TOWARDS A COMMUNITY-ENDORSED DATA STEWARD DESCRIPTION FOR LIFESCIENCES

Mijke Jetten[1], Inge Slouwerhof[1], Salome Scholtens[2], Jasmin Böhmer[3], Marije van der Geest[2], Christine Staiger[4] & Celia W.G. van Gelder[4]

(1) Radboud University, (2) UMC Groningen, (3) UMC Utrecht, (4) DTL (Dutch Techcentre for Life Sciences)/ELIXIR-Netherlands

Contact: Mijke Jetten, m.jetten@ubn.ru.nl

1 BACKGROUND

To create FAIR data management, it is important to build sufficient, high-quality data steward expertise in organisations.

Based on an analysis of competency frameworks and initiatives (e.g. EOSCpilot, EDISON, Purdue and HANDS), this project aims to professionalise data stewardship, including a function description and training.

The outcomes of the project can be found at https://zenodo.org/communities/nl-ds-pd-ls.

2 RESULTS

This ZonMw project (Aug. '18 - Sept. '19) has delivered a function description for three data steward roles: policy, research and infrastructure.

For each data steward role, competence areas with tasks were defined, i.e. 1) policy/strategy, 2) compliance, 3) alignment with FAIR, 4) services, 5) infrastructure, 6) knowledge management, 7) network, and 8) data archiving. Table 1 shows a section of the matrices (full version via Zenodo). Tasks were translated into learning objectives based on Bloom (example shown in table 2).

In two workshops (June '19) with Dutch data stewards (± 60 participants), existing and desired training was mapped. These mappings are included in the final report.

3 RECOMMENDATIONS

- Embed data steward roles and competencies in a formal function profile including function levels (junior, senior), and add formal job evaluation and grading
- Add discipline-specific knowledge, skills and abilities to the function description
- Develop a self-assessment tool for data stewards to assess responsibilities, tasks and competencies, combined with navigating directions to training and materials
- Use the matrices and the (self-)assessment tool to assess data stewardship roles in organisations
- Develop certified training for data stewards

The project originally focused on the lifesciences. However, the outcomes show to be relevant for other domains as well. Starting September 2019, the project will be continued in a one year NPOS (National Platform Open Science, the Netherlands) funded project on professionalising training in open science and data stewardship.

Table 1: Examples for the *policy* data steward

RESPONSIBILITIES	EXAMPLES OF ACTIVITIES/TASKS
POLICY/STRATEGY Responsible for advice on and development, implementation and monitoring of a RDM policy and strategy for the research institute, which includes the complete research data life cycle and supports FAIR data and Open Science, in alignment with the relevant stakeholders and within financial and legal constraints, within the institute and in the context of the institute. The policy is the basis for (project) DMPs.	• Develops, implements and monitors the institute's RDM policy. • Advises the institute's management on short- and long-term actions to advance RDM in the institute. • Assesses and monitors the institute's time and financial investments in relation to the institute's needs for RDM. • Explores new needs, opportunities and trends in RDM.
COMPLIANCE Responsible for compliance of the RDM policy to the Netherlands Code of Conduct for Academic Practice, the Netherlands Code of Conduct for Research Integrity and the General Data Protection Regulation (GDPR), as well as continuous alignment with legal and ethical standards.	• Ensures compatibility of the RDM policy and monitors compliance. • Contacts the institute's privacy officer, legal advisors or ethical board in case of questions regarding compliance. • Translates policies from legal/privacy officer to the institute's practice.
ALIGNMENT WITH FAIR DATA PRINCIPLES Responsible for alignment of the RDM policy to the FAIR data principles and the principles of Open Science.	• Pursues and advises on the findability (F) of data, including adequate data-infrastructure and tools, persistent identifiers and rich (discipline-specific) metadata standards. • Pursues and advises on the accessibility (A) of (meta)data to potential (re)users. • Pursues and advises on the interoperability (I) of data, including broadly applicable languages, vocabularies and other standards. • Pursues and advises on the reusability (R) of data, including documentation and licenses with the conditions for reuse, data licenses, and IP rights.

