

Towards FAIR Data Steward as profession for the Lifesciences

Report of a ZonMw funded collaborative approach built on existing expertise



S. Scholtens¹, M. Jetten², J. Böhmer³, C. Staiger⁴, I. Slouwerhof², M. van der Geest¹, C.W.G van Gelder⁴

1 Genomic Coordination Centre, UMCG, Groningen

2 Radboud University Nijmegen

3 UMC Utrecht

4 Dutch Techcentre for Life Sciences (DTL), Utrecht

5 oktober 2019

This project is funded by the ZonMw Personalised Medicine Programme (Dossier number: 80-84600-98-3007), Zilveren Kruis en KWF Kankerbestrijding. Additional funding was provided by UMCG, UMCU, Radboud University Nijmegen, Radboudumc and DTL/ELIXIR-Netherlands

All output of this project (including this final report) can be found at Zenodo:
<https://zenodo.org/communities/nl-ds-pd-ls/>

Cite this report as:

S. Scholtens, M. Jetten, J. Böhmer, C. Staiger, I. Slouwerhof, M. van der Geest, C.W.G van Gelder (2019, October 5). Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. Zenodo.

<http://doi.org/10.5281/zenodo.3471708>



Content

Summary	5
Abbreviations	7
1. Background, aim, deliverables and assumptions	8
1.1 Background	8
1.2 Aim	8
1.3 Approach	8
1.4 Assumptions	8
2. Definition of data stewardship	10
3. Method	12
3.1 General	12
3.2 Function description	12
3.3 Competencies and learning objectives	12
3.4 Training	13
3.5 Workshops	13
3.6 Documentation	13
4. Results	14
4.1 Organisational context and shared responsibility	14
4.2 Function descriptions	14
4.2.1 Data steward - policy	15
4.2.2 Data steward - research	16
4.2.3 Data steward - infrastructure	16
4.3 Positioning in the organisation	17
4.4 Alignment with other data professions	18
4.4.1 Data manager	18
4.4.2 Data scientist	18
4.4.3 Data consultant	18
4.5 Competencies and learning objectives	18
4.6 Training	19
4.6.1 Existing data steward training	19
4.6.2 Desired data steward training	19
4.6.3 Outline for a self-assessment/navigation tool for data stewards	20
5. Conclusion and discussion	21
6. Recommendations and directions for further development	23
References	24
Acknowledgements	26
Funding information	26
Appendix 1 – Deliverables	27
Appendix 2 - Project team and consultation committee	28
Appendix 3 - Function and competencies matrices	30
Appendix 4 – Mapping	48
Appendix 5 - Data steward workshops	50

Appendix 6 - Overview existing and desired training reported by data stewards	51
Area 1: Policy / strategy	51
Area 2: Compliance	52
Area 3: Alignment with fair data principles	53
Area 4: Services	54
Area 5: Infrastructure	54
Area 6: Knowledge management	55
Area 7: Network	56
Area 8: Data Archiving	57
Appendix 7 - Recommended basic training for data stewards	59
Appendix 8 - Proposals FAIR Data Stewardship Knowledge Skills & Abilities Tool	62
Appendix 9 - Outreach activities	75

Summary

Data stewardship and data expertise are essential in today's research projects. Currently, the lack of consensus on the function of data stewards, including responsibilities and tasks, as well as their required knowledge, skills and abilities (KSAs), hampers adequate data steward capacity building in organisations and complicates efficient data management and handling in the various scientific disciplines.

The aim of the project is to professionalise the data steward function in the lifesciences domain in the Netherlands. The project works towards a common job description (including responsibilities and tasks) and an agreement on the required competencies (knowledge, skills and abilities, KSAs) for the data steward function. Translating these responsibilities and tasks, and required knowledge, skills and abilities into practical learning objectives will provide the basis for developing tailored training. The project aims to contribute to the wider need of increasing data steward capacity in research institutes, both in quality (by training current data stewards) and quantity (by training new data stewards).

This project shows that the data steward function can be seen as a role that differs from existing data experts and research roles. Three different, partly overlapping stakeholder fields (or working areas) of the data steward were characterised, which all have their own focus and thus different data steward role: policy, research and infrastructure. Together they form the data stewardship landscape. Each data steward role has eight competence areas:

- Policy/strategy
- Compliance
- Alignment with FAIR data principles
- Services
- Infrastructure
- Knowledge management
- Network
- Data archiving

The responsibilities, tasks and KSAs were defined per competence area and differ between the data steward roles. The data steward role is often experienced as a role that is 'in between' different disciplines and professionals. Translation between different stakeholders and professionals is seen as a key element of the function of a data steward.

The knowledge, skills and abilities that current data stewards have and that are required for the function are often developed 'on the job'. Apart from a couple of national trainings, to our knowledge, there does not exist dedicated formal data steward training. From the available training, online courses, books and other materials, the quality and appropriateness are unclear and also the findability of suitable trainings is not high. Also, for many data stewards it is unclear what knowledge, skills and abilities they are missing and should develop to advance their career.

Based on this project, we formulate the following recommendations:

- Embed the data steward roles in the formal university function profile structure. This does not have to result in a separate function profile but could also be realised by adding roles to existing function

profiles. Assess the possibility to define different function levels (junior/senior or HBO/university level) and add formal job evaluation and grading.

- Develop a self-assessment tool for defining the data steward roles and the accompanying KSAs, allowing data stewards to assess their responsibilities, tasks and competencies and combine it with a navigation toward training, courses, and learning materials. In time, develop a certified data steward training curriculum.
- For further development and implementation of the data steward function and training, align with the Open Science movement on a national and international level, including its various stakeholder/expert groups.

This project defined, to our knowledge for the first time, the data steward function for the lifesciences domain at the level of detail necessary to implement the function in research institutes and to appoint data stewards. The overview of existing training and the outline for new training give a basis for formal data steward education to further professionalise the data steward function in order to increase the capacity and quality of data stewards in the institutes and the lifesciences domain in general.

As a concluding remark it has to be noted that although the project set out to define the data steward description for the lifesciences, the current work has also high relevance for other domains, as the results of this project already have been picked up and commended outside the lifesciences.

Abbreviations

DAMA-DMBOK	The Data Management Association. Data Management Body of Knowledge (https://dama.org/content/body-knowledge)
DCC	Digital Curation Centre (http://www.dcc.ac.uk/)
DTL	Dutch Techcentre for Life Science (www.dtls.nl)
EDISON	Education for Data Intensive Science to Open new science frontiers
ELIXIR	Europe's Life Science Data Infrastructure (https://elixir-europe.org)
EOSC	European Open Science Cloud (www.eosc-portal.eu)
FAIR	Findable, Accessible, Interoperable and Reusable
GDPR	General Data Privacy Regulation
HANDS	Handbook Handbook for Adequate Natural Data Stewardship (https://data4lifesciences.nl/hands2/data-stewardship)
Health-RI	Health Research Infrastructure (www.health-ri.org)
KSA	Knowledge, skills and abilities
LCRDM	National Coordination point Research Data Management (www.lcrdm.nl/en)
LERU	League of European Research Universities (www.leru.org)
LO	Learning objective
NFU	National association of collaborating university hospitals (www.nfu.nl)
NPOS	National Platform Open Science (www.openscience.nl)
RDA	Research Data Alliance (www.rd-alliance.org)
RDM	Research Data Management
VSNU	National association of collaborating universities (www.vsnu.nl)
ZonMw	The Netherlands Organisation for Health Research and Development (www.zonmw.nl)

1. Background, aim, deliverables and assumptions

1.1 Background

Including sufficient data and data steward expertise is essential in today's research projects ([EOSC-HLEG, 2016](#); [KNAW, 2018](#)). Particularly data-intensive research areas such as genomics and personalised medicine dealing with highly sensitive data will profit from well-educated and specialised data experts. Currently, the lack of consensus on the function of data stewards, including responsibilities and tasks, as well as their required knowledge, skills and abilities, hampers adequate data steward capacity and complicates efficient data management and handling in the various scientific disciplines ([EOSC-HLEG, 2016](#)). The lack of a clear, visible function makes recognition of the data steward as an obvious and essential part of the research process difficult. Moreover, to our knowledge, there does not exist a coherent approach and dedicated training for data stewardship on the Dutch national level in the lifesciences domain nor in general. Adequate funding for the implementation of the data steward function in the Dutch research domain is still lacking. This lack of a clear and visible data steward function and adequate training, seriously limits vital developments such as recruiting procedures and defining career paths for the current and future data stewards that the various scientific disciplines require.

1.2 Aim

The overall aim of the project is to professionalise the data steward function in the lifesciences domain in the Netherlands (see project deliverables in Appendix 1).

1.3 Approach

The project has worked towards a common function description (including responsibilities and tasks) and an agreement on the required competencies (knowledge, skills and abilities) for the data steward function. This common function description can be used for recruiting new data stewards. Translating the responsibilities and tasks, and required knowledge, skills and abilities, into practical learning objectives, will facilitate the future training of data stewards, as the formulated learning objectives may serve as a basic training model. By taking this essential step in the Dutch lifesciences context, the project aims to contribute to the wider need of increasing data steward capacity in research institutes, both in quality (by training current data stewards) and quantity (by training new data stewards). This applies to data stewards at the more overarching institute and infrastructure level as well as data stewards at the department or project level. To ensure a sustainable endorsement and implementation of the project deliverables, we actively worked with a broad community of experts (including current data stewards) from the Dutch research domain, to be able to align with existing educational initiatives and translate the outcomes of the project into a formal data stewardship function description. The project included close collaboration with local, national and international projects and developments, such as the [LCRDM](#) taskforce on data stewardship ([LCRDM, 2019](#)).

1.4 Assumptions

There are some main assumptions underlying the project, as stated below. We will reflect on these assumptions in the Results section.

- The prime focus of the project is data stewardship in academia, not specifically in industry or government, although the outcomes of the project are expected to be useful in those settings too.
- Although the initial focus of the project is the lifesciences domain, by working out the basics of the data steward function, we expected to contribute to professionalising data stewardship for other scientific disciplines as well. In the course of the project it became clear that this indeed is the case.
- In contributing to professionalising data stewardship, we considered developments and concepts such as Open Science, the FAIR (Findable, Accessible, Interoperable and Reusable) data principles, the research data life cycle approach (Figure 1), and regulations such as Scientific Code of Conduct¹ and the General Data Privacy Regulation (GDPR), as they inevitably influence the data stewardship landscape.
- The project builds on and aligns with ongoing activities, existing theoretical frameworks and reports on defining data stewardship and data stewardship competencies, both from the domain-agnostic side, such as the EDISON programme ([EDISON, 2017](#)), the Data Management Body of Knowledge framework ([DAMA-DMBOK2 Framework, 2014](#)), the report from Purdue University ([Purdue University, 2016](#)) and the domain-specific side (among others [ELIXIR, 2017](#)). Practically, we will translate these existing principles and approaches into concrete data stewardship responsibilities and tasks.
- We combine the analysis of existing theoretical frameworks with a bottom-up approach, that is starting in the field of data stewardship practices in the lifesciences domain in the Netherlands. This bottom-up approach was explicitly mentioned as a potentially useful contribution and a next step in the existing theoretical frameworks and reports we build on.

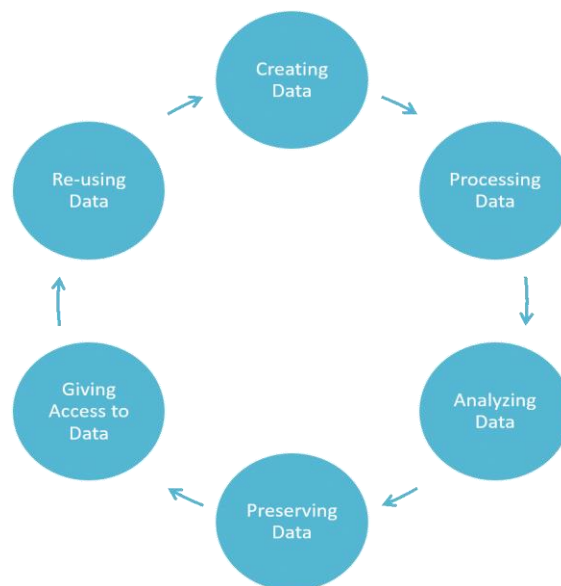


Figure 1. The research data life cycle (UK Data Archive)²

¹ Scientific Integrity Policy Netherlands.

<https://www.nwo.nl/en/policies/scientific+integrity+policy/netherlands+code+of+conduct+for+research+integrity>

² <https://qdr.syr.edu/guidance/managing/planning-data-management>

2. Definition of data stewardship

Scientists have been managing data ever since they collected specimens and shared the results of their study with their peers. The advent of computers has allowed for digital data acquisition in unprecedented amounts, and for that reason Research Data Management (RDM) nowadays is a crucial part of good scientific practice and research excellence for both individual researchers and for research organisations and institutes.

Good scientific practice includes scientific honesty and diligence (professionalism, forthrightness, transparency) in the management of data, samples, and research results. Also, external pressure is rising on researchers in general to make sure their experiments are reproducible and that the resulting data is accessible and reusable for others for a longer time. This pressure puts even higher demands on the data management practice. Data management plans will be a prerequisite for most (public) funding soon. Furthermore, researchers need skills to be able to interact with emerging national and international data infrastructures like [EOSC](#) and [Health-RI](#), which aim to provide researchers with a trusted virtual environment enabling open and seamless services for data storage, management, analysis, sharing, and reuse, across disciplines.

Although good data stewardship is crucial, it is important to note that data stewardship is not a goal in itself but a means to support the final aim of science: knowledge discovery and knowledge utilization. There are many definitions for data stewardship and a selection of them is given in Table 1. To date, there is no consensus on the proper use of the terms 'data stewardship' and 'data management' in the international community. For this project we follow the [DTL](#) definition of data stewardship:

Data stewardship is the responsible planning and executing of all actions on digital data before, during and after a research project, with the aim of optimising the usability, reusability and reproducibility of the resulting data.

It differs from data management, in the sense that data management concerns all actual, operational data-related activities in any phase of the data lifecycle, while data stewardship refers to the assignment of responsibilities in, and planning of, data management.

Good data stewardship has to adhere to the FAIR data principles and by definition implies long-term and sustainable care across multiple research cycles. For this reason, data stewardship is a collective endeavour, with actions (and competencies) needed from the individual researcher, from other scientists in the research project, from the institute, and also from the research discipline(s) and funders involved. Each organisation will implement its RDM policy tailored to its own situation and there will be a variation in how funder and institutional policies (or legislation) define roles and responsibilities, e.g. by placing differing expectations on the research team, host institution and third-party organisation (such as data centres).

Table 1. Definitions of data stewardship

Definitions of data stewardship	Reference
The responsible planning and executing of all actions on digital data before, during and after a research project, with the aim of optimising the usability, reusability and reproducibility of the resulting data	DTL
It encompasses all the different tasks and responsibilities that relate to caring for data during the various phases of the whole research life cycle	LCRDM, 2019
The process and attitude that makes one deal responsibly with one's own and other people's data throughout and after the initial scientific collaboration and discovery cycle	Mons, 2018
The entire process that deals responsibly with one's own and other people's data throughout and after the scientific discovery process	HLEG-EOSC, 2017
It refers to the long-term and sustainable care for research data. Data stewardship implies professional and careful treatment of data throughout all stages of your research project (i.e., the design, collection, processing, analysis, long-term preservation, and sharing of your research data)	NFU HANDS Handbook, 2018
The formalised management and oversight of an organisation's data assets/resources (by a data steward) to help provide business users with high-quality data that is easily accessible in a consistent manner	RDA
The formalisation of accountability for the management of data resources	Seiner, 2006
The formalisation of roles and responsibilities and their application to ensure that research objects are managed for long-term reuse, and in accordance with FAIR data principles	EOSCpilot, FAIR4S framework (2018-2019)

3. Method

3.1 General

The responsibilities, tasks and the knowledge, skills, abilities and learning objectives were defined in multiple iterative cycles, following the steps below. Between these steps the definitions were adjusted and reviewed again.

- Analysis of 40 international job descriptions for data specialists mainly in the lifesciences domain (for example data manager, data expert, data consultant), published between 2016 and 2019.
- Analysing and mapping existing data stewardship competency frameworks and the FAIR data principles.
- Consultation of experts via the Consultation Committee (see Appendix 2) and via diverse outreach activities (see Appendix 9).
- Consultation of persons with data steward functions, among others via the Dutch Data Stewards Interest Group and via knowledge exchange workshops (details in Appendix 5).

3.2 Function description

The definition of the function description started from the data stewardship definition as defined in recent publications and reports ([EOSCpilot WP7 D7.1](#), 2017; [EOSCpilot WP7 D7.3](#), 2018; [DAMA-DMBOK2](#), 2014; [EDISON](#), 2017; [ELIXIR](#), 2017; [Teperek et al](#), 2018 and sources displayed in Table 1) and the experience of the project team. Eight competence areas were defined for the function and per area responsibilities and tasks were outlined. This draft function description was compared with an analysis of 40 function profiles and job openings in the field of data stewardship, and was adjusted and complemented based on this information. These 40 function profiles came from the institutes of the project group and from an internet search on job openings. The defined function description was presented twice in the course of the project to the Consultation Committee of the project with national experts on several aspects of data stewardship (see Appendix 2), as well as to [LCRDM](#) and to the Dutch Data Steward Interest Group³. During the data steward workshops in June 2019 (Appendix 5) community feedback was gathered and the definitions were further improved. The final version of the function description was submitted for review and approval to the Consultation Committee of the project (Appendix 3).

