



# Professionalising data stewardship in the Netherlands: competences, training and education

*Dutch roadmap towards national implementation of  
FAIR data stewardship*

NPOS 2021, end report of the NPOS-F project team “Professionalising data stewardship”, part of the NPOS FAIR data programme line



## Colophon

This version [1.1] of the report has been formally accepted by the NPOS Steering Committee [February 11, 2021]

Copyright © 2021 National Programme Open Science / programme line FAIR data

### Authors:

Mijke Jetten (Dutch Techcentre for Lifesciences, DTL)

Marjan Grootveld (Data Archiving and Networked Services, DANS)

Annemie Mordant (Maastricht University)

Mascha Jansen (GO FAIR Foundation, GFF)

Margreet Bloemers (ZonMw)

Margriet Miedema (National Coordination Point Research Data Management, LCRDM)

Celia van Gelder (Dutch Techcentre for Lifesciences, DTL)

Title: Professionalising data stewardship in the Netherlands: competences, training and education. Dutch Roadmap towards national implementation of FAIR data stewardship

Reference: Mijke Jetten, Marjan Grootveld, Annemie Mordant, Mascha Jansen, Margreet Bloemers, Margriet Miedema, & Celia W.G. van Gelder. (2021). Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship. Zenodo.

<https://doi.org/10.5281/zenodo.4320504>

Use, reproduction and distribution of the content of this report is permitted if provided with reference to this report.

## Content

Foreword by Karel Luyben, National Coordinator of Open Science and President of the EOOSC Association	4
Preamble: The urgency of a coherent approach towards professionalising data stewardship	6
Executive summary	9
Chapter 1: Introduction and background. Towards data stewardship as a profession	12
1.1 The urgency of capacity building in data stewardship	12
1.2 Challenges for professional data stewardship	15
1.3 The National Programme Open Science (NPOS)	15
1.4 NPOS-F: professionalising data stewardship. Competences, training and education	17
Chapter 2: The data stewardship landscape	20
2.1 Data stewardship and data stewards	20
2.2 Differences between training and education	23
2.3 Data stewardship competences and skills	24
2.4 Stakeholders in data stewardship	28
Chapter 3: Landscape analysis of data stewardship education and training	31
3.1 Case studies as first fact-check	31
3.2 Challenges emerging from the landscape analysis	33
3.3 Recommendations: lessons learned from the case studies	33
Chapter 4: Job profiles for data stewards and research software engineers	38
4.1 Emerging field of expertise	38
4.2 Challenges for job profiling	39
4.3 Basic job profile components of a data steward	40
4.4 Recommendations for job profiling	42
Chapter 5: Data stewardship training, education and training certification	44
5.1 Identification and findability of (certified) training and education is key	44
5.2 Challenges for training, education and certification	45
5.3 Annotating and certifying training and education	45
NPOS (2021) Professionalising data stewardship in the Netherlands	2

5.4 Recommendations	47
Chapter 6: Design of a data steward skills tool	49
6.1 Connecting job profiles, competences, and training	49
6.2 Challenges for a data steward skills tool	50
6.3 Basic features of a data steward skills tools	51
6.4 Recommendations for a data steward skills tool	53
Chapter 7: Towards implementing FAIR data stewardship in the Netherlands	56
7.1 The role of data stewardship in the transition to FAIR data	56
7.2 Summary of the recommendations	57
7.3 NPOS-F recommendations for specific stakeholders	59
Annex 1: Glossary and abbreviations	62
Annex 2: Case studies	67
Annex 3: Domain areas, responsibilities and tasks of a data steward	95
Annex 4: Components of the research software engineer job profile	104
Annex 5: Components of the data steward job profile	109
Annex 6: Good practices	122
Annex 7: UAS data steward job profile	132
Annex 8: RDM trainings	148
Annex 9: Annotating trainings	152
Annex 10: Certification of data stewardship	163
Annex 11: Existing tool examples (Competency Hub, CILIP PKSB)	166
Annex 12: Data steward personas and pathways	172
Annex 13: Other features of a data steward skills tool	182
Annex 14: Composition of project team	187

## Foreword by Karel Luyben, National Coordinator of Open Science and President of the EOSC Association



The Dutch National Programme on Open Science has become mature in the last five years. This coincides with the development towards a European Open Science Cloud (EOSC) that aims at creating a “Web of FAIR Data and Related Services”. For these developments to materialise, the contribution of data stewards is essential. Data Stewardship contributions will need to range from those embedded in the active research, directed to the careful handling and curation of data, to those that are much more strategic and are directed towards policies, culture change and/or infrastructures needed for developing such a web of FAIR data.

In the recently developed Strategic Research and Innovation Agenda (SRIA) of the EOSC Association one of the recommendations, by the Working Group on Skills and Training, is to establish so-called Data Stewardship Competence Centres (DSCCs). These DSCCs are foreseen at the local and at the national level to facilitate cross-disciplinary and cross-national networks of experts, leveraging the existing (inter)national approaches in the field.

The recommendations in this report are very much aligned with those at the European level. What is needed, is a coordinated (local - national - European) approach to: formalise the positions of data stewards and to build the capacity of data steward support primarily in the research performing organisations, but also at the research funding, the policy making and at the service providing organisations. This coordinated approach should entail exchange of experience and further alignment on the roles, competences and positions of data stewards. Furthermore, standardisation in training and education, and thus curricula and certifications, will be beneficial to all. The monitoring and evaluation of these developments will have to take place at the local, national and European level.

This report clearly describes the present status in the Netherlands with respect to the issues relevant for professionalising data stewardship. It gives concise and relevant recommendations for the various stakeholders responsible for the development of the data stewardship profession. It helps in defining data stewardship, gives insight in the education and training needed and it provides a framework for planning the further implementation. While this report focuses on the Dutch situation, its findings are likely to apply to other countries as well and therefore to contribute to professionalising data stewardship across Europe.



So, let us not hesitate or ask for new studies and reports, but let us act and implement locally in the research performing organisations and jointly with the service providing and research funding organisations organise the national landscape. While doing so we can monitor, check and revise where necessary.

I firmly believe that we will get far if we implement together and link successfully to the European Open Science Cloud. In my role as National Coordinator of Open Science, I have the privilege of working with a community that is as vibrant as it is diverse. Looking back on what we have achieved in the Netherlands so far, it makes me proud to be part of this Open Science community and I am confident in our ability to tackle further challenges and grasp opportunities by working together.

Many thanks to the team of the NPOS-project F for delivering this excellent result.

## Preamble: The urgency of a coherent approach towards professionalising data stewardship

“Invest 5% of research funds in ensuring data are reusable. It is irresponsible to support research but not data stewardship”, said Barend Mons recently in a Nature article<sup>1</sup>. “Students in PhD programmes spend up to 80% of their time on ‘data munging’, fixing formatting and minor mistakes to make data suitable for analysis — wasting time and talent. With 400 such students, that would amount to a monetary waste equivalent to the salaries of 200 full-time employees, at minimum. So, hiring 20 professional data stewards to cut time lost to data wrangling would boost effective research capacity. Many top universities are starting to see that the costs of not sharing data are significant and greater than the associated risks. Data stewardship offers excellent returns on investment.”

