Managing the Data Landscape: A Case for Data Stewardship

Dierkes, Jens | Glasenapp, Yvana | Kalová, Tereza | Lindstaedt, Birte

OPEN SCIENCE FESTIVAL

MEET. SHARE. INSPIRE. CARE.

#OSF2023DE

Jens Dierkes

Yvana Glasenapp

Tereza Kalová

Birte Lindstädt

Gefördert vom

Bundesministerium für Bildung und Forschung

Managing the Data Landscape: A Case for Data Stewardship

Contents

Time	Action
14:30-14:50	Welcome and icebreaker
14:50-15:10	DataStew
15:10-15:50	Group work on profiles and discussion
15:50-16:15	Break
16:15-17:15	Group work on Data Stewardship models and discussion
17:15-17:30	Wrap up

Contents - DataStew

Project in a Nutshell

DataStew:
Data Stewardship in German Academic Research Institutions - Investigation of the Status Quo and Recommendations for Training and Profiling

- funded by Ministry of Science and Education BMBF Feb. - Nov. 2022 (10 month)
- projectteam:

USB Köln

Dr. Jens Dierkes (PI)
Fabian Hoffmann
Ralf Depping

ZB MED

Birte Lindstädt (PI)
Dr. Eva Seidlmayer
Prof. Konrad Förstner

- report DOI 10.4126/FRL01-006441397

Project in a Nutshell

Aims:

- Situation and needs analysis
  - Analysis and classification of the current national RDM landscape in Germany
  - Needs of practitioners and responsible persons
  - Identify individual structures
- Recommendations for policy and practice
  - on structures and contents of education and training opportunities
  - on the establishment of a data stewardship structure at universities and non-university research institutes

Curdt, Constanze, Jens Dierkes, Kerstin Helbig, Birte Lindstädt, Jens Ludwig, Janna Neumann, and Uta Parmaksiz. "Data Stewardship im Forschungsdatenmanagement - Rollen, Aufgabenprofile, Einsatzgebiete: Überblick: 11. DINI/nestor Workshop, 16. und 17.11.2020." *Bausteine Forschungsdatenmanagement*, no. 3 (October 21, 2021): 70–81. https://doi.org/10.17192/bfdm.2021.3.8347.

Steps and Methods

1. Literature and information research/analysis
2. Job advertisement research/analysis (N=60)
3. Expert interviews (N=9)
4. Thesis generation
5. Focus group discussions (2 rounds, N=23)
6. Data Steward profiling
7. Derivation of recommendations for action