3.3 Competencies and learning objectives

Based on the function description (i.e. the tasks and responsibilities) the competencies for the data steward were defined by the project team. For this, the structure of knowledge, skills and abilities (KSA) was used and learning objectives were defined, including the levels of the taxonomy of Bloom⁴. Subsequently, these KSAs were compared with the before mentioned 40 Dutch function profiles and job openings in the field of data stewardship and adjusted and complemented based on this information. An education expert from Elevate-Health⁵ reviewed the KSAs and the learning objectives and the suggested

³ <https://www.dtls.nl/community/interest-groups/data-stewards-interest-group/>

⁴ <https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/>

⁵ <https://elevatehealth.eu/>

adjustments were adopted. For the KSAs and learning objectives the same review procedure was followed a for the function description.

3.4 Training

Two approaches were taken to get information about existing training and about experienced training gaps. During the two workshops in June 2019 feedback was collected from data stewards (details in Appendix 5). Furthermore, the Data Steward Interest Group had already started to make an inventory of available training resources. Both sources of information were combined. Based on the feedback from the community a proposal was made to develop an online tool that would encompass the information in all matrices developed and that would allow an individual to navigate through the knowledge areas, identify both the skills needed and the accompanying training that could be useful to gain those skills. As part of the project, Elevate-Health⁶ made a proposal for developing this tool (Appendix 8).

3.5 Workshops

On June 11 and 12, 2019 two one-day workshops were organised (with identical programmes). The focus of these workshops was to create working sessions with data stewards and other interested stakeholders. There were 22 and 29 participants respectively, with broad representation from various research organisations, universities and HBOs (universities of applied sciences). Participants originated from different domains. The variety in participant roles shows how diverse the current data steward landscape is (e.g. data data stewards, data managers, data librarians, RDM/data officers, project managers, PhD candidates, data consultants, RDM trainers and policy makers). The aim of the workshops was twofold: 1) to test the intermediate deliverables with the data steward community and the main stakeholders, and 2) to get input for improvement and further steps to be taken with regard to training. A detailed description of the workshops can be found in Appendix 5.

3.6 Documentation

During the course of the project, all outputs have been published in a Zenodo community⁷ that was initiated for this project. This way, the (intermediate) results were shared with the community right from the start of the project.

⁶ <https://elevatehealth.eu/>

⁷ <https://zenodo.org/communities/nl-ds-pd-ls/>

4. Results

4.1 Organisational context and shared responsibility

The defined function description for data steward needs to be viewed in the scope of a shared responsibility on data management. Data stewardship is a collective responsibility and endeavour of the institute and the project, in which individual researchers, data specialists, policy makers “have responsibilities to each other as colleagues, to their research project, and to their organisation, as do each of these in return to the individual” (EOSCpilot WP7, 2017 D7.1 page 44-45). Although data stewards perform a key function in data stewardship, data stewards bear only part of the responsibility for data stewardship in the institute or the project. In the end, the director or board of the institute is responsible and fully accountable for data stewardship in the institute, which may be delegated to the head of department or specific scientists in case of a specific research project. Also, apart from the data steward functions, good data stewardship needs to be embedded in the institute, supported by policies, procedures, infrastructure and qualified people (Savage and Cadwallader, 2019).

4.2 Function descriptions

Based on the literature and consultations, three different data steward stakeholder-fields were defined in the data stewardship landscape depending on their interaction with specific stakeholders. All three stakeholder-fields together ensure that research carried out at institutes and in projects produces FAIR data along the data life cycle. Data stewards find themselves positioned at different positions in this landscape as they interface and communicate between these different stakeholder groups.

The three stakeholder-fields are characterised by their own expertise and terminology. Data stewards can communicate between these stakeholder groups and hence provide the interface between fields of expertise. To indicate the particular or main focus of a data steward, three different data steward roles can be defined (Table 2 and Figure 2). A data steward with a focus on policy makers can be called a ‘data steward - policy’, a data steward with a focus on researchers can be called a ‘data steward – research’, and a data steward with a focus on the IT infrastructure can be called a ‘data steward - infrastructure’ (see details below). An individual data steward may have to deal with all stakeholder-fields, but in practice often one of the defined stakeholder-fields stands out.

Table 2. Data steward roles and their focus

Data steward role	Main stakeholder field
Data steward - policy	Policy makers, funders, management, universities board, and the institutes’ deans. This group of stakeholders has a say in, and are ultimately responsible for, how data should be handled
Data steward - research	Researchers and data scientists who produce data and work with the data on a daily basis with a focus on research. Researchers need to be empowered to handle data in a way which is policy compliant without losing the power to execute their research
Data steward - infrastructure	Data and IT infrastructure providers, e.g. IT staff, technicians and application managers. This stakeholder group provides tools to enable implementation of certain data policies and hence make it easier for researchers to manage their data in a policy-compliant way

For all three data steward roles, eight competence areas were defined (Table 3). The responsibilities and tasks in these eight competence areas can be seen as different aspects of the function that the data steward has. In practice, one data steward in a department may have all roles or, in case of multiple data stewards, roles may be divided between different persons. The responsibilities and tasks per competence area are formulated according to the focus of the data steward (see details in Appendix 3). In case a data steward combines the focus on for instance policymakers (data steward – policy) and researchers (data steward – research), this data steward will also have a combination of the responsibilities and tasks defined for these two data stewardship roles. The competence areas were carefully mapped on the existing frameworks and models (Appendix 4).

Table 3. Overview of the eight defined competence areas for all data steward roles

Competence area	This concerns
Policy/strategy	Development, implementation and monitoring of research data management policy and strategy for the research institute
Compliance	Compliance to the Netherlands Code of Conduct for Academic Practice, the Netherlands Code of Conduct for Research Integrity, the General Data Protection Regulation (GDPR), and other relevant legal and ethical standards
Alignment with FAIR data principles	Alignment to the FAIR data principles and the principles of Open Science
Services	Availability of adequate support on research data management, in staff or services
Infrastructure	Availability of adequate data infrastructure for research data management
Knowledge management	Adequate level of knowledge and skills on research data management within the institute, department or project
Network	Obtaining and maintaining a network of aligned expertise areas and relevant departments and organisations inside and outside the institute, department or project
Data archiving	Adequate support and data infrastructure for FAIR and long-term archiving of data of the institute, department or project

4.2.1 Data steward - policy

The ‘data steward - policy’ role has a focus on policies, procedures, regulations and principles relating to data, coming from inside or outside the institute, for instance the institute’s Research Code or EU legislation. ‘Data steward - policy’ translates these general policies, procedures, regulations and principles into guidelines for data and data stewardship and initiates and oversees implementation and monitoring. This role also ensures suitable and sufficient capacity with regard to infrastructure and services. ‘Data steward - policy’ interacts with directors, policy makers and funders, often has a coordinating role and supervises diverse data stewards and related functions. The responsibilities and tasks of this data steward focus around policy development, implementation and alignment (Appendix 3).

4.2.2 Data steward - research

The 'data steward - research' role has a focus on supporting the main data producer, processor and user in academia: the researcher, ranging from undergraduate students, PhD students and postdocs to full professors. The 'data steward - research' is responsible for implementation of the specific data guidelines, translating them to project and domain specific procedures to data management plans. 'Data steward - research' advises researchers and makes sure that within the project data is handled in a manner compliant to the institute's policy. This role works in close collaboration with the researchers, is often connected to a specific scientific project and may perform hands-on work in the project. The responsibilities and tasks of this data steward focus around translating the scientists and project needs with regard to data to infrastructural and service requirements (Appendix 3).

4.2.3 Data steward - infrastructure

The 'data steward - infrastructure' role has a focus on liaising with the people involved in the IT infrastructure, technicians, application managers and other service providers, inside and outside the institute. 'Data steward - infrastructure' easily interfaces with IT personnel and has an overview of the possible IT solutions. This role translate the requirements of policies and science into suitable IT solutions and provides advice. This role facilitates implementation of the IT infrastructure, can give access to data and software and may also perform hands-on work in the project. The responsibilities and tasks of this data steward focus around defining requirements for infrastructure and tools and fit them to IT solutions (Appendix 3).

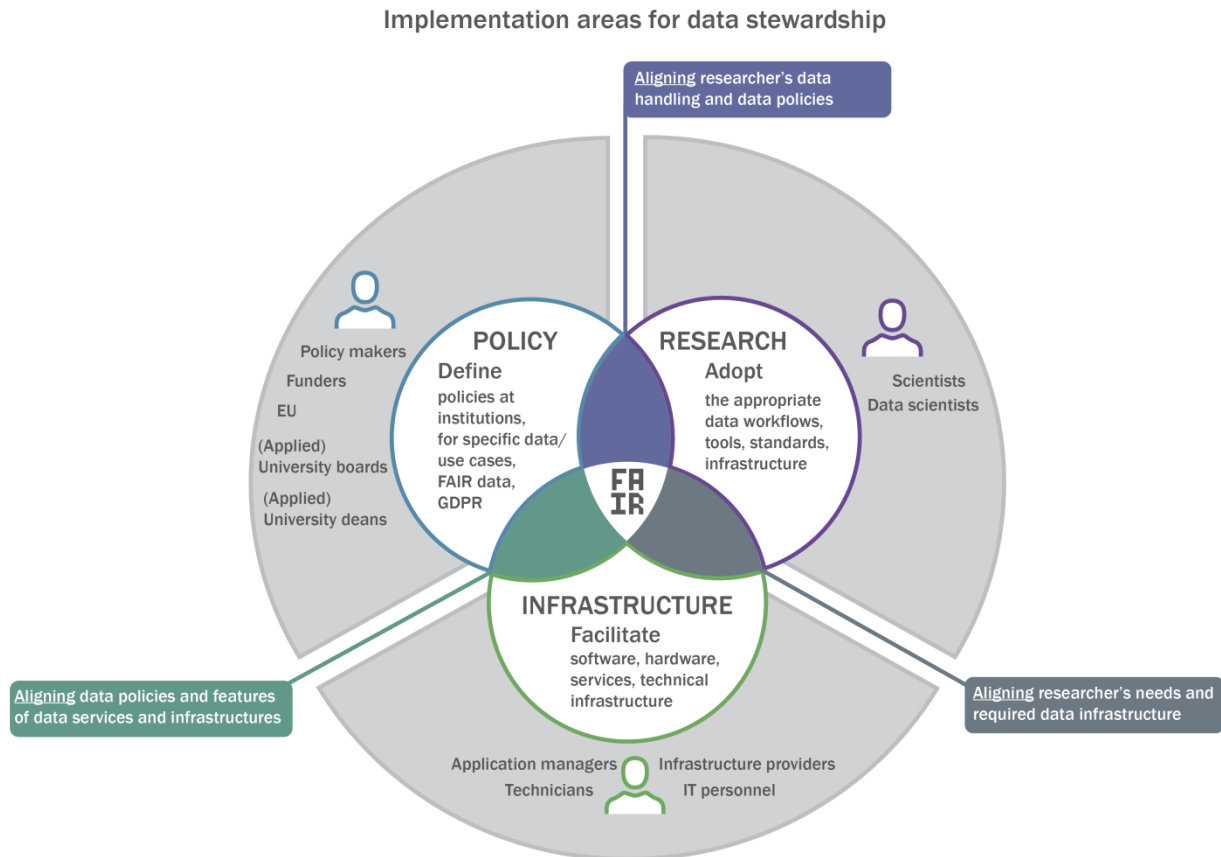


Figure 2. The focus areas of data stewards and mapping of different data steward roles
(<https://doi.org/10.5281/zenodo.3460552>)

Figure 2 shows the three main focus areas and stakeholder-fields of the three different types of data stewards and their interrelatedness with each other and other functions. The three intersecting circles show the three main focus areas of a data steward. The figure does not exclude the possibility that an individual who fills in the role of a data steward can have expertise in several areas of the three focus areas. Ideally, an institute has all three roles covered by internal staff members or in collaboration with other institutes. Other existing data functions, like a data consultant or data manager, can be mapped to the focus areas landscape. Box 1 shows two examples of the practical implementation of data steward roles in institutes.

Box 1. Translation to practice - examples

Implementation of a policy

Assume a 'data steward - policy' concluded that it is necessary to create data publications for each research article publication at the research institute, because of a new code of conduct or because of new guidelines from funders. The data steward with the 'data steward – policy' role defines a new RDM policy with regard to this issue. Subsequently, the data steward with a 'data steward - research' role implements this policy and translates it into the context and practices of the research group or project. The 'data steward – research' defines what should be considered to be a data publication and implements ways of handling of data such that it is known which file belongs to which data publication and research article publication throughout the whole research project. With the knowledge of infrastructures of the 'data steward - infrastructure' role, data publication services are chosen which fulfil the requirements. Furthermore, the 'data steward - infrastructure' supports the 'data steward - research' to build a script to automatically upload a data collection to the data publication service such that several collections can be uploaded in an efficient way.

Implementing metadata policies in infrastructure

Assume a researcher desires to increase the possibilities to combine datasets in a specific discipline. A 'data steward - research' translates this need to a new standard for metadata definition and metadata handling for this discipline. Having the knowledge and overview of technologies, the 'data steward - infrastructure' can define new features and extensions for existing services such that existing metadata is automatically transferred to the new standard on a technical level and that newly created metadata are exposed according to that new standard. Acknowledging the use and benefit of this metadata handling, a 'data steward - policy' may define a policy on metadata in accordance to the institute's FAIR data policy, which requires that data of different research areas need to be interoperable.

4.3 Positioning in the organisation

The positioning of the data steward in the organisation, centralised or decentralised, has been a topic of debate ([Teperek et al, 2018](#); [LCRDM, 2019](#)). Currently, we see a mixture of centralised data steward departments and decentralised, embedded data stewards. The described data steward roles can be implemented at different places in the organisation, so both centrally and locally. Although it is to be expected that the 'data steward - policy' and the 'data steward - infrastructure' are more often positioned at a central department and 'data steward - research' more often in a research project.

4.4 Alignment with other data professions

Besides data stewards, many other functions are involved in aspects of data management and handling. Some functions are already in use for a long time, like a data manager. Here, we describe three functions often found in the research context.

4.4.1 Data manager

The data manager is responsible for managing, curating and coordinating the collection, processing, storage and access of research data generated by researchers ([EDISON, 2017](#); [ELIXIR, 2017](#)). The data manager oversees operationalisation of data quality control and data access. The role is often an integral part of the research team or linked to a specific (research) project, to support the researcher. General activities of the data manager include checking data, preparing documentation about the datasets, handling metadata, providing access to datasets for internal and external users, and preparing datasets for analysis, secondary analysis and archiving. The function may have a functional application management component, with the difference that the data manager has focus on the data, while the functional application manager has focus on the application.

4.4.2 Data scientist

Data science is the extraction of actionable knowledge directly from data through a process of discovery, or hypothesis formulation and hypothesis testing. Data science can be understood as “the activities happening in the processing layer of the system architecture, against data stored in the data layer, in order to extract knowledge from the raw data” (in [EDISON, 2017](#)). This includes activities such as data cleaning, data management, analytics, visualisation and engineering, in the context of among others big data. The data scientist processes and analyses large and complex amounts of data, using advanced statistical modelling and analysis techniques to uncover insights in the data for scientific, organisational or business problems. A data scientist may be a domain-specific scientist or can operate in a support role for a domain-specific research team. They may design and build (together with IT developers) the IT infrastructure needed for these complex analyses, support the researcher in performing these analyses and assist with high-performance computing, cloud computing, parallelisation, data visualisation and privacy-enhancing technologies.

4.4.3 Data consultant

The data consultant identifies and articulates the data and IT needs for research in the institute and translates the requirements to concrete measures and IT solutions. This is done while considering among others the type of data, the specifics of the research, the data analysis tooling, the wishes and needs regarding (long-term) data access, the legal, ethical and privacy context, and the needs and requests regarding data security. The data consultant is connected with a network of experts and liaises with multiple experts, like information architects, database and application developers, privacy by design experts, and security experts. The data consultant guides and assists the development team to create and implement the information architecture and tools. The data consultant provides advice and support for researchers during the complete data life cycle and for the duration of the research, and proactively informs the researcher about (new) possibilities, solutions and demands during the course of the research project.

4.5 Competencies and learning objectives

Data stewardship competence can be seen as the ability to apply the knowledge, skills and abilities needed to meet responsibilities for data stewardship. For the three data steward roles the

responsibilities and tasks were translated into the required knowledge, skills and abilities (KSAs) that are regarded as necessary to fulfil these responsibilities and tasks. Each data steward has competencies on all eight areas, but the content of the KSAs differ between the data steward roles. The KSAs are in turn translated into learning objectives (including Bloom's taxonomy⁸) to be able to fit the KSAs with existing training or design new training. See Appendix 3 for a detailed overview of the KSAs and learning objectives for each of the three data steward roles.