This quote and underlying estimations<sup>2</sup> should be read as a call for action in the current transition of science to Open Science. For Open Science to become the norm, research-performing organisations, funding organisations and policy makers should ensure that they can count on sufficient expertise on data stewardship. Professionalising data stewardship is therefore a crucial step in the transition to Open Science.

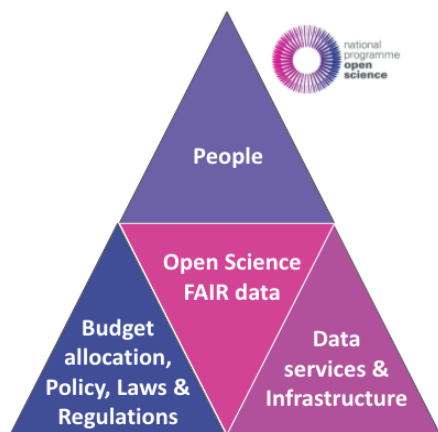


Figure Preamble 1. Dimensions of the route to Open Science

The route to Open Science needs to address three dimensions. Firstly, in terms of policy, regulations, and allocation of budgets. Secondly, research infrastructures and services. And

<sup>1</sup> Mons, B. (2020) Invest 5% of research funds in ensuring data are reusable. Nature, Vol. 578, p. 491 <https://doi.org/10.1038/d41586-020-00505-7>

<sup>2</sup> European Commission High Level Expert Group on the European Open Science Cloud (2016). [https://ec.europa.eu/research/openscience/pdf/realising\\_the\\_european\\_open\\_science\\_cloud\\_2016.pdf](https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf)

finally, cultural change within the research and research support community. For each of these dimensions, various projects in the Dutch National Programme Open Science (NPOS; see also Section 1.3) provide input for the changes that need to take place<sup>3</sup>.

The NPOS-F project, for which the results are described in the current report, addresses the dimension of the people side of Open Science, and focuses on professionalising data stewardship. The results of the project provide arguments for urgent decisions and activities to ensure that the Netherlands can count on adequate capacity of expertise to realise its ambitions with respect to Open Science. The NPOS-F project involves many Open Science themes outside the obvious tasks of supporting researchers in FAIR data stewardship, e.g., policy issues, the position of academic and non-academic professionals in science, and recognition of and reward for the efforts both data stewards and researchers make with respect to data stewardship.

As exemplified by the above quote, the most obvious way to express data stewardship capacity is by **the number of data stewards** who are - or should be - employed at research institutes. The estimated numbers are dazzling. Realising such large numbers has a strong impact on national funds for science and institutional budgets, and it requires changes in the research-performing organisation, including HR management. Also, policy-making organisations and research-funding organisations need to be involved in this change. To allow all stakeholders to make well-informed choices, they need to take aspects into account that go beyond calculating the number of data stewards relative to the number of researchers.

First, research-performing organisations, policy-making organisations, and research-funding organisations need a thorough and well-established understanding of what a data steward actually is. This report therefore provides an analysis of the current situation of data stewardship and what is needed for Open Science and FAIR data. It includes an overview of work domains, competences, roles and tasks that form the basis for recommendations - on job profiles, training and education - that contribute to professionalising data stewardship in all research institutions.

Second, the number of data stewards and their competences depends on the context in which they operate. What is the type of research and data intensity involved? What data services and data infrastructure are available? How well is an organisation prepared to apply Open Science practices? What are the knowledge and skills of researchers with respect to data stewardship and how can data stewards optimally support them? From that perspective, there is still much to be gained. Even a researcher's awareness of Open and FAIR data and software is not yet

---

<sup>3</sup> [www.openscience.nl](http://www.openscience.nl)



*obvious: the FAIR principles are relatively unknown to the research community with 52% of respondents who are frequent data-sharers never having heard of them<sup>4</sup>.*

Third, professionalising data stewardship is inseparable from the position of the data steward within institutions. The NPOS-F project recognises that, apart from gaps in expertise, many research institutes still struggle with the function of data stewards in relation to the researcher's tasks and responsibilities for data. In other words, what choices does a research institute make when composing its research groups? Some argue that the same research output can be obtained with less researchers when they are supported by data stewards. With the broad array of competences provided by this report, it will become possible to take strategic decisions for the composition of a research group, considering research-oriented data stewards as well as data-oriented researchers.

Data stewardship is a new profession. To ensure a data steward's position within a research institute, the profession should become part of HR planning of the organisation and include a vision on career development. In addition, research institutes should collaborate with policy makers and research-funding organisations to realise a common approach to recognise and reward data stewards as full members in research groups, and to make sufficient budgets available to maintain data steward positions in the long run.

The interplay between these three aspects determines whether researchers in a particular research setting are sufficiently equipped and supported to be able to perform data stewardship in the context of Open Science. By having a clear view on each of these aspects, it will be finally possible to answer the question that each research-performing institute should ask itself: **how many data stewards do we need where in the organisation with which competences?**

We are convinced that the recommendations of the NPOS-F project will help stakeholders in the field of data stewardship to take the necessary actions to ensure adequate data steward capacity in the Netherlands to realise its ambitions in Open Science.

---

<sup>4</sup> Digital Science, Fane, B., Ayris, P., Hahnel, M., Hrynaszkiewicz, I., Baynes, G., et al. (2019). The State of Open Data Report 2019. Digital Science. Report. <https://doi.org/10.6084/m9.figshare.9980783.v2>

## Executive summary

This report was commissioned by ZonMw on behalf of the National Programme Open Science (NPOS) to provide arguments for urgent decisions and activities to ensure adequate data stewardship capacity in the Netherlands to realise its ambitions with respect to Open Science. The NPOS-F project, for which the results are described in the current report, focuses on professionalising data stewardship. The report recommends implementation steps that should lead to optimal data stewardship capacity.

In spring 2019, the NPOS steering committee agreed to carry out project F: *Professionalising data stewardship: competences, training and education*<sup>5</sup>. It is one out of ten projects that are aimed at accelerating the implementation of Open Science. In 2020, NPOS defined three key areas<sup>6</sup>: 100% Open Access publishing, optimal reuse of research data, and corresponding evaluation and valuation systems. Project F is associated with the second key area, together with project E on exploring the Dutch data landscape<sup>7</sup>.

The report gives an analysis of the current situation in the Netherlands with regard to data stewardship competences, education and training, and draws attention to the urgent need for a nationally coordinated action by the main stakeholders on the following aspects:

- Defining data stewardship and research software engineering competences as well as formalising the corresponding job profiles via national job classification systems.
- Defining, developing and delivering tailored training programmes to match these required competences.
- Building a data steward skills tool, as a single point of reference for up-to-date information on agreed competences, job profiles, and training opportunities, and allowing for (self-)assessment and identification of career development options.

