

Data Stewardship and Soft Infrastructures for Research Data

Liise Lehtsalu

RDA Data Steward Career Tracks WG/ RDA Professionalising
Data Stewardship IG

Data stewards are professionals who guide researchers (and research organizations) in managing, storing, and sharing research data to meet FAIR (and CARE) data principles. Data stewards can be in advisory roles, or embedded in research groups with more hands-on roles. Data stewards play a key role in the implementation of Open Science principles and sustainable research data management.

2016

FAIR data principles

2018

Turning FAIR into reality report (EC Expert Group FAIR Data)

TU Delft data stewardship program

2020

Open Science Policy Platform report

“It is irresponsible to support research but not data stewardship” - Barend Mons

Professionalising Data Stewardship IG

2021

EC’s Horizon programme makes data sharing and DMPs mandatory

2022

sonrai Irish Data Steward Network

Skills4EOSC project (2022-25)

University of Vienna Data Stewardship certificate program first intake

2025

EOSC SRIA + MAR -> data stewardship must to be institutionalized at European level

RDA Data Steward Career Tracks WG

Data Stewardship Services: Organizational Models

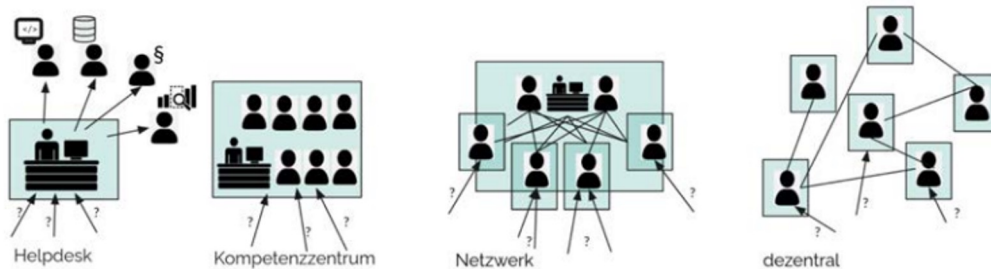


Abbildung 9: Mögliche Umsetzungen von Data Stewardship an einer Institution.

Eva Seidlmayer et al., *Forschung Unterstützen - Empfehlungen für Data Stewardship an Akademischen Forschungsinstitutionen : Ergebnisse des Projektes DataStew* (2023), <https://repository.publisso.de/resource/frl:6441397>.

- A number of different organizational models exist
 - help desk, competence centre, network, decentralized service (Seidlmayer et al. 2023)
 - infrastructure/ repository-centered service, centrally coordinated network of data stewards, centrally advised network of researchers as part-time data stewards (Rousi, Bohem, Wang, 2024)
- Organizational models depend on organizational strategies, available resources incl staff, skills and competences of staff, context of service development
- Organizational model shapes the role and functions of the data stewards

Data Stewardship Services: Scope of Service

No two Services are (exactly) alike?

- Decentralized services more likely to be general/multi-domain
- Federated services more likely to be provided at national/ international level

Organizational structure and context are shaping the scope of data stewardship services