Next to the data steward specific competencies, general professional competencies, like good communication skills, ability to collaborate with a wide variety of disciplines and fluency in English, are required as well. Also, the data steward role is often complemented with more general roles, such as (project) management and coordination. This makes the competencies required for these roles also needed for a data steward.

4.6 Training

4.6.1 Existing data steward training

Data stewards that shared their insights on their training and professional experiences during the workshops in June 2019 (details in Appendix 5), display a variety in backgrounds and seniority. Thus, when questioned what kind of training and knowledge they have incorporated in their career as data steward and in related roles, it became clear that the currently available (inter)national training and advanced training opportunities are frequently referred to. The details of this inventory during the workshops are presented in Appendix 6. The most popular are the 'Essentials 4 data support' by Research Data Netherlands ([RDNL](#) - a joint service from 4TU.Centre for Research Data, Data Archiving and Networked Services (DANS), and SURFsara) and the online training resources from the Consortium of European Social Science Data Archives (CESSDA). Digital material about research data management, Open Science, and data stewardship provided by national (e.g. [LCRDM](#), [NPOS](#)) and international organisations (e.g. EOSCpilot, [DCC](#), [LERU](#)) are used by data stewards and form a strong foundation for their comprehensive daily duties.

One interesting aspect to note is that the majority of the data stewards that attended the workshops indicated that they learned their current skills 'on the job', either by themselves or from a colleague. This makes sense since structured data stewardship training has only been emerging since a few years and many of the people that now are in those roles had to acquire it before that time.

4.6.2 Desired data steward training

In addition, the data stewards reported their desired training during the workshops, according to the eight defined responsibility areas. When it comes to the desired training and educational opportunities to meet the identified gaps in skills and knowledge, the data stewards reported to need professional training. Also, policy making, creating services and training for specific target audiences, data science and informatics skills are considered as decisive competencies to achieve in order to better fulfil the role as data steward. Another aspect that emerged is the differences in skills and training needs depending on the seniority of the data steward. Where a freshly started data steward needs to complete all the groundwork first to get a good body of knowledge for data stewardship specific matters, a more seasoned person seeks more professional training to elevate their current skill sets. In Appendix 6 and 7 we provide an overview of the existing and desired training for data stewards, as well as a curated list of

⁸ <https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/>

online available digital training opportunities for new data stewards and ones that want to deepen certain topics.

4.6.3 Outline for a self-assessment/navigation tool for data stewards

What was clear from the community feedback during the workshops was that, although there is much training available, it is not always properly annotated and findable, let alone that it was clear for which specific skill set a specific training was intended. It was indicated by the data stewardship community that guidance to identify proper training for specific skills would be valued. This led to the proposal to explore developing such a navigation tool combined with a self-assessment tool, linked to the identified KSAs and learning objectives.

The aim of the self-assessment/navigation tool is to convert all information about the competencies, KSAs and learning objectives for the three types of data stewards into an online, easily accessible, interactive tool. The tool will provide a clear visual structure that can be discovered in an interactive way. It has to be suitable for both existing data stewards and for those who are new in the field and want to explore possible career options and needed upskilling. An important part of the tool will be the mapping of specific subareas of the matrices to existing training. Finally, the tool will contain a (self) assessment part, to allow data stewards and others to assess their gaps in knowledge, skills and skills. Elevate-Health⁹ has made a proposal for the development of such a tool, which includes both a description of the foreseen content as well as a budget that would be needed for the implementation (Appendix 8). Development and implementation of this tool will be included in the upcoming project under the umbrella of NPOS.

⁹ <https://elevatehealth.eu/>

5. Conclusion and discussion

This study shows that the data steward in lifesciences can be seen as a role that differs from existing data experts and research roles. Three different, partly overlapping stakeholder-fields of the data steward were characterised, which all have their own focus. Each data steward stakeholder-field has eight competence areas, ranging from policy and compliance to infrastructure, service and network. The content of responsibilities, tasks and KSAs were defined per competence area and differ between the data steward roles. The same goes for the learning objectives that were added to each KSA.

The data steward role is often experienced as a role that is 'in between' different disciplines and professionals. Translation between different data aspects and between different stakeholders and professionals is a key element of the roles. This was also reported by others (EOCS-HLEG, 2016). Therefore, next to the specific data steward competencies, good communication skills and a customer service-oriented attitude are essential for the function.

The knowledge, skills and abilities that current data stewards have and that are required for the function are often developed 'on the job' or via colleagues. Apart from some available, informal national training, no formal trajectory for professionalising data stewardship exists. From the available training, online courses, books and other materials the quality, feasibility and appropriateness is difficult to assess. Additionally, for many data stewards it is unclear what knowledge, skills and abilities they are missing and should develop.

In this project, a practical approach was chosen; it was built on existing knowledge. Two main sources of information were used: reports from high-level international expert groups and projects describing theoretical competency frameworks on the one hand and practical, day-by-day experiences and job descriptions from people involved in the field of data stewardship on the other hand. The project transforms both information sources into a detailed function and competence description as well as a training overview. Already during the project, the deliverables were used for vacancy texts for new data steward positions and for hiring new data stewards.

We explicitly followed this practical approach as well, instead of starting from one specific theoretical or conceptual framework (ESOC, 2017; [EDISON, 2017](#); [DAMA-DMBOK2, 2014](#)). It helped us in creating a coherent data stewardship profile based on the daily experiences and practices of data stewardship, instead of delivering various compositions of competence areas and responsibilities, as the original reports all have different conceptual frameworks. Additionally, many of these reports did not focus on data stewardship as such, but on data management ([DAMA-DMBOK2, 2014](#)), Open Science ([EOSC, 2017](#)) and data science ([EDISON, 2017](#)). In our project, for instance, the competence area 'data archiving' was added on explicit request from data stewards, because they regarded this an important part of their daily job that also required specific competencies that were not covered by the other areas. In the original reports, data archiving often is regarded to be part of 'data storage and operations' ([DAMA-DMBOK2, 2014](#)).

Data stewardship is a new function in lifesciences and in general. Therefore, the function description and competence definition should be seen as first versions and require continuous evaluation. In particular the outline for future training is a starting point and gives direction for concrete actions, but needs further development and implementation. During this implementation phase new insights may arise. So, we advise to continue to have various iterative cycles to further improve the function

description and training. The current description lacks a differentiation in prior educational level (bachelor/master/PhD) and does not include different function levels (junior/senior). This requirement needs to be assessed in future projects.

Although the function description was developed for the lifesciences domain, most parts of the function and training are not specific for this domain but are applicable to other domains as well. The function description corresponds with a recent publications by the [LCRDM](#) on data steward tasks and job profiles and institutional positioning of the function in the academic landscape ([LCRDM, 2019](#)). For domains or projects not dealing with human subject data, parts of the areas, responsibilities and tasks may be left out, such the competence area ‘compliance’ (knowledge on GDPR). In practice, the data steward role often may be complemented with domain-specific knowledge ([Teperek et al., 2018](#); [EOSC, 2017](#)). In addition, responsibilities, tasks and competencies may be added on Open Science, because of the strong link to this field ([EOSC, 2017](#)).

The academic community is highly international and many initiatives on Open Science and data management are coordinated on the European level. This made national and international alignment essential to this project. The project was aligned with the work for the ELIXIR-NL implementation study ‘Towards Data Stewardship in ELIXIR - Training & Portal’ and parts of the project were already presented on an international stage on several occasions, such as the Liber conference (library community), the Open Science community (Open Science Fair Porto), the ELIXIR All Hands meeting and the ISMB conference (both lifesciences domain). The next stage will be the worldwide Research Data Alliance (RDA) plenary. On all occasions, our deliverables, conclusions and recommendations were very well received. This shows that the function description and competencies are being recognised by the community and indicate that they can be used in the international research arena, regardless of the discipline.

This project defined, for the first time, the data steward function for the lifesciences domain on the level of detail necessary to implement the function in research institutes and to appoint data stewards. The overview of existing training and the outline for new training provide a strong basis for future formal data steward education, to further professionalise the data steward function and contribute to the increase in capacity and quality of data stewards in the institutes and the lifesciences domain.

6. Recommendations and directions for further development

In line with the conclusions of this report, we recommend the Dutch lifesciences community to embed the data steward roles and competencies in the formal university function profile structure. This may not have to result in a separate function profile but could also probably be implemented by adding roles to existing function profiles. Our detailed recommendations are:

Competencies and KSAs:

- Assess the possibility to define different function levels (junior/senior or HBO/university level) and add formal job evaluation and grading.
- Complement the data steward competencies with discipline specific knowledge and skills.
- Organisations that have data stewardship support in place, can use the matrices to assess if their team contains all roles needed for the institute or project.

Training and self-assessment/navigation tool:

- Develop a self-assessment tool for defining the data steward role and the KSAs, so current data stewards can assess their responsibilities, tasks and competencies.
- Combine this self-assessment tool with navigating directions toward training, courses, and learning materials based on the assessment, needs and preferences of the data steward.
- This assessment tool could also be used on an organisational level to assess whether the organisation is up to standards in relation to the functions and activities the organisations has, especially when an organisation distributes the data competencies over a number of different people.
- In time, develop a certified training for data stewards.

Alignment with national and international initiatives:

- For further development and implementation of the function and training, align with the Open Science movement on a national and international level through the National Platform for Open Science ([NPOS](#)) and [EOSC](#).
- Remain in close collaboration with the national stakeholders, the national data management community (e.g. [LCRDM](#), [DTL](#) Data Steward interest group, the Health-RI community, and the NFU Data management group), and the international community ([ELIXIR](#), [EOSC](#)).
- Explore further exchange and alignment with [GO FAIR](#) and GO TRAIN, with the [ELIXIR training programme](#), and with data education in the institutes.

Although the project set out to define the data steward description for the lifesciences, the current work has also relevance for other domains, as the results of this project already have been picked up and commended outside the lifesciences.

At the moment of writing of this report, information about continuation of the work is already available. Starting sept 2019, part of the core team will continue to work on these recommendations in a small 1-year [NPOS](#) funded project on data stewardship.

References

- DAMA-DMBOK2 Framework. 2014. Data Management Body of Knowledge. The Data management association (DAMA) International. <https://dama.org/content/body-knowledge>
- EDISON. 2017. Education for Data Intensive Science to Open new science frontiers. EDISON Data Science Framework. Part 1. Data Science Competence Framework (CF-DS) Release 2. https://github.com/EDISONcommunity/EDSF/blob/master/data-science-professional-profiles/EDISON_DSPP-release3-v07.pdf
- ELIXIR Implementation Study: Towards Data Stewardship in ELIXIR: Training & Portal (ongoing). <https://elixir-europe.org/about-us/implementation-studies/towards-data-stewardship>
- ELIXIR position paper on FAIR Data Management in the Life Sciences. 2017. <https://f1000research.com/documents/6-1857>
- EOCS. European Open Science Cloud Declaration. 2017. https://ec.europa.eu/research/openscience/pdf/eosc_declaration.pdf
- EOCS, FAIR4S. 2019. EOCSpilot framework of FAIR data stewardship skills for science and scholarship, and draft recommendations on FAIR training. https://www.eoscipilot.eu/sites/default/files/fair4s_eoscipilot_skills_framework.pdf
- EOCSpilot (<https://eoscipilot.eu/node>) WP7 deliverable. D7.1: Skills landscape analysis and competence model. 2017. <https://eoscipilot.eu/sites/default/files/eoscipilot-d7.1.pdf>
- EOCSpilot (<https://eoscipilot.eu/node>) WP7 deliverable. D7.2: Interim report and catalogue of EOCS skills training and educational materials. 2017. https://eoscipilot.eu/sites/default/files/eoscipilot-d7.2_0.pdf
- EOCSpilot (<https://eoscipilot.eu/node>) WP7 deliverable. D7.3: Skills and Capability Framework. 2018. <https://eoscipilot.eu/sites/default/files/eoscipilot-d7.3.pdf>
- EOCSpilot (<https://eoscipilot.eu/node>) WP7 deliverable. D7.4: Report on Training Workshops. 2018. https://eoscipilot.eu/sites/default/files/eoscipilot_d7.4_submitted.pdf
- European Commission. Providing researchers with the skills and competencies they need to practise Open Science. Report of the Working Group on Education and Skills under Open Science. 2017. https://cdn1.euraxess.org/sites/default/files/policy_library/ec-rtd_os_skills_report_final_complete_2207_1.pdf
- HLEG. European Commission High Level Expert Group on the European Open Science Cloud. 2016. Realising the European Open Science Cloud. https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf
- LCRDM. 2019. Data Stewardship on the map: A study of tasks and roles in Dutch research institutes. <https://zenodo.org/record/2669150#.XS3GkugzaUk>
- LERU. League Of European Research Universities. Advice Paper no.24. 2018. Open Science and its role in universities: A roadmap for cultural change. <https://www.leru.org/files/LERU-AP24-Open-Science-full-paper.pdf>
- KNAW. 2018. big data in wetenschappelijk onderzoek met gegevens over personen. www.knaw.nl/nl/actueel/publicaties/big-data-in-wetenschappelijk-onderzoek-met-gegevens-over-persoonen
- Mons B. 2018. Data Stewardship for Open Science: Implementing FAIR Principles. <https://www.crcpress.com/Data-Stewardship-for-Open-Science-Implementing-FAIR-Principles/Mons/p/book/9780815348184>

- NFU HANDS handbook. Data4LifeSciences. 2018. <http://data4lifesciences.nl/hands/handbook-for-adequate-natural-data-stewardship/what-is-data-stewardship/>
- Purdue University. 2016. Pilot Data Information Literacy Competencies Matrix Scaffolded Across Undergraduate, Graduate and Data Steward Levels. Libraries faculty and staff scholarship and research. https://docs.lib.purdue.edu/lib_fsdocs/136/
- RDA. Research Data Alliance. https://smw-rda.esc.rzg.mpg.de/index.php?title=Data_Stewardship
- Savage JL and Cadwallader L. 2019. Establishing, Developing, and Sustaining a Community of Data Champions. *Data Science Journal*, 18(23): 1-8
- Seiner R. 2006. The Data Stewardship Approach to Data Governance: Chapter 1. The Data Administration Newsletter (TDAN.com). <http://tdan.com/the-data-stewardship-approach-to-data-governance-chapter-1/5037>
- Teperek M et al. 2018. Data Stewardship: Addressing Disciplinary Data Management Needs. *International Journal of Digital Curation*, 13(1): 141-149
- Wilkinson et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* volume3, Article number: 160018 <http://www.nature.com/articles/sdata201618>

Acknowledgements

The authors acknowledge Paula Andrea Martinez from ELIXIR-Belgium who provided a script to automatically analyse job vacancies and derive the described competencies, knowledge and skills. This script was used in addition to the manual analysis described in the Methods section.

The authors acknowledge Nienke Verdonk from Elevate-Health¹⁰ for her work related to the analysis and formulation of the competencies, KSAs and learning objectives.

Funding information

This project is funded by the ZonMw Personalised Medicine Programme (Dossier number: 80-84600-98-3007), Zilveren Kruis en KWF Kankerbestrijding. Additional in-kind funding was provided by UMCG, UMCU, Radboud University Nijmegen, Radboudumc and DTL/ELIXIR-Netherlands. The project period was Aug 2018 - July 2019.

¹⁰ <https://elevatehealth.eu/>

Appendix 1 – Deliverables

The initial deliverables formulated in the project proposal are stated below. During the course of the project the content of deliverable 3, 4 and 5 changed. These changes were endorsed by the consultation committee. The content of deliverable 4 was worked out slightly different, because during the project it became clear that we first needed to have input from data stewards and where they got their training and subsequently develop new training together. Instead of a blueprint for a curriculum and an eLearning module it became clear that a self-assessment tool that was connected to a navigation to existing teaching and training was more useful for the community at this moment. Deliverable 6, a sustainability plan, took shape in 1-year [NPOS](#) funded project on data stewardship.

- Deliverable 1: a function description for a data steward specific for the lifesciences domain. The description includes the responsibilities and activities of a data steward and entails minimally two different function levels. This function is supported by NFU partners, [DTL](#), [LCRDM](#), and [Health-RI](#) and is built based on existing efforts.
- Deliverable 2: a definition of the knowledge base, skills and competencies for lifesciences data stewards based on the function description.
- Deliverable 3: a report on a blueprint for an education line (course trajectory) that describes the required education and training for data stewards in lifesciences, with reference to existing courses. Accompanying training material for a maximum of two courses-components are developed. All training materials are made modular, and are set up in such a way that they can be optimally reused and implemented by education institutes after the project ends.
- Deliverable 4: using the developed training materials, one two-day hands-on courses for 20 data stewards is organised, aiming for a national and discipline wide coverage. This course functions as a pilot to test the training and the training materials.
- Deliverable 5: a design and script for an eLearning module is developed as a starting point for full implementation in the following phase, after more funding has been acquired.
- Deliverable 6: A sustainability plan to ensure national embedding of the expertise, to ensure further development of the data steward function and continuation of the training.