The recommendations presented in this report tackle the challenges and needs that are experienced at present, both at local and at national level, and thus meet the ambitions that are anticipated in the transition to Open Science. Although this report focuses on data stewardship, research software is increasingly important when it comes to data re-use and research reproducibility. Therefore, also the area of research software engineering, as another activity in the realm of digital and Open Science, is incidentally touched upon in this report. This is reflected in the part focusing on job profiles in particular, which also includes the area of

---

<sup>5</sup> <https://www.openscience.nl/en/projects/project-f-professionalising-data-stewardship-competences-training-and-education>

<sup>6</sup> <https://www.openscience.nl/en/national-platform-open-science/the-key-areas>

<sup>7</sup> <https://www.openscience.nl/en/projects/project-e-exploring-the-dutch-data-landscape>

research software engineering. However, both areas differ substantially in many ways, and therefore research software engineering in its full dimension is not a part of the scope of this report. Furthermore, the report takes into account relevant Open Science issues such as recognition and reward of data professionals. The recommendations target not only developers and providers of training and education, but also developers of institutional job classification systems and Human Resource managers.

Based on the analysis of the current situation, it is recommended that local organisations and umbrella organisations VSNU, NFU and VH together:

- Formalise, adopt and implement the basic components of job profiles for data stewards and research software engineers, both nationally - in job classification systems - and locally: in HR departments of HEIs and RPOs.
- Recognise and reward data stewards and research software engineers and secure their position in close proximity to other stakeholders in research (i.e., PhD candidates, researchers and teachers), including professional development and adequate remuneration.
- Create an open, continuous learning-on-the-job culture for data stewards that also includes soft skills, as well as networking and practice exchange with peers nationally and locally. To formalise the diverse, existing training efforts, in time, develop a - certified - data steward educational curriculum, as a joint effort of local and umbrella organisations.
- Develop a skills tool for data stewards and organisations to assess responsibilities, tasks and competences, combined with navigation to training and training materials.

The chapters of the report give more detailed recommendations for each of these topics, including practical suggestions for short-term implementation by the various stakeholders in the field of data stewardship.

Chapter 1 opens with a variety of international developments regarding data stewardship and then introduces NPOS and in particular this NPOS-F project. In Chapter 2 we provide working definitions of data steward, data stewardship, training and education and we summarise the findings of two earlier Dutch projects about data stewardship competences and skills. The next four chapters address the four angles from which we approached the project: while Chapter 3 analyses the landscape of data stewardship training and education, based on case studies, Chapter 4 discusses the components of a job profile for data stewards and additionally for research software engineers, because their work areas and required competences are related. Chapter 5 takes a closer look at existing data-related trainings and at certification of education and training. Next, Chapter 6 introduces a data steward skills tool as a single point of reference



for information on competences, job profiles, and training opportunities. Whereas each of the Chapters 3 through 6 contains recommendations, the final Chapter 7 addresses higher-level recommendations to specific stakeholders.

The NPOS-F project team consists of over thirty representatives of universities, university medical centres, universities of applied sciences, and service providers, complemented by representatives of the major stakeholders VSNU, VH, NFU, PNN, SURF and ZonMw. Thanks to active involvement of so many partners and the practical applicability of the recommendations, we are convinced that this will result in the necessary decisions and activities to ensure adequate data steward capacity in the Netherlands.

## Chapter 1: Introduction and background. Towards data stewardship as a profession

This chapter addresses the following topics:

- National and international initiatives around Open Science, data stewardship and research software engineering (Section 1.1)
- Challenges for professional data stewardship (Section 1.2)
- NPOS and the NPOS-F project in particular (Sections 1.3 and 1.4)

### 1.1 The urgency of capacity building in data stewardship

Data stewardship<sup>8</sup> and the skills to manage digital data and software are essential in research<sup>9,10,11,12,13</sup>. The urgent call for data sharing in the recent Covid-19 pandemic is only one example that illustrates the relevance of Open Science, research data management, FAIR data and research software<sup>14</sup>. Data stewards and research software engineers play a crucial role in stimulating academia and supporting researchers in the development towards FAIR data and open and sustainable software. In the past years, it has become clear that there is a large need for, and shortage of, individuals with Open Science and data stewardship expertise within research organisations. This is felt in all research domains and also transcends the institutional

---

<sup>8</sup> Data stewardship can be defined as 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. <https://www.lcrdm.nl/en/glossary>, based upon <https://www.dtls.nl/fair-data/research-data-management/research-data-management>

<sup>9</sup> European Commission High Level Expert Group on the European Open Science Cloud (2016). [https://ec.europa.eu/research/openscience/pdf/realising\\_the\\_european\\_open\\_science\\_cloud\\_2016.pdf](https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf)

<sup>10</sup> KNAW (2018). Big data in wetenschappelijk onderzoek met gegevens over personen. <https://www.know.nl/nl/actueel/publicaties/big-data-in-wetenschappelijk-onderzoek-met-gegevens-over-personen>

<sup>11</sup> Scholtens, S., Jetten, M., Böhmer, J., Staiger, Ch., Slouwerhof, I., Van der Geest, M. & Van Gelder, C.W.G. (2019, October 3). Final report: 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.3474789>

<sup>12</sup> Verheul, I., Imming, M., Ringersma, J., Mordant, A., Van der Ploeg, J., Pronk, M. (2019). Data Stewardship on the map: A study of tasks and roles in Dutch research institutes <https://doi.org/10.5281/zenodo.2669150>

<sup>13</sup> Mons, B. (2020) Invest 5% of research funds in ensuring data are reusable. Nature, Vol. 578, p. 491 <https://doi.org/10.1038/d41586-020-00505-7>

<sup>14</sup> <https://www.zonmw.nl/en/research-and-results/fair-data-and-data-management/open-science-in-covid-19-research>

and even the national level.

Dutch universities, university medical centres, universities of applied sciences, as well as their umbrella organisations (VSNU<sup>15</sup>, NFU<sup>16</sup> and VH<sup>17</sup>, respectively), stressed the importance of professionalising data stewardship in order to put Open Science and FAIR data into practice. This is recognised in for instance the recent NWO impulse financing for locally based Digital Competence Centres (DCCs)<sup>18</sup> and the Dutch National Platform Open Science (NPOS) exploration of the Dutch data landscape<sup>19</sup>, which recommends consolidating national alignment and coordination in data services, including data stewardship competences and education.

Moreover, Dutch national funders, such as ZonMw<sup>20</sup> and NWO<sup>21</sup>, require FAIR data management, including the appointment of a data steward, for the projects they fund. Dutch communities such as the DTL Data Stewards Interest Group<sup>22</sup> and the LCRDM Pool of RDM Experts<sup>23</sup> attract more and more attention and commitment from data stewards and related data professionals, even beyond the Netherlands. The recently published recommendations for FAIR software<sup>24</sup>, initiated by the Netherlands eScience Center and DANS, are endorsed by a growing number of organisations.

The importance of data stewardship is also reflected in recent international developments. Four large data organisations - CODATA, GO FAIR, the Research Data Alliance and World Data System - joined forces as Data Together to 'optimise the global research data ecosystem'<sup>25</sup>. International communities for data stewardship have arisen, such as the RDA Professionalising Data Stewardship Interest Group<sup>26</sup> and the Virus Outbreak Data Network (VODAN) Data Stewardship Cluster<sup>27</sup>.