Data Stewards

infrastructure

political consultants

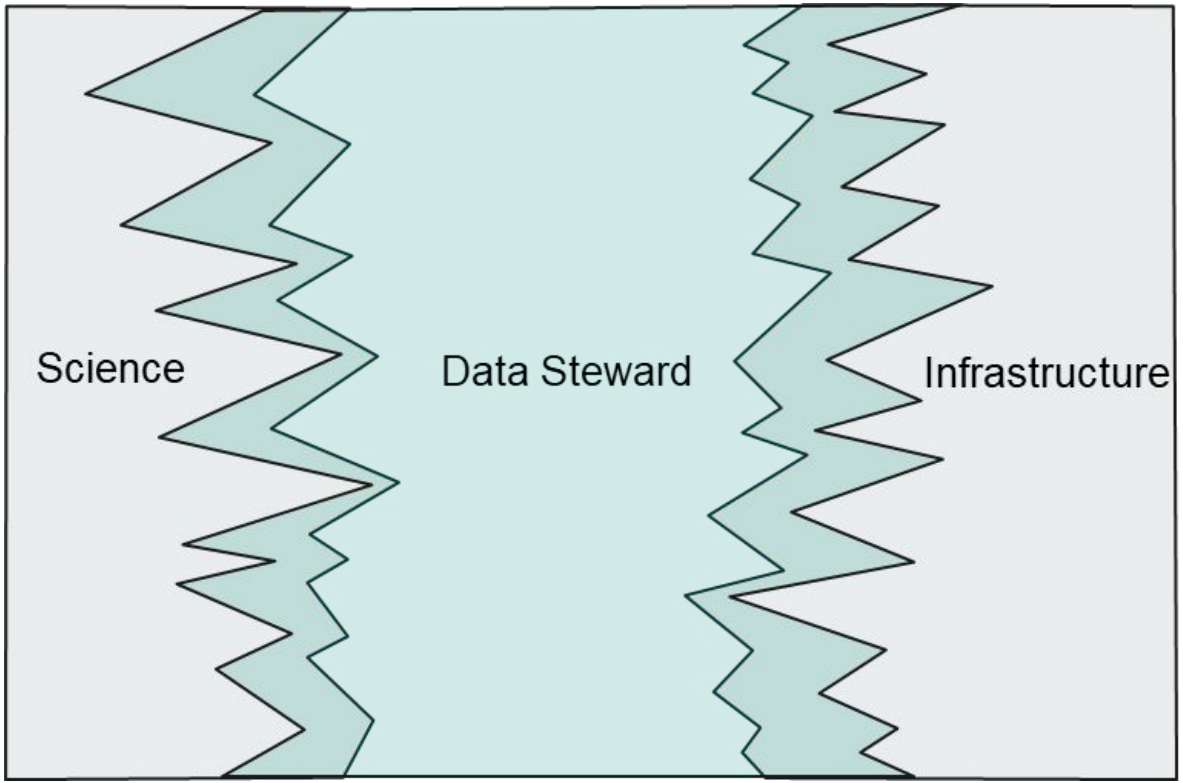
researchers

education

Overview Project Results

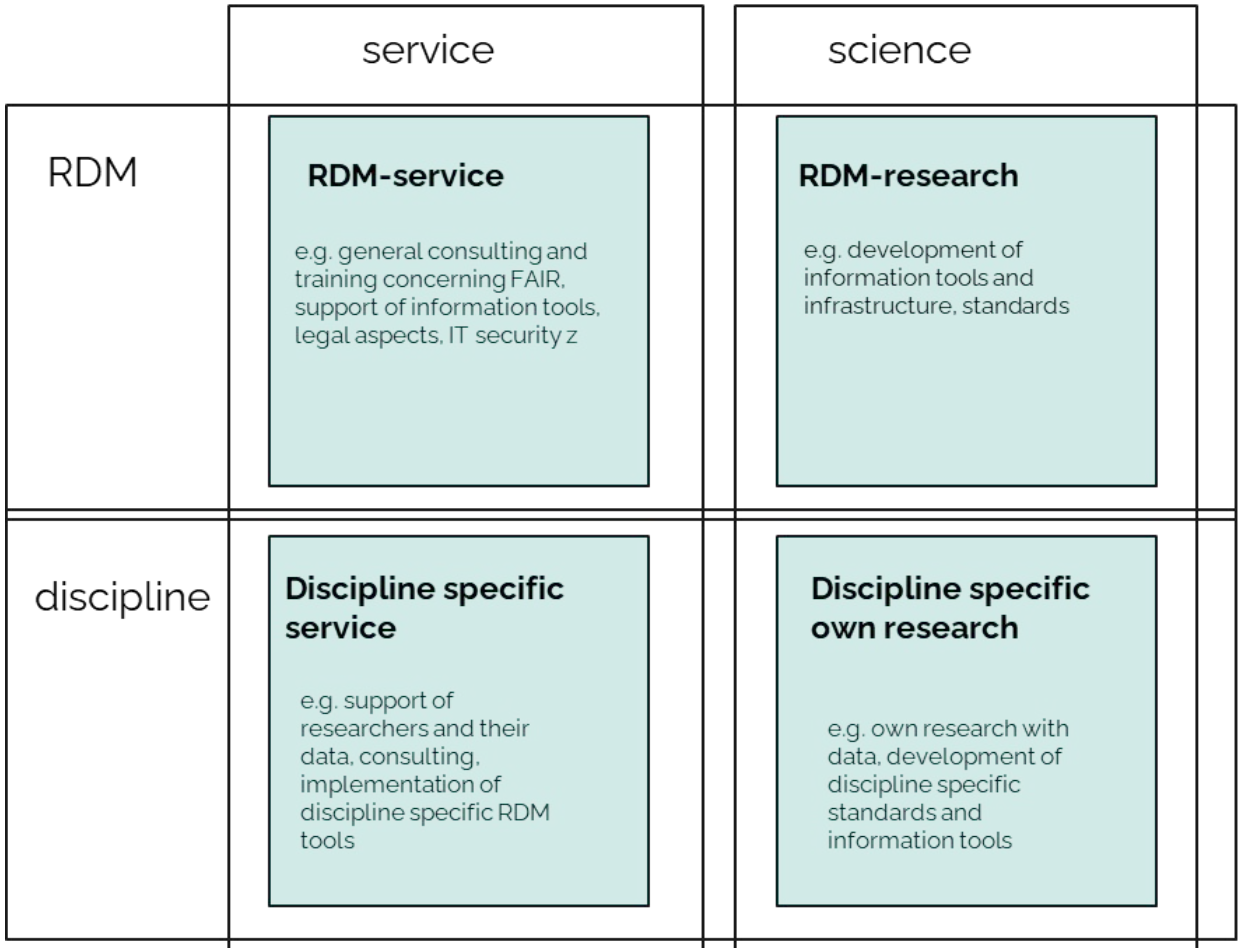
- There is no such thing as "the" Data Stewardship
- Find local answers to local needs: assemble a team from profiles using a modular approach
- Continuous learning
- Recommendations
  - Instead of generic data steward training:
    Basic training + specializations
  - More RDM across curricula
  - Clarify quantitative needs of the different data steward profiles