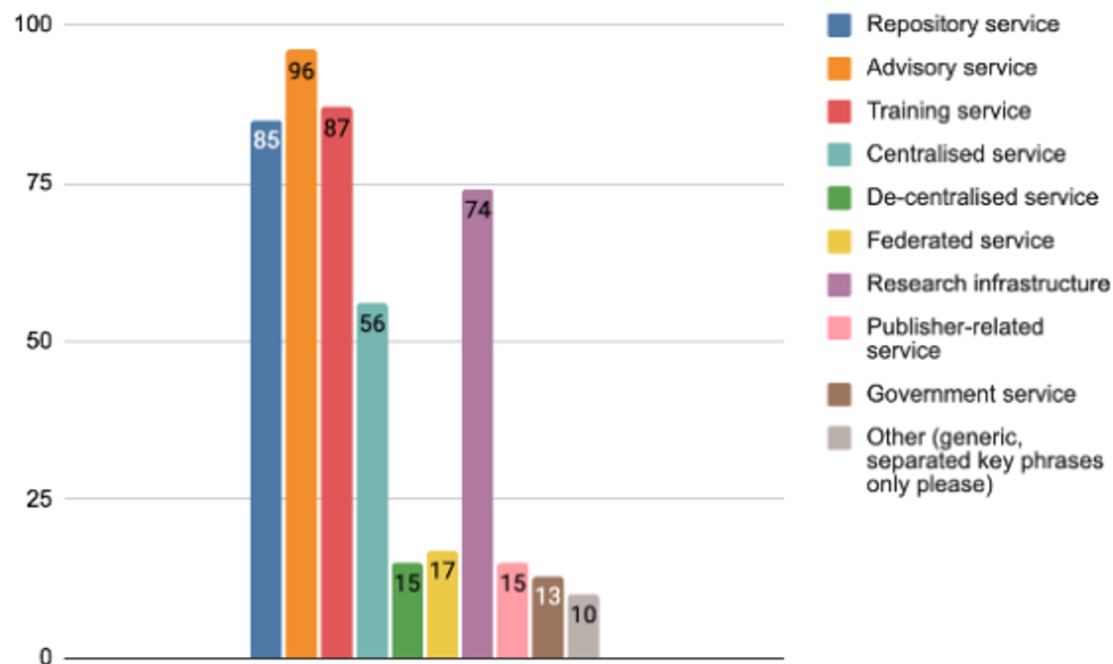
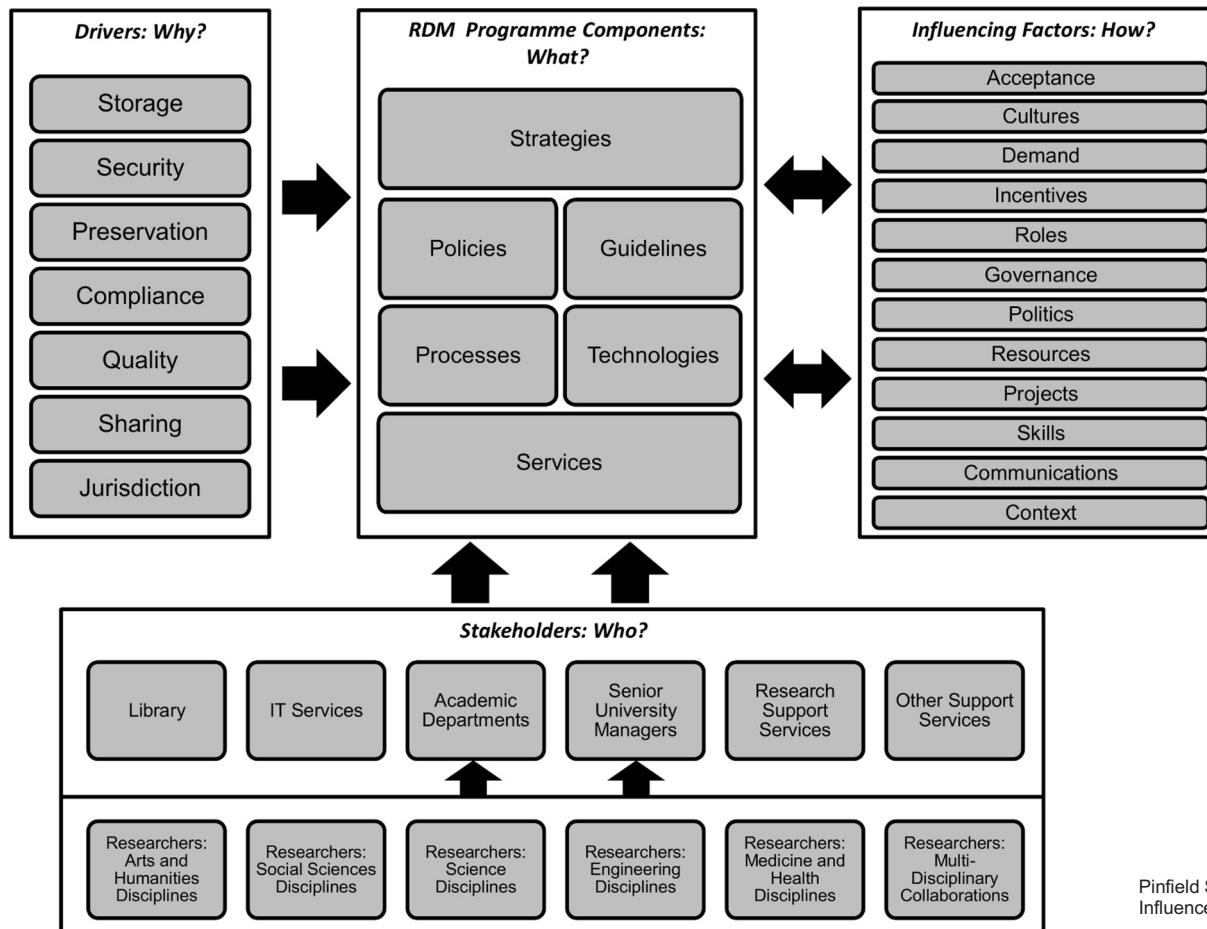