---

<sup>15</sup> <https://www.vsnu.nl>

<sup>16</sup> <https://www.nfu.nl>

<sup>17</sup> <https://www.vereniginghogescholen.nl>

<sup>18</sup> <https://www.nwo.nl/en/calls/local-digital-competence-centres>

<sup>19</sup> <https://www.openscience.nl/en/projects/project-e-exploring-the-dutch-data-landscape>

<sup>20</sup> <https://www.zonmw.nl/en/research-and-results/fair-data-and-data-management/data-management-in-your-project>

<sup>21</sup> <https://www.nwo.nl/en/policies/open+science/data+management>

<sup>22</sup> <https://www.dtls.nl/community/interest-groups/data-stewards-interest-group>

<sup>23</sup> <https://www.lcrdm.nl/en/about-lcrdm/pool-of-experts>

<sup>24</sup> <https://fair-software.eu>

<sup>25</sup> CODATA, Research Data Alliance, World Data Systems, GO FAIR (2020). [https://www.go-fair.org/wp-content/uploads/2020/03/Data-Together\\_March-2020.pdf](https://www.go-fair.org/wp-content/uploads/2020/03/Data-Together_March-2020.pdf)

<sup>26</sup> <https://rd-alliance.org/groups/professionalising-data-stewardship-ig>

<sup>27</sup> <https://www.go-fair.org/implementation-networks/overview/vodan>

In the context of establishing the European Open Science Cloud (EOSC) several recommendations touch on capacity building in data stewardship. The *Turning FAIR into Reality*<sup>28</sup> report (2018) included two specific recommendations. Recommendations 10 and 11 respectively are “Professionalise data science and data stewardship roles and train researchers” and “Implement curriculum frameworks and training”. And its follow-up report, *Six Recommendations for Implementation of FAIR Practice*<sup>29</sup> (2020) from the EOSC FAIR Working Group, explicitly recommends funding awareness-raising, training, education and community-specific support. Moreover, in European Commission funded projects and initiatives - such as the EOSC Skills & Training Working Group<sup>30</sup> and FAIR4S<sup>31</sup> - and bottom-up initiatives like terms4FAIRskills<sup>32</sup>, data stewardship plays an important role as well. Recently, the Danish e-Infrastructure Cooperation (DeIC) and the Danish National Forum for research data management (DM Forum) have communicated the urge for national coordination on data stewardship education in Denmark<sup>33</sup>.

The importance of research software is also reflected in several international developments. The Research Software Alliance (ReSA) is specifically focused on recognising software as a fundamental and vital component of research<sup>34</sup>. RDA has a workgroup focused on developing and applying FAIR guiding principles for software<sup>35</sup>. EOSC also has a dedicated Architecture Task Force for Scholarly Infrastructures for Research Software as part of the Architecture Working Group<sup>36</sup>. Additionally, for a number of years now national Research Software Engineer (RSE) networks have emerged in various European countries: UK, the Netherlands, Germany, Belgium and the Nordic countries, as well as in the USA, New Zealand and Australia.

---

<sup>28</sup> European Commission Expert Group on FAIR data (2018). Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR data. <https://doi.org/10.2777/1524>

<sup>29</sup> European Open Science Cloud FAIR Working Group (2020). Six Recommendations for implementation of FAIR practice. <http://doi.org/10.2777/986252>

<sup>30</sup> <https://www.eoscsecretariat.eu/working-groups/skills-training-working-group>

<sup>31</sup> Whyte, A., Leenarts, E., de Vries, J. et al. (2019). Strategy for Sustainable Development of Skills and Capabilities, EOSCpilot D7.5. <https://eoscipilot.eu/content/d75-strategy-sustainable-development-skills-and-capabilities>

<sup>32</sup> <https://terms4fairskills.github.io>

<sup>33</sup> Wildgaard, L., Vlachos, E., Nondal, L., Larsen, A., Svendsen, M. (2020). National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management (DM Forum). <https://doi.org/10.5281/zenodo.3609516>

<sup>34</sup> <https://www.researchsoft.org>

<sup>35</sup> <https://www.rd-alliance.org/groups/fair-4-research-software-fair4rs-wg>

<sup>36</sup> <https://software.ac.uk/news/have-your-say-scholarly-infrastructures-research-software>

## 1.2 Challenges for professional data stewardship

In the Netherlands, vital developments such as defining recruitment procedures and career paths for current and future data stewards and research software engineers are hampered by the lack of consensus on their responsibilities, knowledge and skills. The confusion about their roles hinders developing adequate data steward capacity. Furthermore, higher-education institutions (HEIs) such as the universities, university medical centres (UMCs) and universities of applied sciences each have their own job classification system: UFO, FUWAVAZ, and FUWA-HBO and HAY, respectively. In addition, a coherent approach to dedicated training and education for data stewardship and research software engineers on the national and international level is lacking. All this, combined with insufficient funding for data stewardship at many institutions, frustrates an efficient approach to extending data steward and research software engineer capacity. This in turn complicates efficient data management, which is necessary for Open Science and reusable data. The lack of a clear career path makes it difficult to recruit and retain talented data stewardship professionals.

To achieve the next steps needed for professionalising data stewardship national coordination is essential, resulting in coherent education, training and HR policy. It is important that researchers and data support professionals work together side by side, each executing their own profession. The term “professionalisation” in this report refers to all activities that broaden a data steward’s competences, which comprises knowledge, skills and abilities. For instance, training, refresher courses and further education fall under the generic term professionalisation, but also coaching, intervision, attending a conference or contact with peers.

## 1.3 The National Programme Open Science (NPOS)

In 2019, the National Programme Open Science<sup>37</sup> (NPOS) proposed the current project (NPOS-F) on professionalising data stewardship, with a focus on data stewardship competences, training and education<sup>38</sup>.

The NPOS sets out to accelerate on the road towards Open Science, along three programme lines:

- Open Access: making all research output (articles etc.) accessible for everyone without costs.
- FAIR data: making all research data FAIR: Findable, Accessible, Interoperable and Reusable.

---

<sup>37</sup> <https://www.openscience.nl/en/open-science>

<sup>38</sup> <https://www.openscience.nl/en/node/171>



- Citizen science: promoting the involvement of citizens in science programmes.

In all programme lines of NPOS attention is paid to the fair recognition and rewarding of researchers with respect to their contribution to Open Science. The FAIR data programme line, composed of NPOS-E (see further below) and NPOS-F, is aligned with the international developments to achieve the necessary facilities and other preconditions for the optimal use and reuse of research data in the Netherlands:

- A consistent ecosystem for FAIR research data: practical implementation of FAIR criteria against technical and policy preconditions.
- Sustainable storage of research data for reuse: research data must be stored in a consistent, reliable and sustainable manner.
- Standards to achieve interoperability of data sets across the boundaries of disciplines, where possible already making use of the interoperability between data management systems in organisations.

