective of long-term preservation and reuse. Their activities enable data to be discovered and accessed, including ensuring that data coherence and quality, so that it can be accessed andand reused over time (Cragin, undated). These activities include data collection, selection, enrichment, preservation, maintenance and archiving. These activities are often carried out in the context of a data archive by the curators within a specific infrastructure (Johnston 2018).	
Mission(s)	• Support researchers in data collection, evaluation and enrichment • Assist in the selection of high value-added data and data preparation for transfer to the archiving system/repository for preservation and sharing
Examples of tasks	• Visualize data, help evaluate and enrich the metadata • Advise researchers on the selection of data for preservation • Advice on selecting the most suitable formats for preservation • Complete/ enrich metadata • Validate data quality before archiving and publication
Life cycle stages • Collect and capture • Collaborate and analyse • Manage, store and secure • Archive and share	

Table 5: Data Curator job description

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

9

3.6.Data Archivist

Other name(s)	/
Keywords (tags)	Data archiving and preservation
Definition	

Data Archivists carry out the tasks required for the long-term preservation of research data. *Data Archivists* are responsible for archiving research data indiscriminately: they do not deliberately select data but treat it as an asset in its own right which it is important to safeguard.

Mission(s)	- Use archival skills to manage research data (copyright, preservation, etc.) - Facilitate data preservation by researchers
Examples of tasks	- Support researchers in preparing data for archiving - Implement institutional and/or national standards for the preservation of research data as the institution's information assets, in collaboration with the institution's authorities and administration - Help researchers select the right license for dataset access and reuse
Life cycle stages - Manage, store and secure - Archive and share - Promote for reuse	

Table 6: Data Archivist job description

4. Conclusion

With the results presented here, we better characterize the roles and responsibilities of key data-related jobs. We have also identified the life cycle stages to which they contribute. We have studied the Data Steward profile in more detail by identifying his/her specific tasks at each stage of the data life cycle, as well as the level of contribution for these tasks. In the next step of the SwissDS-ENV project, we will develop a grid of competencies for Data Stewards, that is as detailed as possible, based on the first results presented in the current document. This grid will help to identify the exact needs and to specify the objectives of the certificate course in Data Stewardship.

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

10

swissuniversities

Bibliography

CRAGIN, Melissa H, HEIDORN, P Bryan, PALMER, Carole L and SMITH, Linda C, [undated]. An Educational Program on Data Curation. p. 1.

FERET, Romain, and CROS, Marie, 2019. "The Embedded Research Librarian: A Project Partner." *LIBER Quarterly: The Journal of the Association of European Research Libraries* 29 (1): -120. https://doi.org/10.18352/lq.10304.

FORS, 2023. Introduction to SWISSUBase. Presentation at the University of Lausanne. June 6, 2023.

JOHNSTON, Lisa R, CARLSON, Jacob, HUDSON-VITALE, Cynthia, IMKER, Heidi, OLENDORF, Robert and STEWART, Claire, 2018. How Important Are Data Curation Activities to Researchers? Gaps and Opportunities for Academic Libraries. *Journal of Librarianship and Scholarly Communication*. 2018. Vol. 6, p. 25.

VERHEUL, Ingeborg, IMMING, Melanie, RINGERMA, Jacquelijn, MORDANT, Annemie, van der PLOEG, Jan-Lucas and PRONK, Martine. 2019. "Data Stewardship on the map: A study of tasks and roles in Dutch research institutes". Zenodo. https://doi.org/10.5281/zenodo.2669150.

WUILLEMIN, Claire. *État des lieux des services de soutien à la gestion des données de recherche à l'Université de Genève.* 2022 https://archive-ouverte.unige.ch/unige:162963

Definition and profile of Data professions
SwissDS-ENV Project - Action 1: Definition and profile of Data Stewards
02.08.2023

11