Appendix 2 - Project team and consultation committee

Core team

The project was coordinated and executed by a small operational team consisting of a project coordinator and four project members with affiliations from different organisations. The core team was assisted by a master student who did an internship and wrote a master thesis aligned with the project. Expertise on adult education and didactics, in particular in eLearning development, was included in the project for the design of the education line and education materials (Elevate-Health¹¹). The project was supported by the office of DTL (administration and finances).

The operational core team consisted of:

Role	Name	Affiliation(s)
project coordinator / project member	Salome Scholtens	Genomic Coordination Center (GCC), UMCG; Data Federation Hub (DFH) liaison (RUG)
project member	Celia van Gelder/Christine Staiger	DTL, ELIXIR-NL/ELIXIR, EOSC-Life, GO TRAIN
project member	Jasmin Bohmer	UMCU, Chair DTL Data Stewards Interest Group
	Nelly Anbeek	UMCU
project member	Inge Slouwerhof/Mijke Jetten Mirjam Brullemans-Spansier	Radboud University Nijmegen Radboudumc
project member (intern)	Marije van der Geest	Genomic Coordination Center (GCC), UMCG

Consultation committee

Community building is an essential part of the project, to ensure that the deliverables of the project match the needs of the target group and are also taken up by the target group, thus ensuring sustainability. We accordingly set up a broad consultation group, advising the core team and helping to assure the proper alignment with other local, national and international projects, initiatives and stakeholders. The consultation committee consisted of the major stakeholders in the Dutch national field of data stewardship.

¹¹ <https://elevatehealth.eu/>

The consultation committee consisted of:

Domain	Name	Affiliation(s)
Funders	Margreet Bloemers	ZonMw
Lifesciences research infrastructures	Jan-Willem Boiten	Data4lifesciences, Health-RI, Lygature
UMCU	Paula Jansen	Data4lifesciences, data management workpackage
Policy makers	Petra Drankier	NFU policy
Data steward tooling and IT solutions	Morris Swertz	UMCG, BBMRI, ELIXIR-NL
Universities of applied sciences (HBO)	Bo Blanckenburg	Hogeschool Leiden, Domain Applied Science (DAS, HBO)
e-infrastructures	Hylke Koers	SURF, SURFsara
Data Stewards expertise and network	Marta Teperek	TUD, DTL Data Stewards Interest Group
Data Stewards expertise and network	Ingeborg Verheul	LCRDM

Alignment with (inter)national initiatives

Alignment with SURF, [LCRDM](#), ZonMw, [Data4lifesciences](#), [DTL](#), [Health-RI](#), [EOSC](#), EDISON, [ELIXIR](#) was ensured via the core team and/or the consultation committee. Also, there was a strong liaison with the national Data Stewards Interest Group of [DTL](#). Members from this interest group, from the [DTL](#) network, and data stewards from the UMCs were consulted for the design of the function description and definition of the knowledge base. All these stakeholders were actively involved in the design of the education material and in the workshops.

Appendix 3 - Function and competencies matrices

This Appendix provides the detailed description of the function (responsibilities and tasks), competencies (knowledge, skills and abilities) and learning objectives for all three data steward roles, structured by the eight competence areas.

[This relates to Deliverable 1 and 2 of the project proposal]

Data steward – policy			
<i>Keywords: Institutional, coordinating, policy</i>			
<i>Focus on policy makers, managers, board of directors, financial and legal experts</i>			
Responsibilities*	Activities and tasks	Knowledge, Skills and Abilities	Learning Objectives <i>(after successful completing training you will be able to)</i>
<p>1) 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.</p>	<ul style="list-style-type: none"> - 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. 	<ul style="list-style-type: none"> - Knowledge about the most important elements of a successful RDM policy. - Knowledge about internal policies and financial and legal constraints within the institute, relevant for RDM policy. - Knowledge about relevant internal and external (funder, publishers, government) stakeholders. - Knowledge about the general research process and all aspects of RDM and the research data life cycle. - Develop, implement, monitor and evaluate policies regarding research data that are endorsed by researchers and aligned with internal and external stakeholders and effectuate change management. - Identify RDM related issues in the research data life cycle at the level of the institute and engage the right persons to solve the issues. - Act as a spokesperson for the institute on research data. - Give advice on RDM and formats for RDM plans to a broad audience. - Communicate about the RDM policy, explain implications and create awareness. 	<ul style="list-style-type: none"> - Explain the most important elements of a successful RDM policy (understanding). - Outline and compare existing internal policies, legal constraints and financial and operational implications with regard to the RDM policy (understanding & applying). - List relevant internal and external stakeholders (remembering). - Explain important aspects of the general research process and all aspects of RDM and the research data life cycle (understanding). - Develop and implement RDM policies within the institutes and align with other internal policies, legal constraints and financial and operational implications (creating). - Monitor the implementation of RDM policies among researchers and other stakeholders and effectuate change management (evaluating). - Identify RDM related issues relevant for the institute (creating). - Support the evolution of DMP templates and monitor that captured features are synchronised with aspects of the policies (creating). - Act as a spokesperson for the institute on research data (applying). - Establish an apprehension of the researchers towards the need for and benefit from RDM (creating). - Communicate about the RDM policy and its implications for the institute and for the individual researcher (applying).

2) 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 institutes practice.
- Develops and/or guides standard solutions for recurring data issues and for data classification, including input for the data protection impact assessment (DPIA).

3) Alignment with FAIR data principles.

Responsible for alignment of the RDM policy to the FAIR data principles and the principles of Open Science.

- Ensures compatibility of the RDM policy and monitors alignment 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

- Knowledge about legislation, ethics and code of conducts with regard to research data, (medical) research and privacy (e.g. WGBO, WMO, GDPR, data licences).
- Knowledge about RDM related issues in the field of legislation and ethics in the research data life cycle.
- Translate RDM policy and legislation and codes of conducts with regard to research data to practical implications and guidelines for that researchers can understand.
- Translate legal requirements on data sharing and archiving to practical implications and guidelines that researchers can understand.

- Knowledge about the FAIR data principles and the principles of Open Science, including knowledge about implementation and how to be compliant to the FAIR data principles.
- Knowledge about findability (F) of data, including adequate data-infrastructure and tools, persistent identifiers and rich (institute-specific) metadata standards.
- Knowledge about accessibility (A) of (meta)data to potential (re)users.
- Knowledge about interoperability (I) of data, including broadly applicable languages, vocabularies and other standards.
- Knowledge about reusability (R) of data, including documentation and licenses with the conditions for reuse and IP rights.
- Communicate about the FAIR data principles and create awareness within the institute, to researchers, support staff and relevant

- List relevant legislation, ethical principles, and codes of conduct for RDM (remembering).
- List RDM related issues in the field of legislation and ethics for the research data life cycle (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).
- Explain and evaluate the cause and possible effects of RDM related issues in the field of legislation and ethics in the data life cycle issues (understanding & evaluating).
- Communicate to researchers, students, support staff and relevant stakeholders about the FAIR data principles (applying).
- Create awareness on the FAIR data principles within the institute, tailored to the needs of the researchers, students, support staff and relevant stakeholders (creating).
- Assess and monitor alignment to the FAIR data principles within the institute (evaluating).
- Contribute to the compliance definition for the institute in relation to FAIR data (creating).
- List the RDM issues related to the FAIR data principles in the research data life cycle (remembering).
- Develop a plan for solving the RDM issues related to the FAIR data principles (creating).
- Explain what compliance to the FAIR data principles entails (understanding).

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.

- Advises management, researchers and support staff of the institute on RDM support.

- Initiates or supervises the set-up and update of suitable support facilities or services in the institute.

- Monitors the development and implementation of (standard) solutions for recurring RDM issues and questions.

- Monitors the availability of sufficient support staff and initiates hiring or training of staff.

- Communicates and stimulates the use of available services by researchers.

- Initiates and supervises

stakeholders.

- Knowledge about the required level of knowledge for researchers and relevant stakeholders with regard to FAIR data principles and the implications for research.

- Knowledge about the institute's KPIs for FAIR data management.

- Assess and monitor compliance to the FAIR data principles and the principles of Open Science of the institute.

- Knowledge about facilities and tools to support FAIR data management.

- Solve RDM related issues related to the FAIR data principles in the research data life cycle.

- Knowledge about RDM related issues in the research data life cycle with regard to available support in staff and services at the organization.

- Knowledge about relevant support available for RDM, both internally as well as externally.

- Knowledge about global development of RDM.

- Assess and analyse needs from researchers regarding support on RDM for researchers and define new requirements.

- Formulate requirements for RDM support (staff and services).

- Communicate with technical and support staff about RDM related services.

- Communicate and stimulate the use of available services by researchers.

- Knowledge about relevant facilities, data-

- Explain what kind of support and assistance is needed by researchers and stakeholders with regard to RDM (understanding).

- Assess and implement adequate RDM support and assistance within the institute (applying).

- List existing support (staff and services) on RDM for researchers (remembering).

- List institutional RDM requirements with regard to staff and services (remembering).

- Analyse the current status within the institute with regard to the institutional RDM support requirements (analysing).

- Analyse needs from researchers and stakeholders (analysing).

- Define new requirements based on needs from researchers and stakeholders (creating).

- Communicate with technical and support staff about RDM related services (applying).

- Formulate requirements for institutional RDM support staff (creating).

- List existing data-infrastructure, tools and emerging

4) Services.

Responsible for the availability of adequate support on RDM, in staff or services, for the researchers and research support staff of the institute.

5) Infrastructure.

<p>Responsible for identifying the requirements of adequate data-infrastructure for RDM to comply with the institute's RDM policy and alignment to (inter)national data-infrastructure and tools.</p>	<p>requests and acquires data-infrastructure and tools for RDM within the institute.</p> <ul style="list-style-type: none"> - Supervises monitoring of the need, use and availability of data-infrastructure and tools. - Advises the management of the institute on data-infrastructure and tools. - Is aware of and advises the institute on relevant (inter)national data-infrastructure. 	<p>infrastructure, tools and emerging standards relevant to RDM (e.g. data capture tools, data cataloguing software, HPC environment, statistical programmes).</p> <ul style="list-style-type: none"> - Knowledge about RDM issues related to data-infrastructure and tools in the data life cycle. - Assess and analyse needs regarding data-infrastructure and tools for RDM for researchers and relevant stakeholders. - Translate these needs to requirements and advise for the management of the institute. - Continuously monitor data-infrastructure and tools available inside and outside the institute, in close collaboration with the responsible (IT) department. - Explore new trends in data-infrastructure and tools and translate these to advise for the management of the institute. - Assures proper alignment of the data-infrastructure of the institute to the relevant external data-infrastructure. - Knowledge about relevant RDM topics for researchers and relevant stakeholders. - Knowledge about the required level of knowledge for researchers and relevant stakeholders. - Knowledge about available RDM training opportunities and possibilities. - Knowledge about the institute's staff abilities for developing RDM training in the institute. - Knowledge about policy, legal requirements and infrastructure on a (inter)national level. - Coordinate organization and/or development of tailored training to develop RDM skills in the institute. - Assess RDM knowledge and skills and identify 	<p>standards for RDM (remembering).</p> <ul style="list-style-type: none"> - List relevant data-infrastructure and tools for the research that is being performed within the institute (remembering). - Monitor the data-infrastructure and tools that are in use in the institute (evaluating). - Analyse the needs of researchers in the institute with respect to data-infrastructure and tools (analysing). - Assess needs of researchers in the institute with respect to data-infrastructure and tools (evaluating). - Explore channels for finding new trends and needs in RDM data-infrastructure and tools (applying). - Signal new trends in data-infrastructure and tools that could be useful for the institute's RDM policy to the management (analysing). - Formulate data-infrastructure and tooling advise for the management of the institute in collaboration with the relevant institutional stakeholders (creating). - List the required RDM knowledge and skills among researchers and relevant stakeholders (remembering). - Assess RDM knowledge and skills of staff (evaluating). - Select relevant RDM topics for training corresponding to the listed requirements (applying). - Coordinate organization and/or development of tailored training to develop relevant RDM skills in the institute (applying).
<p>6) Knowledge management. Responsible for determining the adequate level of knowledge and skills on RDM within the institute in order to comply with the institute's RDM policy.</p>	<ul style="list-style-type: none"> - Monitors RDM skills of the researchers and research support staff within the institute. - Identifies knowledge and skill gaps and ensures appropriate training. - Ensures a sufficient level of awareness among researchers and research support staff of the institute. 	<ul style="list-style-type: none"> - Knowledge about relevant RDM topics for researchers and relevant stakeholders. - Knowledge about the required level of knowledge for researchers and relevant stakeholders. - Knowledge about available RDM training opportunities and possibilities. - Knowledge about the institute's staff abilities for developing RDM training in the institute. - Knowledge about policy, legal requirements and infrastructure on a (inter)national level. - Coordinate organization and/or development of tailored training to develop RDM skills in the institute. - Assess RDM knowledge and skills and identify 	<ul style="list-style-type: none"> - List the required RDM knowledge and skills among researchers and relevant stakeholders (remembering). - Assess RDM knowledge and skills of staff (evaluating). - Select relevant RDM topics for training corresponding to the listed requirements (applying). - Coordinate organization and/or development of tailored training to develop relevant RDM skills in the institute (applying).

gaps among researchers and relevant stakeholders.

7) Network.

Responsible for obtaining and maintaining a network of aligned expertise areas and relevant departments and organisations inside and outside the institute with regard to RDM.

- Refers researchers to other RDM related facilities and services (legal, financial and/or operational), inside and outside the institute.
- Liaises with experts inside and outside the institute.
- Maintains a network with RDM related colleagues and other relevant departments and organisations.
- Connects data support people with each other.

- Knowledge about where to find institute decision makers, stakeholders and researchers, including relevant networks.
- Liaise with institute decision makers, stakeholders, researchers, other data stewards and establish an active network.
- Connect data support people within and outside the institute.
- Communicate with a diverse range of stakeholders.

- Explain where to find institute decision makers, stakeholders and researchers (understanding).
- Liaise with institute decision makers, stakeholders and researchers (applying).
- Establish an active network in which regular consultations are taking place (creating).

8) Data archiving.

Responsible for identifying the requirements of adequate support and data infrastructure for FAIR and long-term archiving of data of the institute, stored internally as well as externally, and for sustainable and legitimate

- Develops, implements and monitors the institute's internal and/or external data archiving and access policy.
- Monitors the internal and/or external archiving of data by researchers of the institute.
- Assesses whether internal and/or external data storage

- Knowledge on archiving workflows and on types of archives and repositories with regard to data quality assurance and trustworthiness
- Develop, implement and monitor the institute's internal and/or external data archiving and access policy.
- Communicate to researchers and relevant stakeholders about data archiving policies and the implication for the research workflow, data-infrastructure and tools.

- List available data archiving solutions and procedures for researchers and support staff (remembering).
- List the requirements regarding internal/external data storage and archiving services (remembering).
- Explain the essential factors/elements for data archiving and access (understanding).
- Create, implement and monitor a data archiving and access policy (creating).
- Audit the internal and/or external archiving of

<p>access to data sources of the institute, for the required period.</p>	<p>and archiving facilities meet the applicable requirements.</p>	<ul style="list-style-type: none"> - Interact with internal and external stakeholders on data archiving with respect to requirements by external and internal stakeholders like funder requirements, national legislation and university policy. - Generic knowledge about metadata schemes and vocabularies (e.g. Dublin Core, DataCite). - Have general oversight on existing differences between accepted file formats for archiving and domain specific file formats. - Audit archiving procedures and workflows in the institute (ways of executing archiving duties). - Provide general advice to the institute management on choosing the archive that is in line with the institute's RDM policy. 	<p>research data (evaluating).</p> <ul style="list-style-type: none"> - Communicate with users about data archiving solutions and procedures for researchers and support staff (applying).
--	---	--	---

* *Researcher* includes senior researchers, postdocs, PhD students and students
Data-infrastructure includes all technical infrastructure aspects necessary for data management and data handling for research
Research includes all disciplines and methodologies, and qualitative as well as quantitative research
RDM = Research Data Management
DMP = Data Management Plan
ELSI = Ethical, Legal, Social Implications
FAIR = Findable, Accessible, Interoperable, Re-useable

Data steward – Research			
<i>Keywords: Research, operational, supporting Focus on researchers</i>			
Responsibilities	Activities and tasks	Knowledge, Skills and Abilities	Learning Objectives <i>(after successful completing training you will be able to)</i>
<p>1) Policy/strategy. Responsible for the development and implementation of a DMP for departments, projects or data collections within the institute that is in alignment with the research requirements, specifications and practices, and is in line with the institute's RDM policy and supports FAIR data and Open Science.</p>	<ul style="list-style-type: none"> - Develops DMP templates tailored for the departments, projects or data collections within the institute. - Writes and/or supports researchers in writing a DMP for departments, projects and data collections, in line with the institute's RDM policy. - Implements RDM as a regular aspect of doing research. 	<ul style="list-style-type: none"> - Knowledge about the general research process and all aspects of RDM and the research data life cycle. - Knowledge about the content of a DMPs, including knowledge about the purpose and how to use it within the research process. - Develop tailored DMP templates together with stakeholders that are understood and can be used by researchers and support staff within the institute. - Translate RDM policies and DMPs to practical implications and guidelines that researchers can understand. - Monitor the use of DMPs by researchers. - Design and draft DMPs for research projects. - Review DMPs written by researchers and help adjusting and refining DMPs over time. - Identify services which support the researcher in putting the DMP into action. - Monitor research projects with regard to data management. - Knowledge about relevant stakeholders and how to contact them. - Act as a spokesperson for the project with regard to research data. - Knowledge about giving advice and providing feedback on RDM to researchers. - Communicate about the RDM policy, explain implications and create awareness. 	<ul style="list-style-type: none"> - List the most important tasks of a spokesperson (remembering). - Act as a spokesperson for the project on research data (applying). - Explain important aspects of the general research process and all aspects of RDM and the data life cycle (understanding). - Outline the essential elements for a understandable and usable DMP template (understanding). - Define the needs of researchers and support staff for understanding and using a DMP template (remembering). - Develop tailored DMP templates in line with the RDM policy, together with stakeholders (creating). - Translate RDM policies and DMPs into concrete actions and tasks (applying). - Explain how to monitor the use of DMPs (understanding). - Apply data services to data management problems identified in DMPs (applying). - List the most important aspects of a DMP on which feedback should be provided (remembering). - Explain how a DMP should look like and give advice to researchers about how to write a DMP (applying). - Write a DMP (applying). - Provide advice and feedback on DMPs for researchers (evaluating). - Monitor the use of DMPs by researchers within the institute (applying).