The NPOS-E project *Exploring and optimising the Dutch research data landscape*<sup>39</sup> collected good practices and explored and developed the necessary improvements in the national data landscape to:

- Create better boundary conditions for the NPOS ambition of the optimum reuse of research data.
- Boost cooperation between data-intensive scientific fields and the resulting societal force for innovation.
- Prepare participation in the European Open Science Cloud (EOSC) at national level.

Also, a number of other NPOS projects directly affect the context for professionalising data stewardship. Project NPOS-A<sup>40</sup> on the *Transition costs for Open Science in the Netherlands* mentions the lack of Open Science expertise as a major challenge. The authors point at the need for more data professionals, incentives and a better recognition and position for non-academics or supporting professionals in research, as well as well-coordinated data services and infrastructure. Project NPOS-G<sup>41</sup> about *Indicators for Open Science* coordinates the further development of a new standard for recognising and rewarding both researchers and supporting professionals. Finally, project NPOS-H<sup>42</sup> on *Accelerating Open Science* aims to inform a wide

<sup>39</sup> <https://www.openscience.nl/en/projects/project-e-exploring-the-dutch-data-landscape>

<sup>40</sup> Van der Voorden, R., Dijso, E. (2019). Transition costs for open science in the Netherlands.

<https://www.openscience.nl/files/openscience/2019-12/Report%20Transition%20Costs%20for%20OS%20in%20the%20Netherlands.pdf>

<sup>41</sup> <https://www.openscience.nl/en/node/174>

<sup>42</sup> <https://www.openscience.nl/en/node/176>

audience inside as well as outside academic institutions. It plays an important role in creating awareness, and thereby contributes to the recognition and appreciation of data professionals.

#### 1.4 NPOS-F: professionalising data stewardship. Competences, training and education

This NPOS-F project comes in a timely fashion, as data stewardship can no longer be ignored in the Dutch research landscape, which aims at supporting primary research and data processes in creating Open and FAIR data. This report presents the challenges, outcomes and recommendations of the project and aims to create a common starting point for the further professionalisation of data stewardship in the Netherlands, with and by relevant stakeholders.

##### Objectives, deliverables and ambitions

The objectives, deliverables and long-term ambitions of the NPOS-F project are visualised in Figure 1.1 below. Building on previous work (described in Section 2.2), NPOS-F is focused on making next steps a) towards a national approach related to data stewardship competences, skills and learning outcomes for training and education in the context of Open Science and b) towards a well-annotated searchable overview of relevant data stewardship training.

To realise those objectives, the project has created five internal deliverables, which form the basis for the current report. First, a scoping exercise was done, the results of which are shown in Chapter 2 of this report. Second, case studies for training and education in Dutch HEIs and RPOs were collected (Chapter 3). Third, competences for data stewards and research software engineers were drafted out and transferred into basic components for job profiles, to be implemented into the job classification systems of the universities, UMCs and UASs (Chapter 4). Fourth, an inventory of data stewardship training resources was made, including a pilot annotation with the identified competences for data stewards (Chapter 5). Lastly, competences and training were converged into a design for a data steward skills tool with competences and training (Chapter 6). The long-term ambitions for the project are focused on endorsement of the proposed framework for competences and training, ideally via the development of the sketched data steward skills tool.

## Professionalising data stewardship: competences, training and education

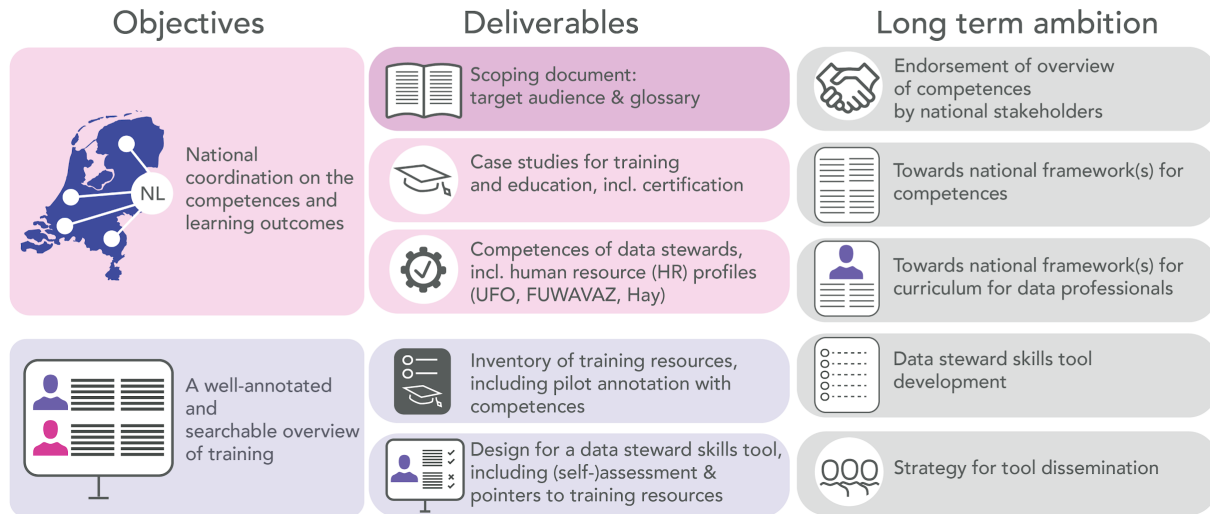


Figure 1.1 NPOS-F project and its objectives, deliverables and ambitions

In the NPOS-F project, we focused on the education levels of the Dutch binary educational system: universities, including university medical centres and universities of applied sciences. Data stewardship is already (partly) recognised and implemented at the university level, including university medical centres, and is quickly becoming so for universities of applied sciences. However, as formal frameworks for data steward profiles and training and education are lacking for both levels of education, there is a need to further professionalise and formalise data stewardship. This need is experienced both bottom-up and top-down and is reflected in the participation of a large and diverse group of representatives of the HEIs and RPOs in the NPOS-F project.

The Dutch umbrella organisations for the different levels of education (VSNU, NFU and VH) are initiators of and partners in the project, as formal members of the NPOS board and as project partners. The long-term ambition of the NPOS members is to make data stewardship competences explicit and develop formal education, embedded in the universities (including university medical centres) and universities of applied sciences. Where relevant, the potential differences between HEIs and RPOs are included in the report, for instance concerning the formal job profiles.

Professionalising data stewardship at the aforementioned levels of education will also stimulate data stewardship at the middle-level applied education, the secondary vocational education and training (represented by the MBO Raad). The latter can profit from the steps already taken in higher-level education, and this level of education can be taken up in future projects.

The NPOS-F project team consists of over thirty representatives of universities, university medical centres, universities of applied sciences, and service providers, complemented by representatives of the major stakeholders VSNU, VH, NFU, PNN, SURF and ZonMw (see Annex 14).

### Recommendations

The output of the NPOS-F project consists of tangible recommendations that will help the aforementioned organisations with professionalising data stewardship, i.e., hire, train and educate data stewards, including clear job profiles and career paths. This way the next step towards professionalising data stewardship can be taken, which will benefit the organisations, but very importantly, also the individual current and future data stewards.