Figure 8. Service keyword selections for all responses (n=136).

The context for Data Stewardship Services



Complex drivers and influencing factors

- Internal AND external drivers and influencing factors
- Existing and planned infrastructures and services
- Current and future staff availability

Data stewardship services/ RDM services respond to/ negotiate with / translate between different drivers, influencing factors, infrastructures, services and stakeholders.

Data Stewardship as Soft Infrastructure

Hard infrastructure = roads, bridges, energy grids, water supply, technical research infrastructure

Social infrastructure = organisations, places and spaces that enable individuals communities to create social connections (e.g. libraries, community-spaces, parks)

Soft infrastructure = intangible institutions and systems needed to develop and maintain hard infrastructures (e.g. policies, regulations, support and service for hard infrastructures, human capital, culture and environment)

In RDM context:

Hard infrastructure: storage, repositories, tools, data, code/ software

Social infrastructure: research performing organizations, professional associations, federating organizations

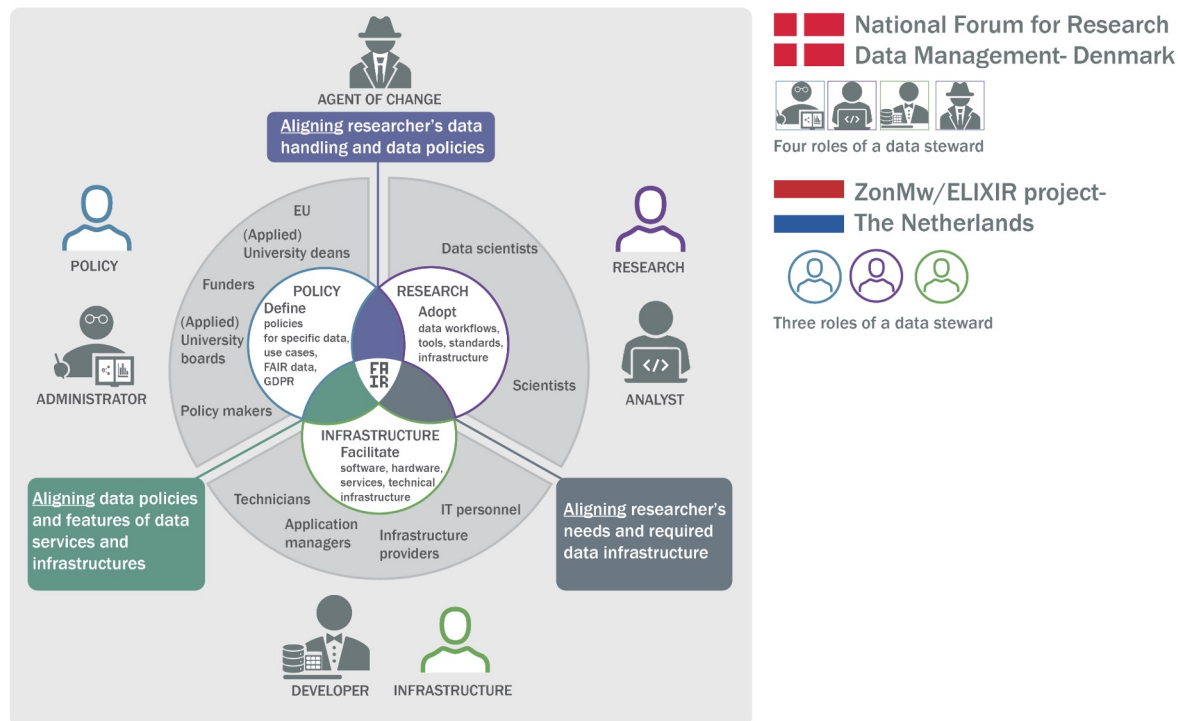
Soft infrastructure: policies, standards, data services, data stewardship services, data stewards

Why do soft infrastructures matter in RDM?