2) Compliance.

Responsible for advice on compliance of the project or data collection with the DMP and the institute's RDM policy, relevant codes of conduct, legislation and field specific standards.

- Monitors and supervises the execution of a project or data collection in line with the DMP and relevant codes of conduct and legislation, including ELSI aspects.
- Identifies gaps and takes action if needed.
- Train researchers on compliance requirements.

- Knowledge about legislation, ethics, code of conducts and research codes with regard to (research)data, (medical) research and privacy (e.g. WGBO, WMO, GDPR, data licences, DPIA).
- Knowledge about RDM related issues in the field of legislation and ethics in the research data life cycle that are relevant for the project.
- Knowledge of field specific standards.
- Support researchers in establishing compliance with the ELSI procedures of the institute.
- Monitor compliance of the DMPs of the projects to the ELSI procedures of the institute and takes action in case of gaps.

3) Alignment with FAIR data principles.

Responsible for alignment of the DMP to the FAIR data principles and the principles of Open Science, and for facilitating and supporting FAIR data.

- Advises, supports and provides guidelines to researchers on the findability (F) of data, including adequate data-infrastructure and tools, persistent identifiers and rich (institute-specific) metadata standards.
- Advises, supports and provides guidelines to researchers on the accessibility (A) of (meta)data to potential reusers.
- Advises, supports and

- Knowledge about findability (F) of data, including adequate data-infrastructure and tools, persistent identifiers and rich (institute-specific) metadata standards.
- Knowledge about accessibility (A) of (meta)data to potential (re)users.
- Knowledge about interoperability (I) of data, including broadly applicable languages, vocabularies and other standards.
- Knowledge about reusability (R) of data, including documentation and licenses with the conditions for reuse and IP rights.
- Knowledge about the FAIR data principles and the principles of Open Science, including knowledge about implementation and how to be compliant to the FAIR principles.

- Explain the RDM policy and its implications (understanding).
- Communicate about the RDM policy (applying).
- Establish an apprehension of the researchers towards the need for and benefit from RDM (creating).
- Explain RDM policies with regard to research data to researchers (applying).
- Examine and list the practical implications of RDM policies with regard to research data (analysing).
- Create guidelines based on RDM policies with regard to research data (creating).
- Explain ELSI aspects and legal licenses relevant for the project data (remembering).
- Guide the selection of data licenses (applying).
- List the RDM related issues in the field of legislation and ethics in the research data life cycle (remembering).
- Give a short explanation for each issue (understanding).
- Explain the FAIR data principles (understanding).
- Communicate about the FAIR data principles (applying).
- Create awareness on FAIR data among the researchers (creating).
- Explain when compliance to the FAIR data principles is reached (understanding).
- Assess and monitor alignment of research DMPs to the FAIR data principles (applying).
- List the RDM issues related to the FAIR data principles for a research project (remembering).
- Develop a plan for solving the RDM issues related to the FAIR data principles for a research project (applying).
- List available services and tools for transformation of research data into FAIR data (remembering).

provides guidelines to researchers on the interoperability (I) of data, including broadly applicable languages, vocabularies and other standards.

- Advises, supports and provides guidelines to researchers on the reusability (R) of data, including documentation and licenses with the conditions for reuse and IP rights.

- Advises and supports researchers on RDM.
- Advises researchers on services and all aspects of the research data life cycle in the department, project or data collection.
- Develops and implements (standard) procedures for recurring issues and questions.

- Provides guidance and instruction on discovery, acquisition and (re)use of data (in the public domain).
- Solves practical RDM problems.

- Monitors the needs regarding data-infrastructure and tools for RDM within the department, project or data collection.

- Communicate about the FAIR data principles and create awareness among researchers.
- Knowledge about the required level of knowledge for researchers with regard to FAIR data principles and the implications for research.
- Assess and monitor compliance of projects to the FAIR data principles.
- Solve RDM related issues related to the FAIR data principles in research projects.
- Capability to transform discipline specific research data into FAIR data with help of available services and tools.

- Knowledge about RDM related issues in the research data life cycle with regard to available support in staff and services.
- Advise and assist researchers on short and long term actions for RDM.
- Knowledge about relevant support available for research data (management), both internally as well as externally.
- Assess and analyse needs from researchers regarding support on research data
- Knowledge about which RDM related services are managed by the technical staff
- Communicate with researchers, technical staff and support staff about RDM related services.
- Knowledge and good understanding of researchers needs, ideas etc.

- Knowledge about relevant facilities, tools and emerging standards available for RDM, both internally as well as externally (e.g. data capture tools, data cataloguing software, HPC environment, statistical programmes).

- Use the available services and tools to successfully transform research data into FAIR data (applying).

- Define the short and long term actions for RDM (remembering).
- Explain what kind of advice and assistance is usually needed by researchers (understanding).
- Advise and assist researchers on RDM (applying).
- List existing support (staff and services) on RDM for researchers (remembering).
- Define new requirements based on needs from researchers (creating).
- Define new features based on needs from researchers (creating).
- Communicate with technical and support staff about RDM related services (applying).

- List existing data-infrastructure and tools for research data (understanding).
- analyse needs of the researchers with respect to data-infrastructure and tools (analysing).
- Assess the quality and purposefulness of data-

4) Services.

Responsible for delivering sufficient adequate support on RDM for the involved researchers.

5) Infrastructure.

Responsible for monitoring the department's or project group's needs regarding data-infrastructure and tools

for RDM, including supporting adequate access, in accordance with the DMP.

- Supports access to data-infrastructure and tools for RDM.
- Sets requirements for data-infrastructure and tools for RDM.

- Knowledge about RDM issues related to data-infrastructure and tools in the data life cycle.
- analyse needs from researchers regarding data-infrastructure and tools for research data.
- Continuously monitor and explore new trends and needs in data-infrastructure and tools for RDM.
- Solve RDM related issues regarding data-infrastructure and tools in the research data life cycle.

- infrastructure and tools for projects (evaluating).
- List channels for finding new trends and needs in data-infrastructure and tools for RDM (remembering).
- Monitor and explore the channels (applying).
- Identify trends and needs (applying).
- List the RDM issues regarding data-infrastructure and tools in the research data life cycle (remembering).
- analyse each issue and the causes and effects (analysing).
- Develop a plan for solving the issues (creating).
- Explain the needs of the researchers and relevant stakeholders (understanding).
- Design training tailored to the needs of the researchers and relevant stakeholders (applying).
- Give training (applying).
- Assess RDM knowledge and skills (evaluating).
- Able to adapt generic data solutions to domain specific data (applying).

6) Knowledge management. Responsible for determining the adequate level of knowledge and skills on RDM within the department or project group in order to comply with the institute's RDM policy.

- Monitors RDM skills in the department or project.
- Identifies knowledge and skill gaps, including understanding of the DMP, and takes action if needed.
- Initiates and provides training on RDM, tailored to needs of the researchers involved in department or project.
- Creates awareness on RDM among researchers and explains the added value of RDM.
- Introduces to new employees the institute's RDM policy as well as the department's or project's DMP.

- Knowledge about the required level of knowledge on RDM for researchers and relevant stakeholders.
- Knowledge about relevant RDM topics for researchers and relevant stakeholders.
- Assess knowledge status of researchers and relevant stakeholders, and if needed give training in RDM.
- Knowledge about domain specific data and domain specific handling of such data.
- Identify knowledge and skill gaps, and takes action if needed.

7) Network.

Responsible for liaison and alignment of RDM within the department or project group and with relevant stakeholders outside the department or project group.

- Refers researchers to other RDM related facilities and services (legal, financial or operational), inside and outside the department or project.
- Liaises with RDM-related experts inside and outside the department or project.
- Maintains a network with colleagues and other relevant departments and projects.

- Knowledge about where to find department or project stakeholders and researchers, including relevant networks.
- Liaise with department or project stakeholders, researchers and other data stewards and establish an active network.
- Communicate with a diverse range of stakeholders.

- Explain where to find department or project stakeholders and researchers (remembering).
- Liaise with department or project stakeholders and researchers (applying).
- Establish an active network in which regular consultations are taking place (evaluating).

8) Data archiving.

Responsible for identifying the requirements of adequate support and data infrastructure for FAIR and long-term archiving of data of the department or project group by researchers, including selection of data, and sustainable and legitimate access to data sources of the department or project group, for the required period.

- Monitors the department's, project's or data collection's internal and/or external data archiving and access policy.
- Monitors the internal and/or external archiving of data by researchers of the department or project.
- Advises and supports researchers in the selection of data to be archived.
- Assesses whether internal and/or external data storage and archiving facilities meet the applicable requirements.

- Knowledge on how to audit archiving procedures and workflows (ways of executing archiving duties).
- Knowledge on how to change management of archiving policies.
- Knowledge on collection management and dissemination ways.
- Audit the internal and/or external archiving of data.
- Monitor the relevance of data policies.
- Execute data archiving together with researchers.
- Develop, implement and monitor the project's internal and/or external data archiving and access policy compliant with institute's policies.
- Communicate about data archiving policies and their implicit requirements to the researcher and data-infrastructure and tools.
- Interaction with external stakeholders like service providers and internal stakeholders.
- Communicate the essential factors/elements for data archiving and access.
- Monitor the execution of archiving workflow

- Develop ways of working for researchers to comply with archiving policies (create).
- Select data to be archived to ensure that the goals of the archiving policy are met (applying).
- Implement collection management (applying).
- Explain the institutes internal/external research data archiving and access policy (understanding).
- Execute the policy (applying).
- Assess internal/external data storage and archiving facilities (evaluating).
- Assess the internal and/or external research data storage and archiving facilities with regard to project needs (evaluating).
- Select available data archiving solutions and procedures for researchers and support staff (analysing).
- Communicate about data archiving solutions and procedures for researchers and support staff (applying).

and use of services.

- Communicate a data archiving and access policy and monitor the compliance.
- Give advice to researchers about data archiving, including guidance on choosing the right archive.
- Select the archiving services which are relevant for a project given the requirements from institute's policies.
- Select available data archiving solutions and procedures for researchers and support staff.
- Communicate about data archiving solutions and procedures for researchers and support staff.
- Knowledge about metadata standards.

* *Researcher* includes senior researchers, postdocs, PhD students and students

Data-infrastructure includes all technical infrastructure aspects necessary for data management and data handling for research

Research includes all disciplines and methodologies, and qualitative as well as quantitative research

RDM = Research Data Management

DMP = Data Management Plan

ELSI = Ethical, Legal, Social Implications

FAIR = Findable, Accessible, Interoperable, Reuseable

Data steward – infrastructure			
<i>Keywords: Data-infrastructure, operations, IT support</i>			
<i>Focus on technicians, infrastructure experts</i>			
Responsibilities	Activities and tasks	Knowledge, Skills and Abilities	Learning Objectives <i>(after successful completing training you will be able to)</i>
<p>1) Policy/strategy. Responsible for identifying the requirements of adequate data-infrastructure and tool landscape that fits with the institute's RDM policy and supports FAIR data and Open Science.</p>	<ul style="list-style-type: none"> - Translates the RDM policy to data-infrastructure and tools requirements, tailored for the departments, projects or data collections of the institute. - Creates an overview of available data-infrastructure and tools of the institute. 	<ul style="list-style-type: none"> - Knowledge on how the RDM policy can be implemented into infrastructure and tools for the institute. - Create an overview of available data-infrastructure and tools of the institute. 	<ul style="list-style-type: none"> - List the institute's RDM policy (remembering). - Map the institute's RDM policy to existing data-infrastructure and tools (applying).
<p>2) Compliance. Responsible for compliance of the data-infrastructure and tool landscape to the institute's RDM policy, relevant codes of conduct and legislation.</p>	<ul style="list-style-type: none"> - Monitors and supervises the use of the data-infrastructure and tool landscape in line with the institute's RDM policy, relevant codes of conduct and legislation. - Identifies gaps and takes action if needed. - Consults with privacy and security officers. 	<ul style="list-style-type: none"> - Knowledge about relevant Open Source solutions. - Knowledge of best practices in data protection. - Knowledge of relevant legislation and code of conducts with regard to data-infrastructure and tools. - Translate ethical and technical requirements for data-infrastructure and tools to technological measures, while understanding the research requirements and limitations. - Identify gaps between the legal requirements for data protection and the institute's RDM policy. - Take action upon identified gaps between data security and the institute's RDM policy. 	<ul style="list-style-type: none"> - List relevant open source solutions (remembering). - List best practices in data protection (remembering). - List relevant legislation and code of conducts (remembering). - List security measures necessary to be compliant to legislation and codes of conduct (remembering). - Assess the infrastructure requirements for researchers (evaluating). - Identify gaps between the legal requirements for data protection and the institute's RDM policy (evaluating). - Take action upon identified gaps between data security and the institute's RDM policy (creating).

3) Alignment with FAIR data principles.

Responsible for alignment of the data-infrastructure and tool landscape to the FAIR data principles and the principles of Open Science, and for facilitating and supporting FAIR data.

- Monitors and supervises the use of the data-infrastructure and tool landscape on alignment with the FAIR data principles.
- Identifies gaps and takes action if needed.

- Knowledge about the FAIR data principles and the principles of Open Science, including knowledge about implementation and how to be compliant to the FAIR data principles.
- Knowledge of data-infrastructure and tools supporting and producing FAIR data.
- Knowledge about metadata data-infrastructure and tools to expose metadata (OAI-PMH, FAIR data points, SPARQL endpoints).
- Knowledge about metadata serialisations (linked data, XML, json, json-ld).
- Translate the FAIR data principles into data-infrastructure and tool requirements.
- Identify gaps between data-infrastructure and tools and the FAIR data principles.
- Take action upon identified gaps between data-infrastructure and tools and the FAIR data principles.

- Explain what compliance to the FAIR data principles entails (understanding).
- List FAIR data infrastructures and tools (remembering).
- List metadata infrastructures and tools (remembering).
- Express metadata (serialisations) with different data-infrastructure and tools (applying).
- Translate the FAIR data principles into data-infrastructure and tool requirements (applying).
- Identify gaps between data-infrastructure and tools and the FAIR data principles (evaluating).
- Take action upon identified gaps between data-infrastructure and tools and the FAIR data principles (creating).

4) Services.

Responsible for identifying the requirements of adequate data-infrastructure for RDM for researchers and responsible for providing access to this data-infrastructure and tools and adequate support for researchers.

- Advises and supports researchers on data-infrastructure and tools.
- Develops and implements (standard) IT solutions for recurring issues and questions.
- Provides guidance and instruction on the use of data-infrastructure and tools.

- Knowledge about RDM related issues in the research data life cycle with regard to available data-infrastructure and tools.
- Knowledge about the research data flow.
- Advise and assist researchers on short and long term actions for data-infrastructure and tools.
- Assess and analyse needs from researchers regarding support on data-infrastructure and tools.
- Communicate with researchers, technical staff and support staff about data-infrastructure and tools.
- Knowledge and understanding of researchers needs.

- Define the short and long term actions for RDM data-infrastructure and tools (remembering).
- Explain what kind of advice and assistance on data-infrastructure and tools is usually needed by researchers (understanding).
- Advise and assist researchers in choosing the best IT solution for their data management (applying).
- List existing IT support (staff, data-infrastructure and tools) on RDM for researchers (remembering).
- Communicate with technical and support staff about RDM related services (applying).

5) Infrastructure.

Responsible for identifying the requirements of adequate data-infrastructure and tool landscape that fits with the needs of the researchers, with the institute's RDM policy and supports FAIR data and Open Science. Responsible for making an inventory for existing data-infrastructure and tools that fit with the institute's needs.