In the various chapters, these recommendations are structured per topic, for the landscape analysis, job profiles, training and education, and for the data steward skills tool.

In summary:

- Recommendations for job profiles for data stewards and research software engineers, both at the level of local organisations and their HR departments, and umbrella organisations (and their job classification systems). This is the topic of Chapter 4.
- Recommendations related to data stewardship training and education in HEIs and RPOs. This is described in Chapter 5 and is based on the Chapter 3 landscape analysis.
- Recommendations for the design and strategy of a data steward skills tool, which integrates - present or aspired - data steward competences (knowledge, skills and abilities or KSAs), appropriate training, and learning outcomes. This is described in Chapter 6.

As we want the various data stewardship stakeholders to take the next steps and implement this report's recommendations, Chapter 7 presents the high-level, overarching recommendations, for each of the stakeholders represented in the NPOS-F project.

## Chapter 2: The data stewardship landscape

This chapter addresses the following topics:

- Working definitions of data steward and data stewardship (Section 2.1)
- The relationship between education and training (Section 2.2)
- Two earlier Dutch studies about data stewardship (Section 2.3)
- The many stakeholders in data stewardship (Section 2.4)

### 2.1 Data stewardship and data stewards

Data stewardship is a catch-all term for numerous support functions, roles and activities with respect to creating, maintaining and using research data. The core responsibilities and tasks vary from policy advising and consultancy, to operational, and technical, ICT-related tasks. They also vary between and among the different HEIs and RPOs.

To illustrate this diffuse situation, we include a list of roles and descriptions. It should be noted that we don't advocate all of them; in this report we only use the description of "data steward". The diffusion is understandable as most descriptions originate from the same evolving landscape of Open Science, research data management and FAIR data, and thus share its newness and fuzziness with that of the data steward.

Role	Description
Data steward	A person responsible for keeping the quality, integrity, and access arrangements of data and metadata in a manner that is consistent with applicable law, institutional policy, and individual permissions. Data stewardship implies professional and careful treatment of data throughout all stages of a research process. A data steward aims at guaranteeing that data is appropriately treated at all stages of the research cycle (i.e., design, collection, processing, analysis, preservation, data sharing and reuse).
Data analyst	This is someone who knows statistics. They may know programming, or they may be an Excel wizard. Either way, they can build models based on low-level data. Most importantly, they know which questions to ask of the data.

Data engineer	Operating at a low level close to the data, they are people who write the code that handles data and moves it around. They may have some machine learning background.
Data manager	A data manager is a person responsible for the management of data objects including metadata. These people think about managing and preserving data. They are information specialists, archivists, librarians and compliance officers.
Data scientist	A practitioner of data science. It is a generic term that encompasses many fields of specialised expertise. In the current report, data analysts, data stewards and research software engineers are considered as sub-groups of data scientists. In certain contexts, data scientist is also sometimes used in a more limited way that makes it equivalent to either the data analyst or software engineer roles.
Research software engineer	A growing number of people in academia combine expertise in programming with an intricate understanding of research. These Research Software Engineers may start off as researchers who spend time developing software to progress their research or they may start off from a more conventional software-development background and be drawn to research by the challenge of using software to further research.
Research support professional	In the context of digitalisation, these are the people who support scientific researchers conducting data-intensive science. They are not necessarily part of a research team and might be considered as service providers. This is a broad category that can include data stewards, RSEs, data managers, librarians and archivists.

Table 2.1 Roles in the data professional landscape<sup>43</sup>

In this report the focus is on the data steward - and partly on the research software engineer, because work areas are related - which is considered a function upon itself with its own competences, responsibilities and tasks, based on specific knowledge, qualifications and

<sup>43</sup> Based on the LCRDM glossary, <https://www.lcrdm.nl/en/glossary>, and on OECD (2020), Building digital workforce capacity and skills for data-intensive science, OECD Science, Technology and Industry Policy Papers, No. 90, OECD Publishing, Paris.

education. Since data stewardship is a newly evolving expertise area there is no universal definition available yet. Several definitions of data stewardship are given in Table 2.2.

Our working definition is the definition of data stewardship that has been put forward by DTL: *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*<sup>44</sup>.

Definition	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 <sup>45</sup>
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 <sup>46</sup>
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 <sup>47</sup>
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 <sup>48</sup>
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 <sup>49</sup>

<sup>44</sup> This report uses "research data management" (RDM) as interchangeable with data stewardship.

<sup>45</sup> <http://www.dtls.nl>

<sup>46</sup> Verheul, I., Imming, M., Ringersma, J., Mordant, A., Van der Ploeg, J., Pronk, M. (2019). Data Stewardship on the map: A study of tasks and roles in Dutch research institutes. LCRDM. <https://doi.org/10.5281/zenodo.2669150>

<sup>47</sup> <https://www.crcpress.com/Data-Stewardship-for-Open-Science-Implementing-FAIR-Principles/Mons/p/book/9780815348184>

<sup>48</sup> European Commission High Level Expert Group on the European Open Science Cloud (2016). [https://ec.europa.eu/research/openscience/pdf/realising\\_the\\_european\\_open\\_science\\_cloud\\_2016.pdf](https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf)

<sup>49</sup> <https://data4lifesciences.nl/hands2/data-stewardship>

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 <sup>50</sup>
The formalisation of accountability for the management of data resources.	Seiner, 2006 <sup>51</sup>
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) <sup>52</sup>

Table 2.2. Definitions of data stewardship<sup>53</sup>

## 2.2 Differences between training and education

Since data stewardship is an emerging profession, this means that both current and future professionals are in need of acquiring data stewardship knowledge, skills and abilities. Chapter 3 and Chapter 5 give insight into the current practices of data stewardship training and education.

Education, on the one hand, refers to formal school curricula at universities, university medical centres, and universities of applied sciences. Typically, these are semester courses, with a strong theoretical basis. Training, on the other hand, refers to point-of-need teaching, is often shorter and tailored to the use of tools and to the implementation of the learning outcomes in daily practice.

A more detailed overview of differences in training and education is given in Table 2.3. Identifying the appropriate format - training or education - for specific individuals typically depends on a number of factors, including the reasons for seeking training, the trainees' background knowledge, time available, and the format of training (including depth, level of difficulty and nature of training required).

<sup>50</sup> [https://smw-rda.esc.rzg.mpg.de/index.php?title=Data\\_Stewardship](https://smw-rda.esc.rzg.mpg.de/index.php?title=Data_Stewardship)

<sup>51</sup> <http://tdan.com/the-data-stewardship-approach-to-data-governance-chapter-1/5037>

<sup>52</sup> EOSCPilot. (2018). EOSCpilot framework of FAIR data stewardship skills for science and scholarship, and draft recommendations on FAIR training.