Bridging the Gap between Science and Infrastructure Providers



Data Stewards: researchers or infrastructure staff?

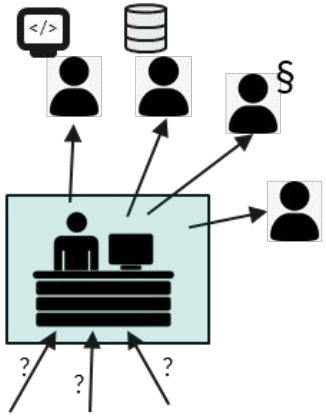
Tasks of Data Stewards

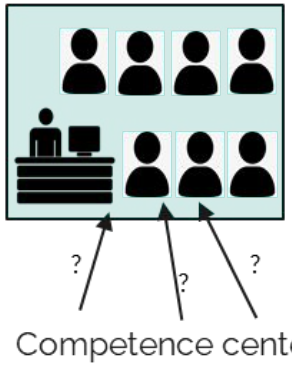


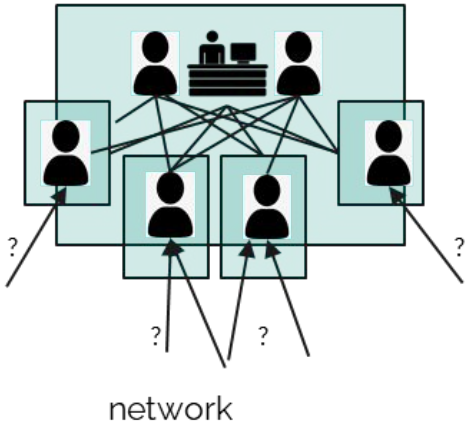
coordination tasks

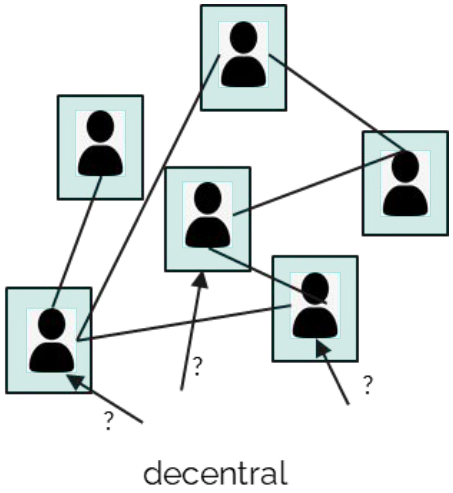
e.g. networking in the institution, regional or national