Figure 1: Data stewardship landscape and its stakeholders

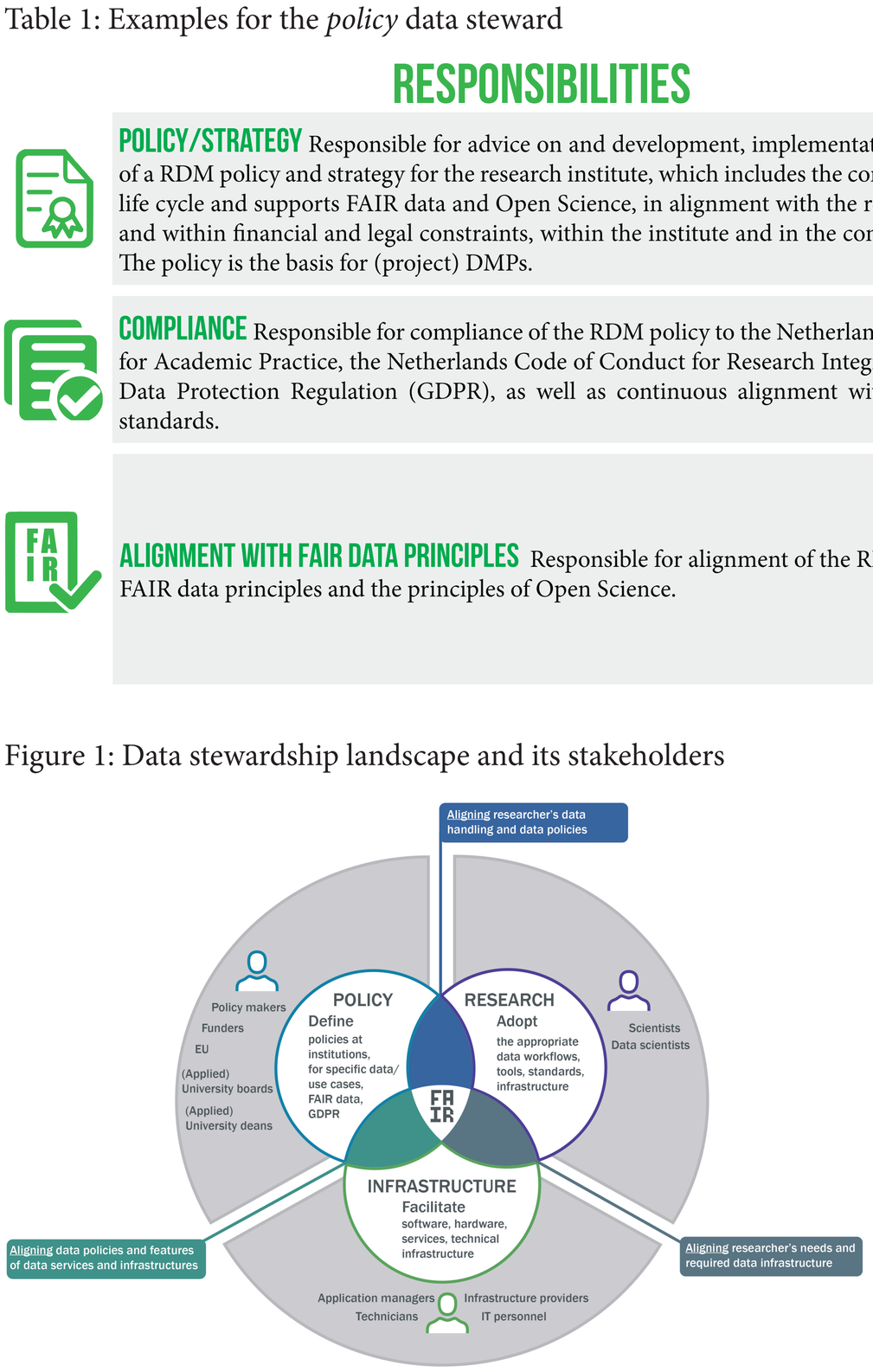

Table 2: Example of learning objectives following Bloom's taxonomy

RESPONSIBILITY

Compliance

SKILL/ABILITY

- Translate RDM policy and legislation and codes of conduct with regard to research data to practical implications and guidelines that researchers can understand.

LEARNING OBJECTIVES

- List relevant legislation, ethical principles, and codes of conduct for RDM (remembering).
- Examine and list the practical implications of legislation, ethical principles, and codes of conduct with regard to research data (analysing).
- Translate RDM policy and legislation, ethical principles, and codes of conduct with regard to research data to researchers (applying).
- Create guidelines and procedures based on legislations, ethical principles, and codes of conduct with regard to research data (creating).

This project is funded by the ZonMw Personalised Medicine Programme under dossier number: 80-84600-98-3007.