[https://www.eoscpilot.eu/sites/default/files/fair4s\\_eoscpilot\\_skills\\_framework.pdf](https://www.eoscpilot.eu/sites/default/files/fair4s_eoscpilot_skills_framework.pdf)

<sup>53</sup> Scholtens, S., Jetten, M., Böhmer, J., Staiger, Ch., Slouwerhof, I., Van der Geest, M. & Van Gelder, C.W.G. (2019, October 3). Final report: 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.3474789>



Training	Education
The term training implies the act of imparting a special skill or behaviour to a person, which is commonly offered to employees of operational level.	Education is a process of systematically learning something in an institution.
The training prepares a person for the present job.	Education is a typical system of learning.
Typically short.	Typically long.
Often point-of-need teaching; almost never part of curriculum.	Scheduled in regular curricula e.g., semester courses; learning path in BSc, MSc or PhD programs.
Proof of participation; often not accredited and no certification.	Formally accredited.
Focus on application and direct usability in the working place; many hands-on parts; not so much theory.	Attention for theory; focus on understanding concepts; broader picture.

Table 2.3 Differences between training and education<sup>54</sup>

### 2.3 Data stewardship competences and skills

In 2019, two complementary projects related to data stewardship were conducted in the Netherlands: *Data stewardship on the map* and *Towards FAIR data steward as profession for the life sciences*.

#### “Data stewardship on the map”

Against the backdrop of local and national attention for data stewardship and via survey and interview research, the LCRDM report<sup>55</sup> investigates the provision of data stewardship tasks,

<sup>54</sup> Adapted from Key Differences, <https://keydifferences.com/difference-between-training-and-education.html>

<sup>55</sup> Verheul, I., Imming, M., Ringersma, J., Mordant, A., Van der Ploeg, J., Pronk, M. (2019). Data Stewardship on the map: A study of tasks and roles in Dutch research institutes <https://doi.org/10.5281/zenodo.2669150>, and Verheul, I., Imming, M., Dekker, W. (2019). Infographic to

roles and needs of data stewards in Dutch research institutes. The report has been written by the LCRDM Task Group Data Stewardship and provides an understanding of what Dutch institutes demand of data stewards and also what has been asked for, implemented and developed by and for them. This offers the basis for a clearer job description for data stewardship roles.



Figure 2.4 LCRDM data stewardship task areas

### Findings

- Professionalising data stewardship is high on the agenda. It is clear that data steward functions are no longer a niche area.
- The ambition is to expand the number of data stewards, particularly those with domain specific knowledge.
- Professionalising<sup>56</sup> requires proper recognition of data stewards, career perspectives, suitable training, visibility, a good position in the organisation, focused coordination and an institutional policy.

### Recommendations

- Invest in basic training and education, including recurrent training programs.
- Greater uniformity in function profiles helps to ensure recognition and career perspectives, and makes the role of a data steward more appealing and better recognisable.
- Formulate those profiles not too tight, with a focus on tasks areas, to allow for new task descriptions and tailor-made vacancy texts in a changing field of expertise.

LCRDM Report Data Stewardship on the Map: A Study of Tasks and Roles in Dutch Research Institutes. <https://doi.org/10.5281/zenodo.3066366>

<sup>56</sup> Please note that in the NPOS-F project “professionalisation” has the more limited definition of “all education and training activities that increase a data steward’s competence”.

- Staff should be easily approachable, since strong community building and knowledge exchange are essential, also cross domains. There is enthusiasm for establishing a pool of data stewards, to offer (possibly domain-specific) support.

### “Towards FAIR data stewardship as profession”

Based on an analysis of the data steward landscape and competency frameworks, the ZonMw/ELIXIR funded project focuses on professionalising data stewardship via a function description and pointers to training. Although the project set out to deliver competences for data stewardship in the life sciences, the outcomes<sup>57</sup> are relevant for all research domains.

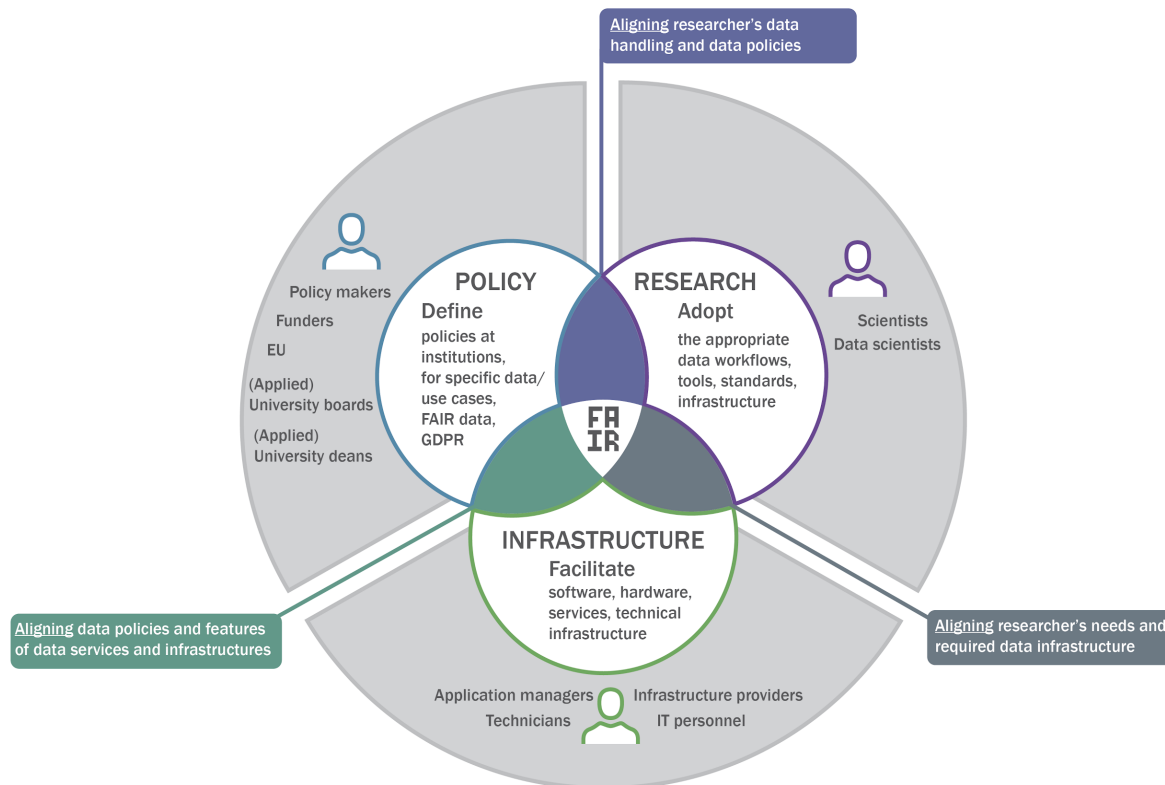


Figure 2.5 ZonMw/ELIXIR data stewardship roles in the data stewardship landscape<sup>58</sup>

<sup>57</sup> Scholtens, S., Jetten, M., Böhmer, J., Staiger, Ch., Slouwerhof, I., Van der Geest, M. & Van Gelder, C.W.G. (2019, October 3). Final report: Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. <http://doi.org/10.5281/zenodo.3474789>

<sup>58</sup> Staiger, C., Jetten, M., Böhmer, J., Slouwerhof, I., Van der Geest, M., Van Gelder, C.W.G., Scholtens, S. (2019). Data stewards function landscape and its stakeholders. <https://doi.org/10.5281/zenodo.3243909>

### *Findings*

- The landscape analysis shows three partly overlapping data steward roles: policy, research and infrastructure (Figure 2.5).
- Eight competence areas for the data stewards can be distinguished: policy/strategy, compliance, alignment with FAIR, services, infrastructure, knowledge management, network, and data archiving.
- Responsibilities and tasks were formulated for each data steward role and competence area, to be used for job descriptions. Knowledge, skills and abilities (KSAs) and learning objectives (including Bloom levels) were described. The resulting three large matrices, one for each of the three identified data stewardship roles, together form the data stewardship competency framework<sup>59</sup>.