- Strategies and frameworks define how RDM infrastructures (e.g. storage, repositories) are developed
- Policies and regulations define how RDM infrastructures are operated and managed
- Standards suggest good-practices for how (meta)data are managed, how storage and repositories are built and how they ingest data, etc.
- **RDM professionals build RDM infrastructures, maintain those infrastructures, and advise and train users on how to use those infrastructures**
- RDM professionals develop/build the soft infrastructures (policies, services)

Soft infrastructures facilitate the use of hard infrastructures, their efficiency and sustainability.

Who are the Data Stewards?



- Data steward are professionals who guide researchers (and research organizations) in managing, storing, and sharing research data to meet FAIR (and CARE) data principles.
- Data steward roles perform a range of function:
 - technical/infrastructure focused
 - data/ code management focused
 - policy focused
 - advisory and training focused
- Skills and competences needed to perform data stewardship role(s)
 - Data stewardship skills
 - Technical skills
 - Transversal (soft) skills

(Some) Current challenges

- Access to data stewardship services
 - Service maturity varies greatly across organizations
- Business cases for data stewardship services
- Lack of clarity on data steward role descriptions
 - At national, European, international levels work towards convergence, definition through definition of skills and competencies training and certificate programs
 - Limits to convergence (national contexts, professionalization of adjacent roles (e.g. librarianship roles), etc.)
- Lack of defined career paths
 - Need to consider not only how to recruit data stewards but also how to retain data stewards
- Rewards and recognition

Bibliography

- Bill Ayres et al., *RDA Professionalising Data Stewardship - Current Models of Data Stewardship: Survey Report* (2022), <https://doi.org/10.15497/RDA00075>.
- Lillian Barratt, 'Social infrastructure' in two minutes, The British Academy Blog, <https://www.thebritishacademy.ac.uk/blog/social-infrastructure-in-two-minutes/>
- Clara Boavida et al., *Inquérito nacional sobre o papel dos profissionais no apoio à gestão de dados de investigação* (Portuguese Data Stewards Network, 2025), <https://doi.org/10.5281/zenodo.17662515>.
- Alexander Gruber et al., "Kompetenzen von Data Stewards an Österreichischen Universitäten," *Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare* 74, no. 1 (2021): 12–32, <https://doi.org/10.31263/voebm.v74i1.6255>.
- Karsten Kryger Hansen et al., *Data Steward: Minimum Viable Skills Profile*, August 7, 2025, <https://doi.org/10.5281/ZENODO.16760455>.
- Kavanagh SA, Hawe P, Shiell A, Mallman M, Garvey K. Soft infrastructure: the critical community-level resources reportedly needed for program success. *BMC Public Health*. 2022 Mar 2;22(1):420. doi: 10.1186/s12889-022-12788-8. PMID: 35236315; PMCID: PMC8889705.
- Stephen Pinfield, Andrew M Cox, Jennifer Smith (2014) Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. *PLoS ONE* 9(12): e114734. <https://doi.org/10.1371/journal.pone.0114734>
- Antti Mikael Rousi, Reid Isaac Bohem, and Yan Wang, "Data Stewardship: Case Studies from North American, Dutch and Finnish Universities," *Journal of Documentation* 80, no. 7 (2024): 306–24, <https://doi.org/10.1108/JD-12-2023-0264>.
- Eva Seidlmayer et al., *Forschung Unterstützen - Empfehlungen für Data Stewardship an Akademischen Forschungsinstitutionen : Ergebnisse des Projektes DataStew* (2023), <https://repository.publisso.de/resource/frl:6441397>.

Please be in touch:

liise.lehtsalu@gmail.com

RDA Professionalising Data Stewardship IG:

<https://www.rd-alliance.org/groups/professionalising-data-stewardship-ig/activity/>

RDA Data Steward Career Tracks WG: <https://www.rd-alliance.org/groups/data-steward-career-tracks-wg/activity/>