Location in the Institution

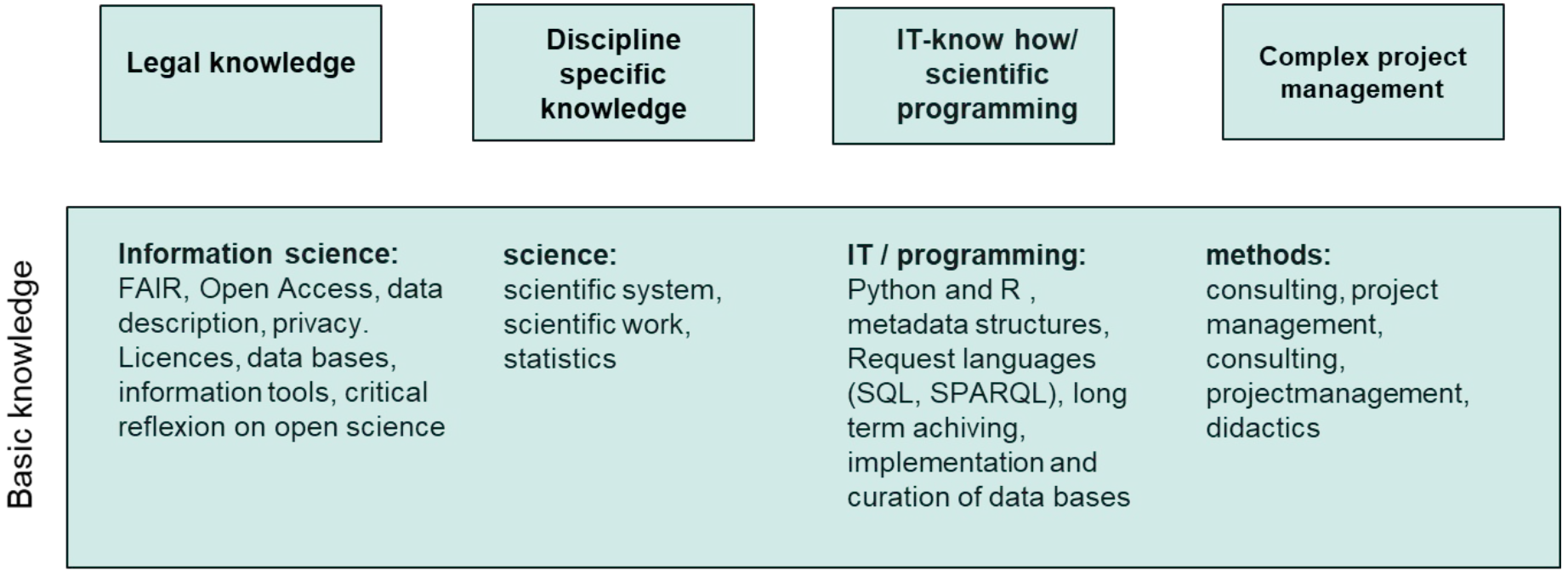








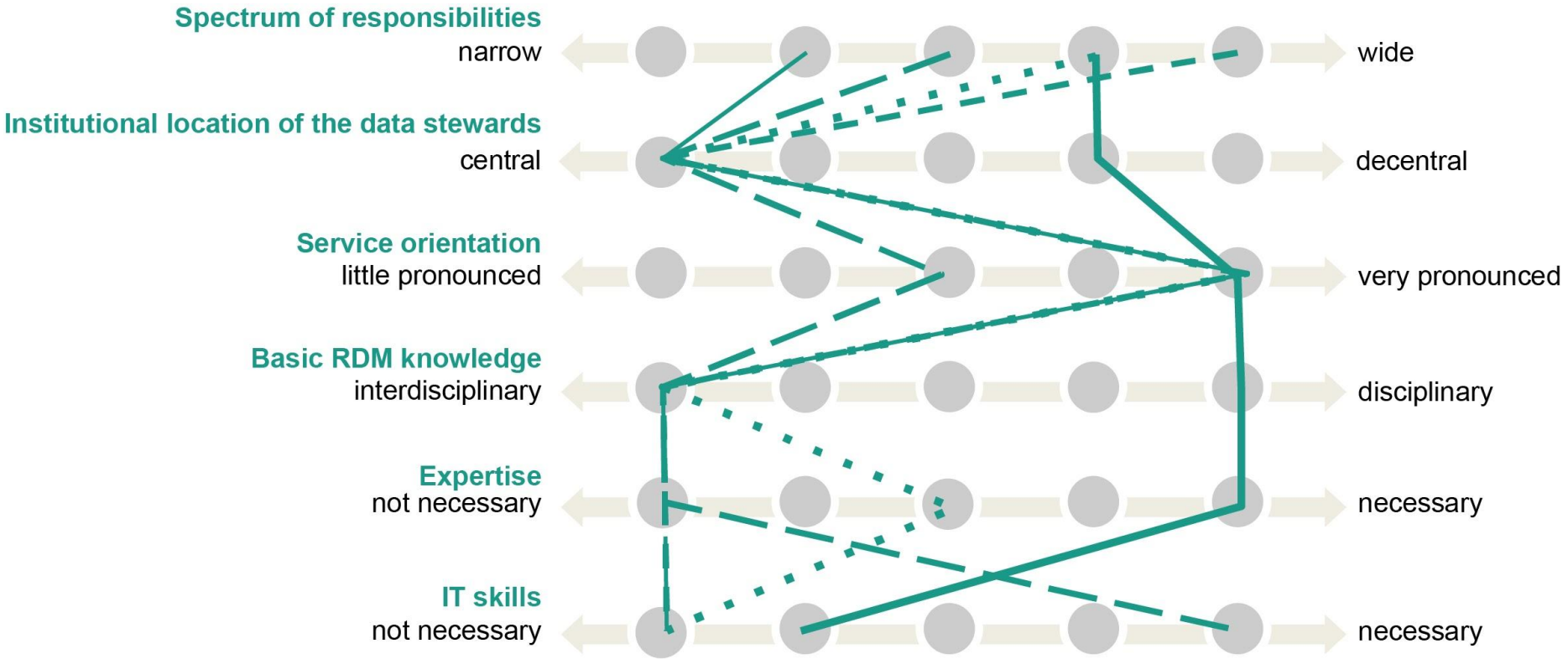
Competence Building



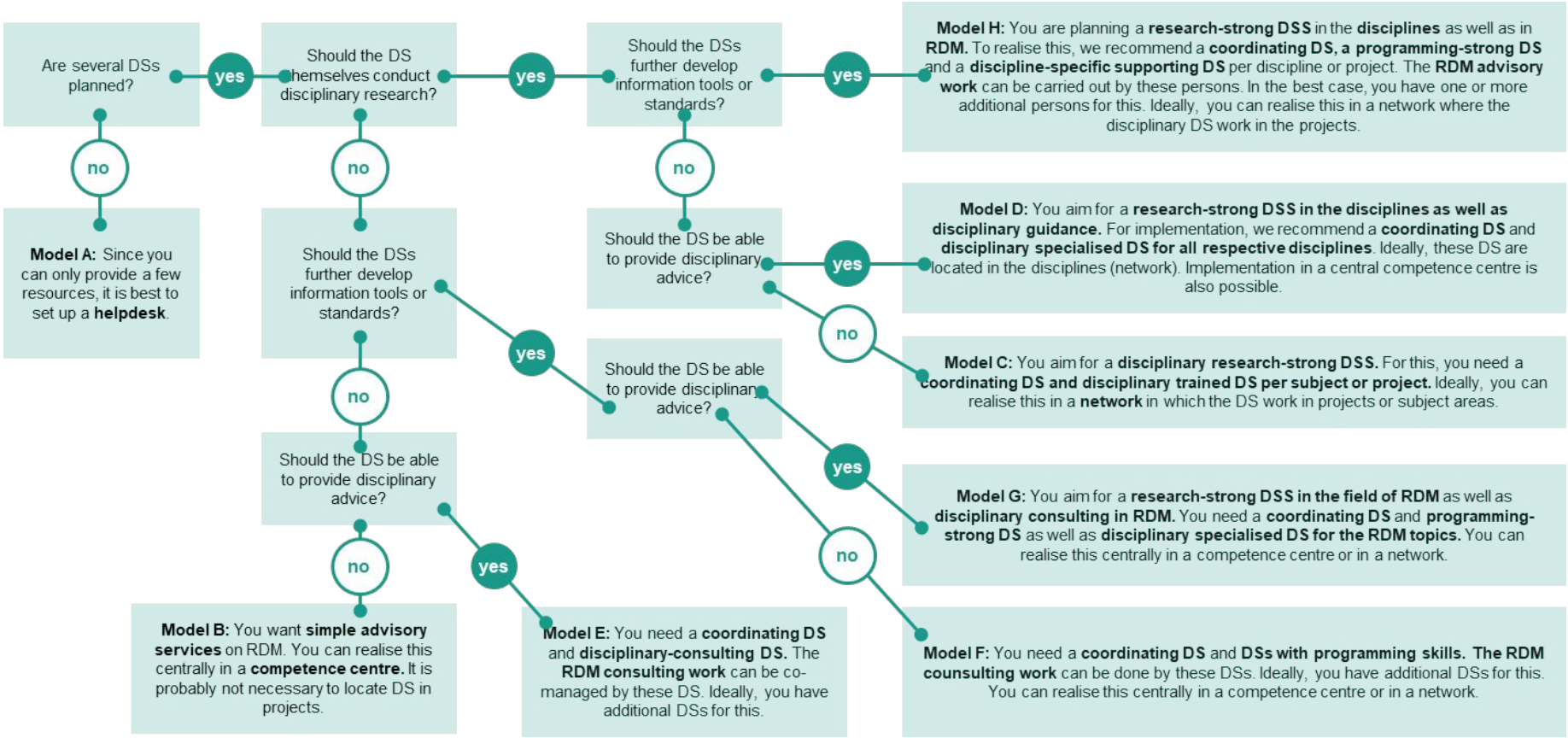
Data-Steward-Profiles