- Requests and acquires data-infrastructure and tools for RDM within the institute.
- Monitors the need, use and availability of data-infrastructure and tools.
- Is aware of and advises the institute, department and researchers on relevant (inter)national data-infrastructure and tools.
- Takes part in initiatives to further develop data-infrastructure and tools.
- Explores new trends in data-infrastructure and tools for RDM.

- Knowledge about RDM related issues in the research data life cycle with regard to available data-infrastructure and tools.
- Assess and analyse needs regarding data-infrastructure and tools for RDM for the institute, department and researchers.
- Translate needs to requirements and advise to the institute, department and researchers.
- Continuously monitor data-infrastructure and tools available inside and outside the institute, in close collaboration with the responsible (IT) department.
- Explore new trends in data-infrastructure and tools and translate these to advise for the institute, department and researchers.
- Assure proper alignment of the data-infrastructure and tools of the institute to the relevant external data-infrastructure and tools.
- Knowledge of encryption of data and data security protocols.

- List existing data-infrastructure and tools for research data (understanding).
- Making an inventory of available data-infrastructure and tools (remembering).
- Analyse needs of the researchers with respect to data-infrastructure and tools (analysing).
- Identify gaps in the data-infrastructure and tool landscape (applying).
- Explore channels for finding new trends and needs in data-infrastructure and tools for RDM (applying).
- Monitor and explore the channels (applying).
- List the RDM issues regarding data-infrastructure and tools in the research data life cycle (remembering).
- Analyse each issue and the causes and effects (analysing).
- Develop a plan for solving the issues (creating).
- Formulate data-infrastructure and tooling advise for the management of the institute in collaboration with the relevant institutional stakeholder (creating).
- Define new IT requirements and features based on needs from researchers (creating).
- Explain the technical RDM related needs of the researchers and relevant stakeholders (including other data stewards) (understanding).
- Design training tailored to the technical RDM related needs of the researchers and relevant stakeholders (applying).
- Give training (applying).
- Assess technical RDM related knowledge and skills (evaluating).

6) Knowledge management.

Responsible for determining the adequate level of knowledge and skills on RDM related data-infrastructure and tools in the institute, department and among researchers.

- Monitors technical RDM related skills in the institute, department or project.
- Identifies technical knowledge and skill gaps, and takes action if needed.
- Initiates and provides training on data-infrastructure and tools for RDM.
- Creates awareness on data-infrastructure and tools among the institute,

- Knowledge about the required level of technical knowledge on RDM for researchers and relevant stakeholders.
- Knowledge about relevant RDM data-infrastructure and tools for researchers and relevant stakeholders.
- Assess technical knowledge status of researchers and relevant stakeholders (including other data stewards), and if needed give training in technical RDM skills.
- Identify knowledge and skill gaps, and takes action if needed.

department and researchers and explains the added value to RDM.

7) Network.

Responsible for liaison and alignment of the data-infrastructure and tools management within the IT group, institute or department with relevant stakeholders outside, such as (inter)national data-infrastructure groups.

- Refers researchers and research support staff to RDM related data-infrastructure and tools, inside and outside the institute.
- Liaises with (technical) RDM experts inside and outside the institute.
- Maintains a network with (technical) colleagues and other relevant stakeholders.

- Knowledge about where to find institute, department or project stakeholders and researchers, including relevant networks.
- Liaise with institute, department or project stakeholders, researchers and other data stewards and establish an active network.
- Liaise and communicate with a diverse range of stakeholders about data-infrastructure and tools.

- Have an overview who at the institute needs or provides RDM related data-infrastructure and tools (remembering).
- Inform researchers and other data stewards about (the usage of) data-infrastructure and tools (applying).
- Understand the researcher's needs and priorities (applying).
- Establish an active network in which regular consultations are taking place (evaluating).

8) Data archiving.

Responsible for identifying the requirements of adequate data-infrastructure and tools for FAIR and long term internal and/or external archiving of data of the institute.

- Advises on data-infrastructure and tools for data archiving services.
- Identifies the institute's, department's or researcher's needs into infrastructural requirements.
- Assesses whether internal and/or external data storage and archiving facilities meet the applicable requirements.
- Monitors and evaluates data-infrastructure and tools that best fit the institute's RDM policy.
- Advises on (meta)data formats for data archiving.

- Knowledge about data archiving, digital preservation and data curation.
- Knowledge about repository and preservation software and hardware.
- Knowledge on (meta)data standards for data archiving.
- Translate needs and requirements into archiving solutions.
- Advise researchers and stakeholders (including data stewards) on archiving solutions, including (meta)data standards.
- Evaluate which data-infrastructure and tools best fit the institute's RDM policy.

- List available data archiving solutions and procedures for researchers and support staff (remembering).
- List the requirements regarding internal/external data storage and archiving services (remembering).
- List available data repository and preservation software and hardware (remembering).
- Communicate with users about data archiving solutions for researchers and support staff (applying).
- Map the institute's RDM policy to data-infrastructure and tools for data archiving (applying).

* *Researcher* includes senior researchers, postdocs, PhD students and students

Data-infrastructure includes all technical infrastructure aspects necessary for data management and data handling for research

Research includes all disciplines and methodologies, and qualitative as well as quantitative research

RDM = Research Data Management

DMP = Data Management Plan

ELSI = Ethical, Legal, Social Implications

FAIR = Findable, Accessible, Interoperable, Re-useable

Appendix 4 – Mapping

This Appendix provides the mapping of the defined data steward function (responsibilities and tasks) and competencies (knowledge, skills and abilities) for all three data steward roles on existing data stewardship principles and approaches.

[This relates to Deliverable 1 and 2 of the project proposal]

Competence area	Mapping (for references see below)				
	Research data life cycle	FAIR principles	Purdue competencies areas	DAMA knowledge areas	EDISON data science framework
1) Policy / strategy	① ② ③ ④ ⑤ ⑥	FAIR	③ ④ ⑧*	① ⑪*	① ⑤*
2) Compliance to legislation and codes of conduct	① ② ③ ④ ⑤ ⑥	FAIR	④ ⑧ ⑫*	⑤ ⑪*	⑤*
3) Alignment with FAIR data principles	④ ⑤ ⑥	FAIR	② ⑤ ⑥ ⑦ ⑨*	④ ⑥ ⑩*	③*
4) Services	① ② ③ ④ ⑤ ⑥	FAIR	③*	①*	① ④*
5) Infrastructure	① ② ③ ④ ⑤ ⑥	FAIR	① ②*	② ④*	① ② ③*
6) Knowledge management	① ② ③ ④ ⑤ ⑥	FAIR	③ ⑧ ⑫*	① ⑨*	① ④*
7) Network	① ② ③ ④ ⑤ ⑥	FAIR	③ ⑧ ⑫*	① ⑨*	④*
8) Data archiving	④ ⑤ ⑥	FAIR	② ⑤ ⑥ ⑦ ⑨*	④ ⑥ ⑩*	③*

* indicates that this topic involves potentially all areas, with emphasis on the indicated areas

References

Research data life cycle (<https://www.ukdataservice.ac.uk/manage-data/lifecycle>)

① planning data; ② creating data; ③ processing & analysing data; ④ publishing & sharing data; ⑤ preserving data; ⑥ reusing data

FAIR principles (<http://www.nature.com/articles/sdata201618>)

(F) findable; (A) accessible; (I) interoperable; (R) reusable data

Purdue competence areas (https://docs.lib.purdue.edu/lib_fsdocs/136)

① databases and data formats; ② discovery and acquisition of data; ③ data management and organisation; ④ quality assurance; ⑤ data conversion and interoperability; ⑥ metadata; ⑦ curation and re-use; ⑧ cultures of practice; ⑨ data preservation; ⑩ data analysis; ⑪ data visualisation; and ⑫ ethics, including citation of data

DAMA knowledge areas (<https://dama.org/content/body-knowledge>)

① data governance; ② data architecture; ③ data modelling and design; ④ data storage and operations; ⑤ data security; ⑥ data integration and interoperability; ⑦ documents and content; ⑧ reference and master data; ⑨ data warehousing and business intelligence; ⑩ metadata; and ⑪ data quality

EDISON data science framework (<https://github.com/EDISONcommunity/EDSF/>)

① plan; ② build; ③ run; ④ enable; ⑤ manage

Appendix 5 - Data steward workshops

This Appendix provides the detailed description of the workshops organised in June 2019. The presentations are available at [Zenodo community pages](https://zenodo.org/communities/nl-ds-pd-ls/?page=1&size=20) (<https://zenodo.org/communities/nl-ds-pd-ls/?page=1&size=20>)

[This relates to Deliverable 4 of the project proposal]

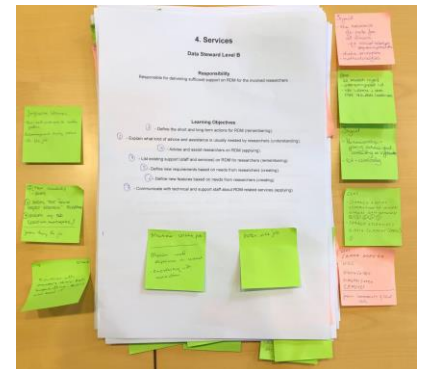
The workshops were structured according to the various elements of the project:

- Working session 1: function description and role
- Working session 2: competencies (knowledge, skills and abilities)
- Working session 3: training
- Working session 4: future steps

Based on the workshops, the main conclusions are:

Function description and competencies

- Data stewardship is often more a role than a function
- Data stewardship is so new that current data stewards are 'five-legged-sheeps'
- There is a high need for clear function profiling
- Moving towards data stewardship (and roles, tasks and responsibilities) in an organisation is challenging, particularly when it comes to capacity and financing
- The project's matrices with competence areas, responsibilities, tasks, KSAs and learning objectives is recognisable, elaborated enough and very adequate to the workshop participants
- The Venn diagram (organisation oriented) could additionally be converted into a spider diagram (focussed on personal competencies) to be used as an assessment tool
- Connect to NFU and VSNU for opening discussion about getting the data steward role in the formal "functie huizen"



Training

- There is a high need for training, as most training is currently done 'on the job', for instance in the context of RDNL/Essentials 4 Data Support. See overview Appendix 6
- On the other hand, there currently is an information overload for the data stewards. Online, much information is available, but it is difficult to navigate and difficult to assess the quality. A condensed, structured way to offer the available information and courses would be very helpful to data stewards
- If courses could get a quality indication including certification, this could be useful for data stewards to show they have the competencies and for employers to be able to hire qualified persons

Future steps

- Work on estimating the capacity (fte) needs for implementing good data stewardship in an organisation
- Most needed: data stewards want to know how to engage with their boards, deans and policy maker, how to convince them of the needs, and how to bring the message about data stewardship

Appendix 6 - Overview existing and desired training reported by data stewards

This Appendix provides an overview of existing and desired training for new and experienced data stewards. This overview was mainly built by the input given by data stewards during the workshops.

The tables below represent the answers from the workshop participants on their existing training and sources of knowledge and desired training according to all eight competence areas described by this project. At the time of the workshops, 'data steward - infrastructure' was not completed yet, therefore the training for this specific role is left out.

In addition, in the next Appendix (Appendix 7) a selection of available training materials and pointers to resources are given. This was based on input during the workshops and on input from the project team and consultation commission. Further analysis of these data will be undertaken in the follow-up project.

[This relates to Deliverable 3 of the project proposal]

Area 1: Policy / strategy

Data steward - policy	Data steward - research ¹²
Existing training	Existing training
<ul style="list-style-type: none"> - Essentials 4 data support (RDNL) - Participating in European projects (EUDAT, EOSCpilot, etc) or reading their output - Digital Curation Centre (online guides) - LEARN project outputs (LERU) - policy examples from other universities in Europe - From colleagues, learning on the job, seeing more and more DMPs, literature, conferences & workshops 	<ul style="list-style-type: none"> - Essentials 4 data support (RDNL) - e-learning from CESSDA - Digital Curation Centre (online guides) - Master educational sciences - By delivering training in data management - Meeting with researchers - Data policies from other faculties - Discussion with own policy makers - info sessions - My own experience as a PhD in a data intensive research project - From colleagues, learning on the job, seeing more and more DMP, literature, conferences & workshops
Desired training	Desired training
<ul style="list-style-type: none"> - Project management - Policy writing and communication 	<ul style="list-style-type: none"> - Monitoring and monitoring role - More hands-on training DMP (what are the criteria)

¹² At the time of the workshops, data steward-infrastructure was not completed yet, therefore the training for this this specific role is left out

	<ul style="list-style-type: none"> - How to bring a policy to researchers successfully. - Networking - Policy development - Group dynamics - Global development in RDM - Training on strategy, communication and engagement - Learn more on specific (technical) requirements of RDM, for different scientific domains. - Learn more about different existing DMPs
--	--

Area 2: Compliance

Data steward - policy	Data steward - research
Existing training	Existing training
<ul style="list-style-type: none"> - NPOS - Essentials 4 data support (RDNL) - Online courses - Publications/literature - Online information at institutes - Own research - BROK course (good clinical practice) - Reading codes of conduct - Talking with the university's data protection officer - CESSDA/uk data science training materials - OPENAire training materials - On the job and from colleagues, conferences/workshops, twitter, textbooks 	<ul style="list-style-type: none"> - NPOS - Essentials 4 data support (RDNL) - Course on GDPR form DANS - Info session GDPR or privacy training at the institute - Discussion with information manager/privacy officer, on the job - Work group with ethical review board - BROK course (good clinical practice) - Privacy training - Previous role as monitor WMO/GCP compliant research - CCMO (centrale commissie mensgebonden onderzoek)
Desired training	Desired training
<ul style="list-style-type: none"> - How to engage stakeholders to comply (sticks & carrots) - Legal requirements (musts) and best practices (shoulds) and institute requirements (local shoulds/musts) and the difference between them. - How to create broad support in the institute - How to translate regulations (like GDPR) into practical steps for researchers - Skills how to translate legal policy/laws into 	<ul style="list-style-type: none"> - Knowledge about legislation and how to create guidelines on RDM - Staying up to date - More knowledge about copyright, privacy, ethics - How to make researchers aware of services/support to become compatible with legal and ethical guidelines/codes. - Course on GDPR and (practical) implications for research

<ul style="list-style-type: none"> - practical guidelines - Need to learn how to read legal text regarding data protection 	<ul style="list-style-type: none"> - List of key skills researchers should have for the highest impact of my proposed solutions - To learn how to register the data and check their compliance with the GDPR.
--	---

Area 3: Alignment with fair data principles

Data steward - policy	Data steward - research
Existing training	Existing training
<ul style="list-style-type: none"> - Info sessions about open science and metadata workgroup - Online courses - Internal experts - Essentials 4 data support (RDNL) - Librarianship - GO FAIR - By doing in my role as coordinator of the UMC & RDM working group - FAIR data expert group (European commission) especially the report on 'turning fair data into reality' - On the job, Literature, knowledge exchange, read a lot about it 	<ul style="list-style-type: none"> - UU data management community - UMCG data management meetings - GO FAIR - FAIRytale meetings with researchers - Resources on website of own and other institutes - Essentials 4 data support (RDNL) - Work experience in research - I work with colleagues who know a lot about PIDs and controlled vocabularies - Article on FAIR data - On the job, journals, guidelines, literature
Desired training	Desired training
<ul style="list-style-type: none"> - Deliver bring your own data FAIRifications workshops for all disciplines - FAIR data & connection to publishing - Deal with "stubborn" researchers (communication level) - Not to misuse the term FAIR data and implement the principles of FAIR data - Learn state of the art related to FAIR data in different domains - Learn about FAIR metrics - Skills to apply, implement/enhance: metadata standards and the use of persistent identifiers 	<ul style="list-style-type: none"> - Learn, know and use tools and services to make data fair - Practical application/use of metadata - Communicate well with researchers - Application of FAIR data

Area 4: Services

Data steward - policy	Data steward - research
Existing training	Existing training
<ul style="list-style-type: none"> - Keeping up with trends and information (information research) - Essentials 4 data support (RDNL) - Workshops/webinars - Community - By delivering training - Creating a service catalogue for our university - Meeting with researchers - Talk with people in conference - On the job, working routine, online/literature/google, colleagues 	<ul style="list-style-type: none"> - By performing research yourself - Data management consultations in the institute - By being involved in an institutional data programme - Discussion with members library, RDM support office, central and local it. - Member of the works council - Librarianship (= giving advice and consulting on information) - ICT curiosity - As coordinator of international data network - Essentials 4 data support (RDNL) - Course on information literacy (TUDelft) - Communicating with scientists (translating) - On the job, from colleagues and by doing
Desired training	Desired training
<ul style="list-style-type: none"> - Strategic skills to build capacity - Stakeholder needs for a service: sending the best possible service experience - More knowledge about data in general and domain specific - Structured introduction to computational reproducibility - Knowing more about RDM training at national level/other institutes. Is there a catalogue? 	<ul style="list-style-type: none"> - Research life cycle for all domains (e.g. social science, experimental, etc) - Data analysis - Methodologies - Learn more on "kos" (knowledge organisation, services) - Improve communication & social skills - Use of docker - Useful services that would be endorsed by researchers - Repositories for relevant disciplines - Tool/platform facilitating requirement specification, collection and communication? - Reproducibility principles - Data reuse - Listing of effective services - examples