- Data stewardship is always a team effort
- Five prototypical profiles of data stewards:
  1. Data stewards as generalists
  2. RDM-consulting data stewards
  3. Discipline specific-supporting data stewards
  4. Data stewards as coordinators
  5. Infrastructure-oriented data stewards
- Six dimensions form the profiles

Data-Steward-Profiles: Overview



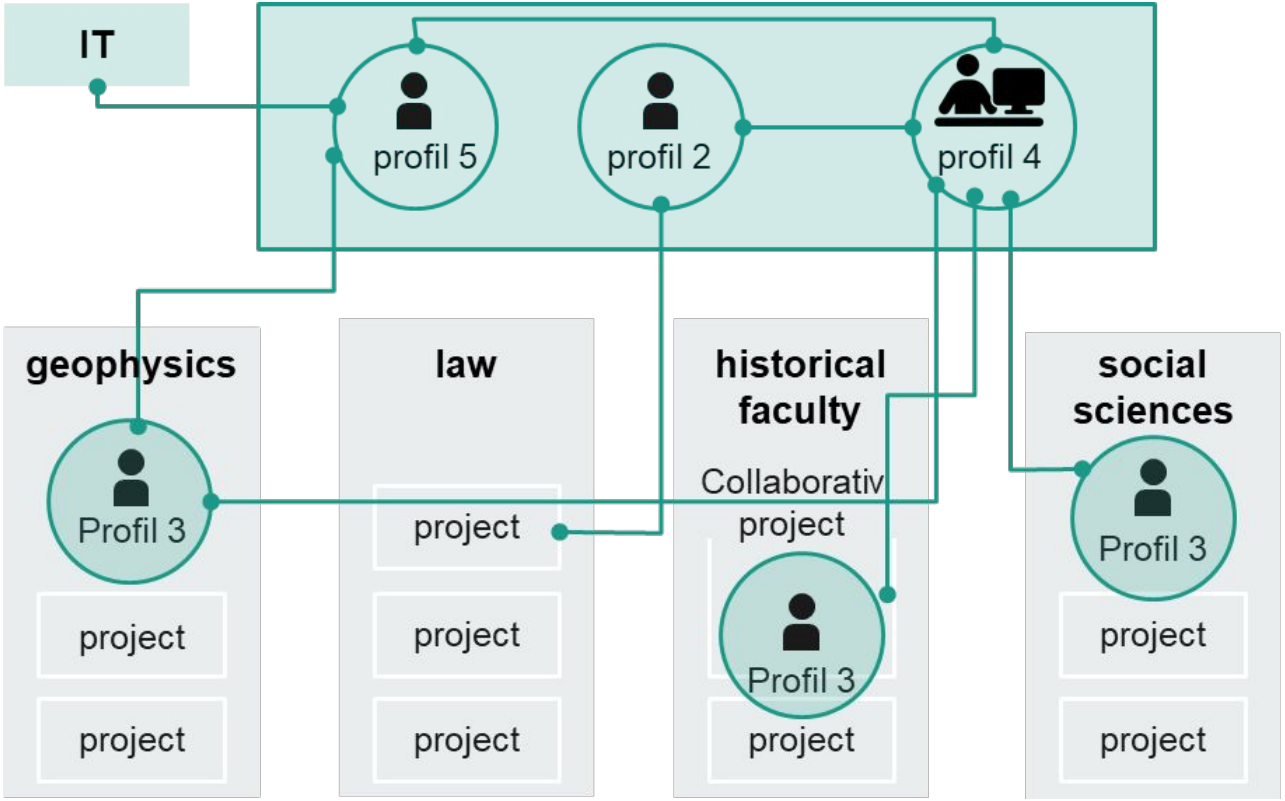
Implementation in Institutions: Decision Tree



Example: **Model G**

Discipline specific support and infrastructure developement

- RDM-consulting (2)
- Discipline specific support (3)
- networking(4)
- infrastructure developement (5)



DataStew **Report**

https://repository.publisso.de/resource/frl%3A6441397

Eva Seidlmayer, Fabian Hoffmann, Jens Dierkes, Birte Lindstädt, Ralf Depping, Konrad U. Förstner (2023): Forschung unterstützen: Empfehlungen für Data Stewardship an akademischen Forschungsinstitutionen, Köln, doi: 10.4126/FRL01-006441397.

Group Work: Data Stewardship Profiles

Step 1: Choose one of the profiles

Step 2: go through the guiding questions

Step 3: Presentation of results and group discussion

Group work: Data Stewardship Profiles - Guiding Questions

- Are the **profiles** explained in an understandable way?
- Is the naming of the **dimensions** understandable?
- Are the **dimensions** relevant?
  - Is one of the dimensions to be changed?
  - Is there a dimension to add?
- Are the **profiles** sufficient?
  - Is one of the profiles to be changed?
  - Is there a profile to add?

Group Work: Data Stewardship Models

Step 1: Choose a research institution (a real example or fictional)

Step 2: Write down the key facts for this institution:

- Type of funding (public or private)
- Number of students
- Number of employees
- Policies (Open Access, RDM, …)
- RDM-related infrastructure (Repositories, electronic lab notebook)
- Already existing RDM/Data Steward staff

Step 3: Go through the decision tree and choose a DSS-model

Step 4: Present your results and discuss with the group

Group Work: Data Stewardship Models - Guiding Questions

- Is the **decision tree** helpful and practicable?

- Would you suggest any changes based on your discussions today?

Example 1 (Tereza): A large university is looking to start a data steward program and hire several embedded data stewards. The program is coordinated by the data steward coordinator at the university library who had offered generic RDM support to the whole university for the past five years. The goal of the university is to further develop the existing services based on the discipline-specific demands of researchers. The university expects the data stewards to focus broadly on designing support services, training and contributing to the further development of IT services for research in order to achieve this goal.

Key facts – university:

- Funding: public
- Students: 40.000
- Staff: 10.000
- Policy on Open Access and RDM available
- Some generic RDM training and support available – provided by the head of RDM /data steward coordinator from the library in collaboration with the IT and legal department; focus on writing data management plans)
- No institutional data repository

Example 2 (Yvana) A middle sized university wants to set up a new data stewardship concept. The spectrum of research fields covers humanities and natural sciences, with a main focus on the natural sciences like physics and engineering. The new data stewardship concept should address the needs of all research fields. The university has an established RDM-service, which works cooperatively with the library and the computer centre. This service offers general training and support. Individual solutions like the setup and implementation of a new software cannot be supported. The new data stewardship concept should offer solutions for field-specific problems, trainings and support the setup of new infrastructures.

Key facts – university:

- Funding: public
- Students: 25.000
- Staff: 5.000
- Policy on Open Access and RDM available
- Some generic RDM training and support available – provided by RDM-team (Research service, library, IT and legal department)
- Institutional data repository is available