Area 5: Infrastructure

Data steward - policy	Data steward - research
Existing training	Existing training
<ul style="list-style-type: none"> - Work experience in research 	<ul style="list-style-type: none"> - Meeting with IT, other data stewards and

<ul style="list-style-type: none"> - 4.TU research data work experience - LCRDM information and meetings - Essentials 4 data support (RDNL) - Meeting with researchers - Meetings with central and local IT departments - By asking many naive questions to our IT team while preparing guidelines for the institute - Postdoctoral research experience - Software engineering experience - Standard development and specification experience - Working in IT, support, maintaining servers - Learning on the job - Re3data.org - ESFRI 	<ul style="list-style-type: none"> researchers - Update meetings with IT - 4TU research data work experience - LCRDM information and meeting - Essentials 4 data support (RDNL) - Master in medical informatics - IT-related education background - Learn by doing/on the job - Work experience in research or during PhD/postdoc job - Work experience in IT, IT support, maintaining servers, etc. - WP lead on IT work package EHRI project (holocaust data & metadata) - MSc Computer Science at the University of Liverpool - Software engineering experience - Standard development and specification experience
Desired training	Desired training
<ul style="list-style-type: none"> - Basic skills (at least good understanding) in AI, Algorithms, machine learning - Standardised evaluation matrices - More on RDM vs. DMPs (even if it is only terminology) - DMP online etc. - How to intervene into/cooperate with researchers about software and hardware choices - Organisational model for RDM-ICT support and applications - Knowing which types of infrastructures are out there for different types of data (clinical, genetic, imaging, audio, etc.) 	<ul style="list-style-type: none"> - Basic skills (at least good understanding) in AI, Algorithms, machine learning - Proper monitoring training - Existing infrastructure/tools - Think about an infrastructure for data collection - Task list for ICT roles on RDM within research groups rather than external steward in 1 person - Discipline specific repositories - Organisation like CLARIN, CESSDA, RDA: what do they offer and how to get in? - DTL FAIR data portal - FAIR Data Point

Area 6: Knowledge management

Data steward - policy	Data steward - research
Existing training	Existing training
<ul style="list-style-type: none"> - Essentials 4 data support (RDNL) - Former jobs: how to form policy, top down vs bottom up, ways of communicating in 	<ul style="list-style-type: none"> - On the job - Working as a PhD/ in research - Lab meetings, meetings/trainings for researchers

<ul style="list-style-type: none"> healthcare organisations - Work experience in research - Read institute's policy - Meeting with researchers - Online resources/publications - On the job - Creating CESSDA data management expert guide (for social sciences) - Creating course at TUDelft (in cooperation with Education Services) - Interaction with colleagues and networking 	<ul style="list-style-type: none"> - Essentials 4 data support (RDNL) - Teacher training - Former jobs: how to assess & improve knowledge and behaviour with employees. - Designing & facilitating workshops and training - Software carpentry instructor training - Books: Non-invasive governance & research data management - Working long enough in scientific environment with sense for process optimisations, facilitation and efficiency - Literature - Online training resources (FOSTER training portal, CESSDA Expert tour guide on RDM)
Desired training	Desired training
<ul style="list-style-type: none"> - Lesson on data/research policy - Policy making - Staying up to date - Knowledge about the different data domains, workflows, research within the institute - How to work with people outside your own university so that you don't have to reinvent the wheel & join forces which improves overall quality of policy matters in RDM knowledge - Professionalise training - Knowledge & skills on development of education and training - How to prepare re-usable, modular training material - How to prepare quizzes on RDM 	<ul style="list-style-type: none"> - Need more training in coding/programming to support researchers - Handle data & scientists. Education skills: what works with scientists; types of learning and teaching - Giving lessons data literacy (learning objectives) - Domain specific knowledge/skills - Design and give training - Practical application of metadata (which one? how?) - Get more insight into tools and legal aspects. This should be a policy/knowledge database in the NL

Area 7: Network

Data steward - policy	Data steward - research
Existing training	Existing training
<ul style="list-style-type: none"> - UKB working group Research data - RDM programme at university - LCRDM - RDA - Liber 	<ul style="list-style-type: none"> - On the job/learned by doing - Colleagues, normal workplace interaction and asking others, Work experience/work socialising - Experience from PhD research - Study psychology

<ul style="list-style-type: none"> - Stakeholder analysis (as part of a training) - Via the organisation structure of my university - Learn by doing, from colleagues and on the job 	<ul style="list-style-type: none"> - Via internal DS network at Radboud University - RDM lists on the internet (email lists) - Conferences (IDCC, RDA) - Data stewards interest group
Desired training	Desired training
<ul style="list-style-type: none"> - Social engineering: how to interact with different personality types - How to influence people without having a manager position 	<ul style="list-style-type: none"> - "Social engineering" how to interact with different personality types - I would like to know more on liaise with departments or project - How to motivate other parties so they have good feeling about my requirements/tasks - Communication skills

Area 8: Data Archiving

Data steward - policy	Data steward - research
Existing training	Existing training
<ul style="list-style-type: none"> - Tailored session at the DANS archive - Talks with 4TU DC - By creating catalogues for Leiden researchers - Meeting with 4 TU, 4TU website - Publications on archiving - Data stewards - Own research - Archive school - Learning on the job and from colleagues. - Examples from other universities and from existing data repositories/archives - Digital curation centre 	<ul style="list-style-type: none"> - UU data management community - On the job (took me years though) and from colleagues - BA + MSc on the topic - DIY research online, network, etc. - Contact 4 TU, DANS, etc. - Reinward Academie/ Archief School University (IT field) - Blog webs: "good/best practice of ..." - Data formats learned by working on specific projects. - Via tailored sessions at data archives DANS and 4TU.DC
Desired training	Desired training
<ul style="list-style-type: none"> - How to create what type of FAIR data with the aid of what type of archive - Matching procedures in the Netherlands 	<ul style="list-style-type: none"> - Learn about metadata & different types of repositories - Workflow management - Know more about external options for archive data, pros and cons - How to archive data + which archives are available and suitable. - Learn about subject specific repositories - Overview repositories, general characteristics

- | | |
|--|--|
| | <ul style="list-style-type: none">- Demands repositories- External demands: how to deal with these- To learn how to store data- Digital preservation and curation |
|--|--|

Appendix 7 - Recommended basic training for data stewards

During the project workshops in June 2019, the following informal training resources were recommended by existing data stewards for new data steward, to acquire the basics of data stewardship and data management:

Foster Open Science	Free e-learning platform with mainly Open Science focussed content https://www.fosteropenscience.eu/ [Last accessed 16.07.2019]
MANTRA	Free online course about research data management https://mantra.edina.ac.uk/ [Last accessed 16.07.2019]
Essentials 4 Data Support	Introductory course for data support and research data management. Free online course, as well as certified full course https://datasupport.researchdata.nl/en/ [Last accessed 16.07.2019]
Research Data Management and Sharing	Introductory course for novices in research data management. Free online course, as well as certified full online course https://www.coursera.org/learn/data-management [Last accessed 16.07.2019]
CESSDA Data Management Expert Guide	Free online guide from European experts on social science research https://www.coursera.org/learn/data-management [Last accessed 16.07.2019]
NFU Data 4 Lifesciences HANDS 20	Free online “Handbook for Adequate Natural Data Stewardship” (HANDS 2.0) https://data4lifesciences.nl/hands2/data-stewardship/ [Last accessed 16.07.2019]

Additionally, for each of the competence areas the following resources were mentioned as useful:

1. POLICY / STRATEGY	Essentials 4 Data Support, chapter 5: Legislation and Policy
-----------------------------	---

<https://datasupport.researchdata.nl/en/start-the-course/v-legislation-and-policy/>
[Last accessed 16.07.2019]

2. COMPLIANCE	RUG guidance on data protection & GDPR
https://www.rug.nl/research/research-data-management/data_protection-gdpr/?lang=en [Last accessed 16.07.2019]	

3. FAIR DATA ALIGNMENT	<ol style="list-style-type: none">1. DANS 'FAIR assessment tool'2. ANDS NECTAR RDS 'FAIR tool'3. FAIRsharing 'FAIR Evaluation Service'4. HANDS 2.0
<ol style="list-style-type: none">1. http://blog.ukdataservice.ac.uk/fair-data-assessment-tool/ [Last accessed 16.07.2019]2. https://www.ands-nectar-rds.org.au/fair-tool [Last accessed 16.07.2019]3. https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/ [Last accessed 16.07.2019]4. https://data4lifesciences.nl/hands2/data-stewardship/ [Last accessed 16.07.2019]	

4. SERVICES	MANTRA chapter about storage and security
https://mantra.edina.ac.uk/storageandsecurity/ [Last accessed 16.07.2019]	

5. INFRASTRUCTURE	Training material for the EUDAT B2SHARE service
https://github.com/EUDAT-Training [Last accessed 16.07.2019]	

6. KNOWLEDGE MANAGEMENT	LCRDM Research Data Management wiki
https://www.edugroepen.nl/sites/RDM_platform/SitePages/Home.aspx [Last accessed 16.07.2019]	

7. NETWORK	Your institutional portfolio on personal development courses
E.g. course on communication, networking, project management etc.	

8. DATA ARCHIVING

Essentials for Data Support, chapter 4:
user phase / archiving data

<https://datasupport.researchdata.nl/en/start-the-course/iv-user-phase/archiving-data/>

[Last accessed 16.07.2019]

Appendix 8 - Proposals FAIR Data Stewardship Knowledge Skills & Abilities Tool

This appendix contains two proposals that are submitted by Elevate-Health:

- Appendix 8A describes the proposed content, timeframe and budget for developing a navigation tool coupled to the matrices that we have created in our project.
- Appendix 8B describes the timeframe and budget for the first phase (Design Phase) of the project described in Appendix 8A)

Both proposals have been submitted by Eva van Ingen, Elevate-Health Business Development Manager, eva.van.ingen@elevatehealth.eu, +31 (0)30-253 7250 / +316-46700258, Heidelberglaan 8, 3584 CS Utrecht

[This relates to deliverable 3 and 5 of the project proposal]

Appendix 8A: Proposal for Navigation Tool - Proposal

To: Dr. Celia W.G. van Gelder
Programme Manager DTL Learning
Postbus 8500, 3503 RM Utrecht
e-mail: celia.van.gelder@dtls.nl

Subject: FAIR Data stewardship Knowledge, Skills & Abilities Tool

Date: August 23, 2019



Introduction

Good data stewardship is rapidly becoming an essential part of modern science. To facilitate good data stewardship and to promote open science, a broad community of international stakeholders have developed the FAIR Data principles. The Dutch Techcentre for Life Sciences (DTL) actively promotes FAIR data stewardship of life science information. In 2018, Elevate and DTL joined forces and set up a partnership, because DTL's network goes very well with the activities and ambitions of Elevate.

In the context of the ZonMW-project entitled '**FAIR Data Steward as a profession in Life Sciences in the Netherlands**' Elevate has led an exploratory advisory process in preparation of the establishment of an education line and training material for data stewards. This advisory trajectory has taken place between January and July 2019 and has resulted in a 'matrix type' blueprint that describes the required and recommended **Knowledge, Skills and Abilities** for data stewards in life sciences, further converted by the project team by formulating specific learning objectives. With this blueprint, the project team aims to increase the data steward capacity in the institutes in terms of quality and quantity.

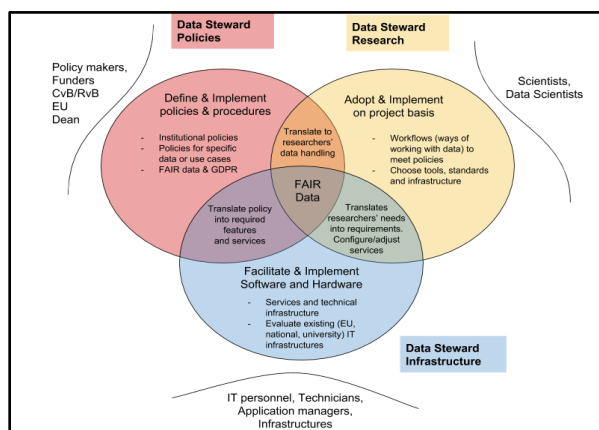


Fig.1 Data steward function landscape and other stakeholders

Request



With the overall objective to professionalize the data steward function within the life-sciences domain, the project team wishes to develop an interactive tool to encompass all the gathered information on Knowledge, Skills and Abilities corresponding to the different levels of data steward professionals. Users of the tool would be both sitting data stewards and new-to-be data stewards in the Netherlands, spread over three levels as visualized in figure 1.

By using this tool, the data steward will be able to:

1. Explore the knowledge, skills and abilities that are connected to their function
2. Self-assess their existing knowledge, skills and abilities and identify possible gaps
3. Find resources to bridge these gaps (e.g. websites, blogs, articles, existing course materials)

Already known **quality criteria** of the requested solution are listed below, although some of them have to be elaborated during the design phase. These criteria have been marked with an *.

	The tool should be:
Accessibility	<ul style="list-style-type: none"> - online accessible - easy to use and navigate - self-instructing
User orientation	<ul style="list-style-type: none"> - suitable for both formal and informal learning (on the job & just-in-time)* - suitable for different levels (junior to senior) and functions within the data steward field*
Technical interoperability	<ul style="list-style-type: none"> - browser or platform based* - platform independent
Visual & Instructional Design	<ul style="list-style-type: none"> - based on a clear instructional design - using explicit learning objectives* - offering self-assessment options - visually appealing
Content	<ul style="list-style-type: none"> - following the set-up of the function description overview (8 competence areas)* - offering mapping of (training) resources



Elevate's Vision & Solution

Based on her many years of experience in solving issues in the field of learning & development, Elevate has developed a clear vision on **continuous education**. In response to the request of the ZonMW-project Elevate envisions a platform independent tool that enables data stewards to explore the elements of the data steward function, and to validate gaps in knowledge and skills through self-assessments. Users will then make a choice from a variety of learning interventions, such as e-courses, literature, blogs and other resources, that are offered in order to reduce their knowledge gap. What has been learned is then applied in practice and, over time, new learning needs will arise. This creates a cyclical learning system, as shown in the figure below.

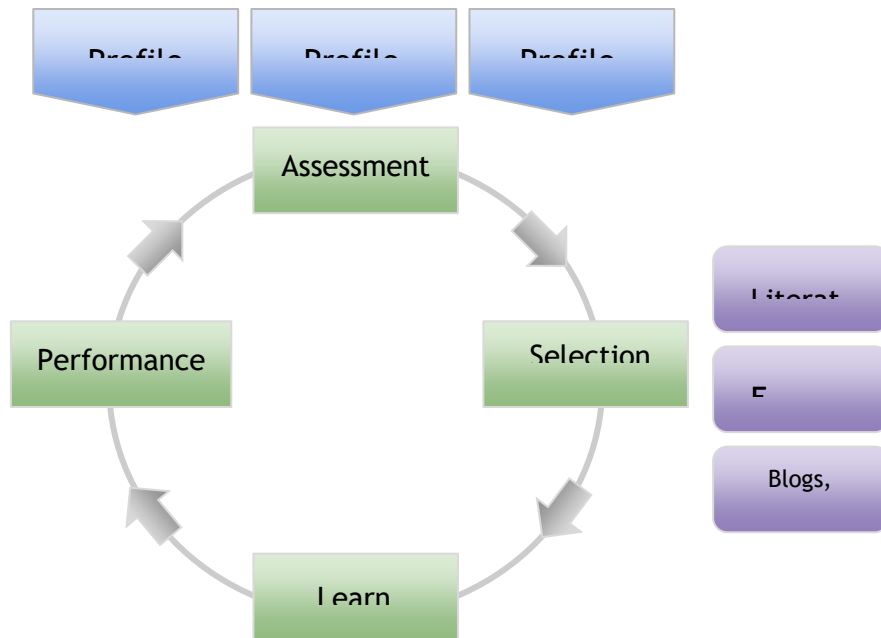


Fig. 2 Proposed cyclical learning system

Technology

Elevate recommends to use SCORM technology (Articulate 360) as a basis for the online tool. This technology offers the opportunity to create a unique look & feel or to incorporate an existing look & feel, can be navigated easily, can be hosted on different platforms, is easily accessible on different devices and is suitable for self-assessment. Next to that, this technology is very well suited for the usage of branching scenarios and cross-referencing.

Milestones & deliverables

1. Design phase

In the design phase the educational experts of Elevate will sit down with the project team to jointly design the structure that will be used in the tool. Based on the function description blueprint and the learning objectives, we will decide on the types of self-assessment that will be used, how and when to connect the results of these self-assessments to the (training) resources and the different paths/routes the tool should contain to connect to the needs of the target audience (data stewards). Next to that, a SCORM test package will be created to enable the client to run a first test implementation of the technology.

→ Deliverable: completed design template (on paper)

→ Deliverable: SCORM test package

2. Development phase

In the development phase all input for the tool will be developed by the client's project team, under guidance of Elevate. This includes instruction texts, self-assessments and information pages, based on the design. The content partner will deliver all necessary input, including (training) resources. In one review round, Elevate will review the complete materials and provide didactical advice. Next to that, a visual design will be created by Elevate, to be reviewed by the client.

→ Deliverable: completed development template (on paper), including all resources

→ Deliverable: draft visual design

3. Creation phase

In the creation phase the Elevate development team will build the tool, based on the deliverable resulting from the development phase. There will be one review round by the content partner, followed by revisions by Elevate. At the end of this phase the tool will be ready for implementation in the desired target website or platform.

→ Deliverable: tool (SCORM package) **Please note that this deliverable needs to be further specified.*

Branching scenario

The technology (SCORM) Elevate suggests to use for the development of the tool, is ideal for creating branching scenarios. As an illustration, one possible scenario, based on the existing blueprint, is visualized below. It could be used as a basis for the design phase of the project.

In this scenario, a data steward (level A) enters the tool. He or she indicates the level and selects the preferred activity (view the data steward profile, check existing knowledge, or find more information). Based on the choice, the data steward will proceed to a next page (either a page with access to the full profile, a self-assessment page or a page with relevant resources and links). On these pages, the data steward again can make a choice (e.g. selecting a category from the profile page and continue to related self-assessment questions / viewing the related resources, visit an external website or resource by selecting a link on the resources page, or receiving a specific score and personalized advice based on the score after going through the self-assessment page(s). We assume the tool will be offering three scenarios, based on the three different levels within the function description.

Examples of how a self-assessment page and scoring pages could look like within the tool can be found in Annex I.

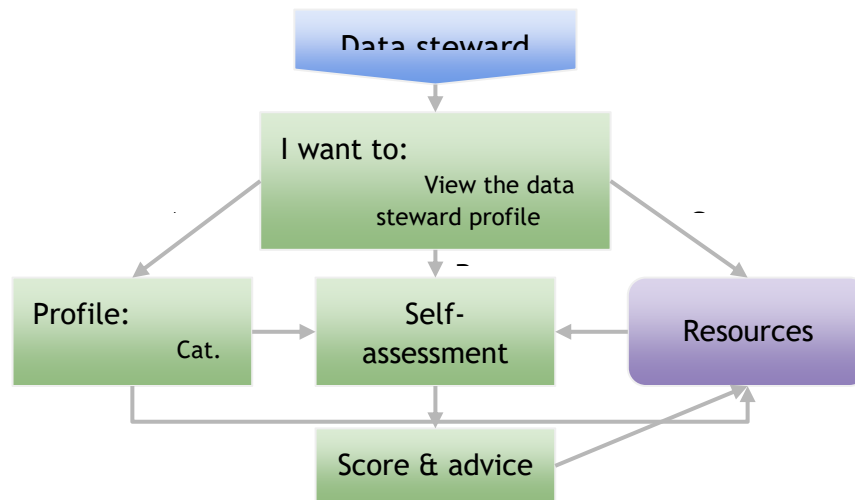


Fig. 3 Branching scenario (one specific data steward level and one specific competency category)

Time frame



For the chosen scenario in this proposal Elevate has estimated a duration of circa 5 months for the design and development of the tool, taking into account the full capacity and the availability of the content partner as agreed upon on at the kick-off of the collaboration.



Indicative investment

For the development of a web based tool, making use of SCORM technology, according to the scenario as described above Elevate has estimated an indicative investment, distinguishing the different design and development phases of the tailor-made tool (one-time costs). For Elevate’s hourly rates we refer to Annex II.

Activities Elevate	<i>One-time costs min - max</i>
Design phase	€ 3.500 - € 5.000
Creation phase	€ 3.500 - € 5.000
Visual design	€ 5.000 - € 7.500
Development	€ 16.500 - € 22.500
Test & revision	€ 2.500 - € 3.500
Projectmanagement	€ 4.500 - € 6.500
Subtotal	€ 35.500 - € 50.000
Total investment incl. 10% unforeseen	€ 40.000 - € 55.000

Please note that all amounts are given in euros, excluding VAT.



Project team & management

Upon agreement on this proposal Elevate will appoint a senior instructional project leader. He or she will take care of guiding the ZonMW-project team during designing & developing the online tool. The project leader will be assisted by other instructional designers (MSc Educational Sciences) from Elevate during the developing process. The project leader will also be your contact person during the project and will monitor planning and budget.

To guarantee smooth collaboration it is important that the ZonMW project team assigns a project leader who is allowed to give approval upon the deliverables as well.

Annex – I Examples

Please read the following case description:

[... case description ...]

Is compliance to the FAIR data principles reached?

Yes

No

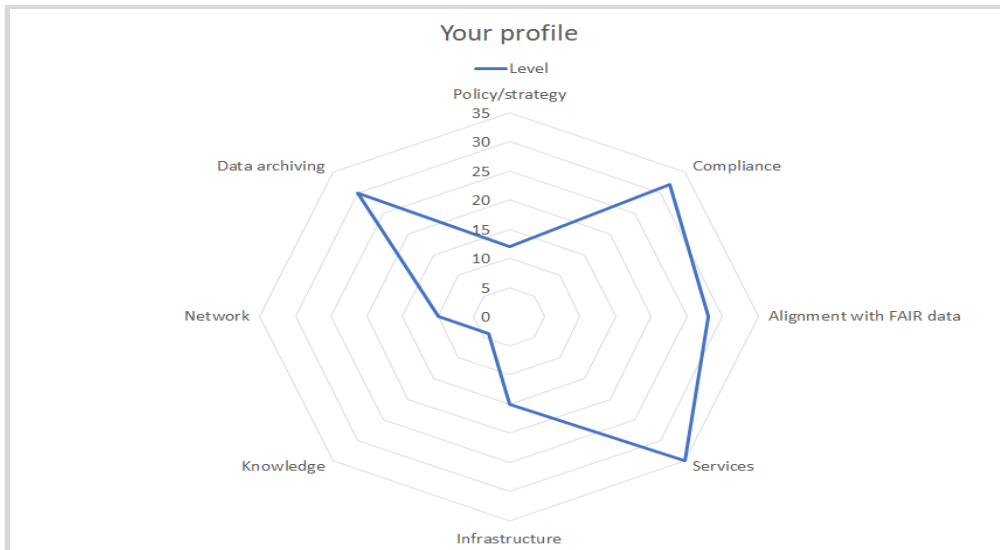


Correct

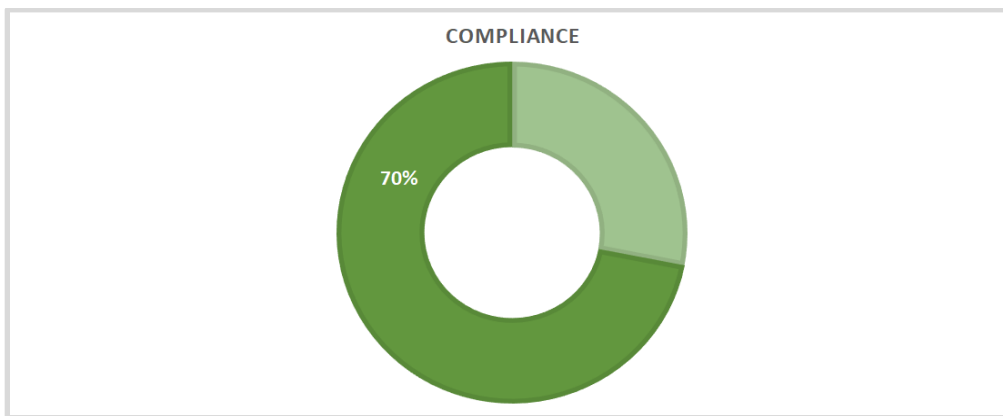
Thank you for your answer. For more information on compliance to FAIR data principles, please proceed to the resources overview.

NEXT

Example 1: self-assessment page



Example 2: score page



Example 3: score page

Annex - II – Hourly rates partners 2019

Junior Instructional Designer / Developer / E-moderator / Operations	€ 75,00
Medior Instructional Designer / Support / Marketing & Communication	€ 92,50
Senior Instructional Designer / Consultant / Project Manager / IT	€ 110,00

Note: Elevate is allowed to adapt her hourly rates each year (from January 1st) based upon the consumer price index, which is determined by the CBS yearly.

Appendix 8B: Proposal for Design Phase of Navigation Tool - Proposal Design phase

To: Dr. Celia W.G. van Gelder
Programme Manager DTL Learning
Postbus 8500, 3503 RM Utrecht
e-mail: celia.van.gelder@dtls.nl

Subject: Design phase FAIR Data stewardship Knowledge,
Skills & Abilities Tool

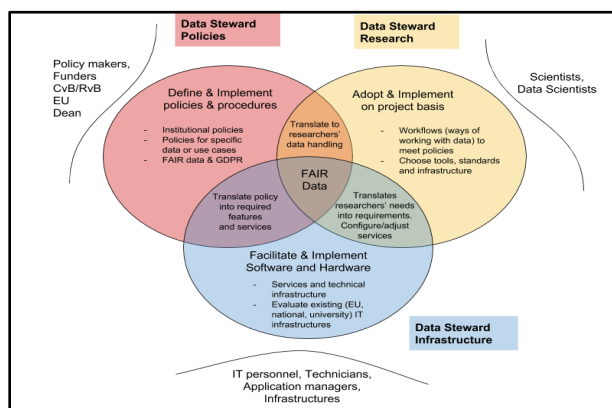
Date: September 12, 2019



Introduction

Good data stewardship is rapidly becoming an essential part of modern science. To facilitate good data stewardship and to promote open science, a broad community of international stakeholders have developed the FAIR Data principles. The Dutch Techcentre for Life Sciences (DTL) actively promotes FAIR data stewardship of life science information. In 2018, Elevate and DTL joined forces and set up a partnership, because DTL's network goes very well with the activities and ambitions of Elevate.

In the context of the ZonMW-project entitled '**FAIR Data Steward as a profession in Life Sciences in the Netherlands**' Elevate has led an exploratory advisory process in preparation of the establishment of an education line and training material for data stewards. This advisory trajectory has taken place between January and July 2019 and has resulted in a 'matrix type' blueprint that describes the required and recommended **Knowledge, Skills and Abilities** for data stewards in life sciences, further converted by the project team by formulating specific learning objectives. With this blueprint, the project team aims



to increase the data steward capacity in the institutes in terms of quality and quantity.

Fig.1 Data steward function landscape and other stakeholders



Request

With the overall objective to professionalize the data steward function within the life-sciences domain, the project team wishes to develop an interactive tool to encompass all the gathered information on Knowledge, Skills and Abilities corresponding to the different levels of data steward professionals. Users of the tool would be both sitting data stewards and new-to-be data stewards in the Netherlands, spread over three levels as visualized in figure 1.

By using this tool, the data steward will be able to:

1. Explore the knowledge, skills and abilities that are connected to their function
2. Self-assess their existing knowledge, skills and abilities and identify possible gaps

- Find resources to bridge these gaps (e.g. websites, blogs, articles, existing course materials)

Already known **quality criteria** of the requested solution are listed below, although some of them have to be elaborated during the design phase. These criteria have been marked with an *.

	The tool should be:
Accessibility	<ul style="list-style-type: none"> - online accessible - easy to use and navigate - self-instructing
User orientation	<ul style="list-style-type: none"> - suitable for both formal and informal learning (on the job & just-in-time)* - suitable for different levels (junior to senior) and functions within the data steward field*
Technical interoperability	<ul style="list-style-type: none"> - browser or platform based* - platform independent
Visual & Instructional Design	<ul style="list-style-type: none"> - based on a clear instructional design - using explicit learning objectives* - offering self-assessment options - visually appealing
Content	<ul style="list-style-type: none"> - following the set-up of the function description overview (8 competence areas)* - offering mapping of (training) resources



Design phase

Based on her many years of experience in solving issues in the field of learning & development, Elevate has developed a clear vision on **continuous education**. In response to the request of the ZonMW-project Elevate proposes to start with the design phase as the first step. The three main objectives of the design phase are as follows:

- Further refine the already formulated learning objectives.
- Decide on the right design structure and architecture of the envisaged tool.
- Define the self-assessment part of the tool.



Division of tasks & responsibilities

Elevate	Project Team
1) Review and revise learning objectives and advice on how to make them sufficiently explicit.	1) Deliver first set of learning objectives. Process the comments and finalise learning objectives.
<p>2) Two co-design sessions led by Elevate in which jointly will be decided upon the right design structure and architecture of the envisaged tool. To give an example, the following questions (among others) will be tackled in these sessions:</p> <ul style="list-style-type: none"> - Layering of main- and subtopics: of how many segments should the design be composed? - How many assessment components are required per category? - How many pages for resources have to be built in? - How many advisory pages have to be built in? - How many and which cross-connections are there between the categories? - What didactical options are there for the various types of user orientation (on-the-job vs just-in-time learning)? - What are the exact quality criteria in terms of technical interoperability? - Which content experts are responsible for the various topics in the follow-up process? <p>These are just examples of questions that will be addressed during a co-design session. Elevate will take the relevant project team by the hand and advise.</p> <p>Each co-design session will be followed by a 'working phase' of the project team, for following-up and to further elaborate on the different outcomes of the sessions.</p>	
3) Draft a proposal for a set of self-assessment components in a suitable template.	3) Review proposal and complete to a final version.
Deliverable: Finalize the complete design structure in online template and submit for approval.	Approval of final design structure in template.



Time frame & investment

For the design phase Elevate proposes a time frame between November 2019 and August 2020, taking into account a predetermined schedule as shown in the following overview:

Activities	Investment Elevate (Hours)	Period
Kick-off & preparation	8	November 2019
Active phase incl. co-design session 1	12	Jan 2020
<i>Work phase project team</i>	-	Jan - March 2020
Active phase incl. co-design session 2	12	March 2020
<i>Work phase project team</i>	-	Apr 2020 - Jun 2020
Finalizing & delivery final design structure	8	July 2020
Total number of hours	40	

Please note that the exact schedule will be discussed during the kick-off phase.

Elevate will take no active role during the two working phases of the project team. According to the budget the main activities of Elevate are strictly limited to the kick-off, both active phases and the final stage.



Indicative investment

For the activities in the above described design phase Elevate has estimated **40 hours à € 110,00**. The total investment, including 10% contingency costs and project management, will be: **€ 4.840 excluding VAT**. For Elevate's hourly rates we refer to Annex I.



Project team & management

Upon agreement on this proposal Elevate will appoint a senior instructional project leader. He or she will take care of guiding the project team during the design phase. The project leader will also be your contact person during the project and will monitor planning and budget.

To guarantee smooth collaboration it is important that the ZonMW project team assigns a project leader who is allowed to give approval upon the deliverable(s) as well.

Annex – I Hourly rates partners 2019

Junior Instructional Designer / Developer / E-moderator / Operations	€ 75,00
Medior Instructional Designer / Support / Marketing & Communication	€ 92,50
Senior Instructional Designer / Consultant / Project Manager / IT	€ 110,00

Note: Elevate is allowed to adapt her hourly rates each year (from January 1st) based upon the consumer price index, which is determined by the CBS yearly.

Appendix 9 - Outreach activities

Upcoming

- Lightning talk at IDCC conference, Dublin, Ireland (February 2020)
- Presentation on ELIXIR-Belgium RDM event (December 2019)
- Talk at CODATA FAIR workshop, Helsinki, Finland (October 2019)
- Poster RDA plenary, Helsinki, Finland (October 2019)
- Birds-of-Feather session RDA plenary, Helsinki, Finland (October 2019)
- Blog Open Aire (invitation; to be written)

Past

- Presentation at Needs for Data Stewards in Denmark, Copenhagen, Denmark (October 2019)
- Presentation Open Science Fair, Porto, Portugal (Sept 2019)
- Presentation at Education Session at ISMB2019, Basel, Switzerland (July 2019)
- Presentation during RD-NL Essentials for Data Support Course, the Netherlands (June 2019)
- Presentation Liber Conference, Dublin, Ireland (June 2019)
- Presentation Liber RDM workshop, Dublin, Ireland (June 2019)
- Poster at ELIXIR All Hands Meeting Lisbon, Portugal (June 2019)
- Presentation at DTL board meeting, Utrecht, the Netherlands (June 2019)
- Presentation ZonMw congress Personalised Medicine, Utrecht, the Netherlands (June 2019)
- Presentation at DTL Partner Advisory Meeting, Utrecht, the Netherlands (May 2019)
- Poster at BioSB Conference, Lunteren, the Netherlands (April 2019)
- Presentation at LCRDM networking day, Utrecht, the Netherlands (April 2019)
- Poster + pitch at Health-RI meeting, Utrecht, the Netherlands (January 2019)
- Regular sharing with Dutch Data Steward Interest Group, the Netherlands
- Regular sharing via LCDRM network, the Netherlands

All the [presentations and posters](#) are deposited in our Zenodo Community (<https://doi.org/10.5281/zenodo.3466720>, at <https://zenodo.org/communities/nl-ds-pd-ls/>)