### *Recommendations*

- Embed the data steward roles in a formal function profile, including junior and senior levels, and add formal job evaluation.
- Develop a self-assessment tool for data stewards and organisations to assess responsibilities, tasks and competences, combined with navigation to training and 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.

The outcomes and recommendations of both projects are widely recognised and formulate the groundwork on which the current NPOS-F work has been built. The projects have distinguished two types of data stewards: the embedded data steward and the generic data steward. In practice, the division between these two types of data stewards reflects an *organisational* perspective on data stewardship, that is, data stewardship strongly depends on where a data steward is positioned. Moreover, three partly overlapping task areas of the data steward were characterised, each with its specific focus: policy, research and infrastructure. The division between these three data steward task areas reflects an *individual* perspective on data stewardship, i.e., strongly depends on the tasks an individual data steward is given in or by an organisation. Together the types and the task areas form the data stewardship landscape.

In the NPOS-F project, these two complementary views on the data stewardship types and roles have been merged. This becomes for instance visible in Chapter 6 on a data steward skills tool,

---

<sup>59</sup> Scholtens, S., Anbeek, P., Böhmer, J., Brullemans-Spansier, M., Van der Geest, M., Jetten, M. ... Van Gelder, C.W.G. (2019). Function and competencies matrices for three types of data stewards 'policy', 'research' and 'infrastructure' (Version 2.1). Zenodo. <https://doi.org/10.5281/zenodo.3239079>

which introduces five different data steward personas: the embedded and generic data steward (from the organisation perspective), and the policy, research and infrastructure data steward (from an individual perspective).

## 2.4 Stakeholders in data stewardship

In the field of data stewardship expertise many stakeholders can be identified. Two important groups are the researchers and the data stewards. The researchers, who are represented in the NPOS-F project by the PhD candidate network PNN<sup>60</sup>, rely on the expertise and support from data stewards, and need basic training in data stewardship as well. Both researchers and data stewards will benefit from further steps to professionalise data stewardship in the Netherlands - and beyond.

One of the strengths of the NPOS-F project is that in addition to representatives of local research organisations, who bring in the researcher and data steward community perspective, there is involvement of stakeholders who are in a position to implement policies and measures to professionalise data stewardship. These stakeholders can directly influence national and local policymaking and/or implement (local) operational processes in Open Science, research data management and FAIR data. Therefore, they are in the optimal position to endorse and implement this report's recommendations.

These stakeholders, many of which are explicitly addressed in the recommendations in this report, are listed below. Chapter 7 gives an overview of the practical recommendations for short-term implementation by these stakeholders.

- Research-performing organisations and higher-education institutes, i.e., the universities, university medical centres, and universities of applied sciences. This stakeholder group includes executive boards, deans, HR managers, libraries and research/RDM/data support desks.
- National umbrella organisations that represent local research organisations, i.e., NPOS-F project members VSNU (the Association of Universities in the Netherlands)<sup>61</sup>, NFU (the Netherlands Federation of University Medical Centres)<sup>62</sup> and VH (the Netherlands Association of Universities of Applied Sciences)<sup>63</sup>.

---

<sup>60</sup> <https://www.hetpnn.nl/en>

<sup>61</sup> <https://www.vsnu.nl>

<sup>62</sup> <https://www.nfu.nl>

<sup>63</sup> <https://www.vereniginghogescholen.nl>

- National umbrella organisation and commissioner of this report NPOS, the National Open Science Programme in the Netherlands<sup>64</sup>.
- Policy-oriented stakeholders such as research funders (e.g., NPOS-F project member ZonMw<sup>65</sup> and NWO<sup>66</sup>).
- Data services and infrastructure organisations, such as NPOS-F project members SURF<sup>67</sup>, the National Coordination Point Research Data Management (LCRDM)<sup>68</sup>, DANS<sup>69</sup>, the Netherlands eScience Center (NLLeSc)<sup>70</sup> and the Dutch Techcentre for Life Sciences (DTL)<sup>71</sup>.

Of course, as stated above, we also take into account the earlier mentioned groups of researchers and data stewards. In the NPOS-F project we have representatives of six universities, five university medical centres and five universities of applied sciences (see Annex 14). The majority of them are working either in a managerial role or in a data stewardship role. They have proven to be highly relevant to get input on local, practice-oriented processes, to make sure that the project's recommendations match the local experiences, expectations and needs. This is essential for getting endorsement for the project outcomes on a local level as well. Furthermore, researchers and data stewards can be engaged in the implementation via discussions in Dutch community groups like the DTL Data Stewards Interest Group (DSIG)<sup>72</sup>, the above-mentioned networking organisation for Dutch PhD candidates PNN, and the Netherlands Software Engineer Community (RSE-NL)<sup>73</sup>.

Given the broad representation of all stakeholders in the NPOS-F project we are convinced that the recommendations of the NPOS-F project will help an even broader group of stakeholders to take the necessary actions to ensure adequate data steward capacity in the Netherlands to realise its ambitions in Open Science. These include future local and thematic Digital Competence Centres, research-performing institutes with their board members, deans and HR managers which are not represented in NPOS-F, such as institutes of KNAW<sup>74</sup>, NWO<sup>75</sup> and

---

<sup>64</sup> <https://www.openscience.nl/en>

<sup>65</sup> <https://www.zonmw.nl/en>

<sup>66</sup> <https://www.nwo.nl/en>

<sup>67</sup> <https://www.surf.nl/en>

<sup>68</sup> <https://www.lcrdm.nl/en>

<sup>69</sup> <https://dans.knaw.nl/en>

<sup>70</sup> <https://www.esciencecenter.nl>

<sup>71</sup> [www.dtls.nl](http://www.dtls.nl)

<sup>72</sup> <https://www.dtls.nl/community/interest-groups/data-stewards-interest-group>

<sup>73</sup> <https://nl-rse.org>

<sup>74</sup> <https://knaw.nl/en>

<sup>75</sup> <https://www.nwo-i.nl/en>



TO2<sup>76</sup>, other data service providers, other research funding organisations, and their non-academic, private counterparts.

---

<sup>76</sup> <https://www.to2-federatie.nl>

## Chapter 3: Landscape analysis of data stewardship education and training

This chapter addresses the following topics:

- Eight case studies of data steward training and education (Section 3.1)
- Challenges from the landscape of case studies (Section 3.2)
- Lessons drawn from the case studies (Section 3.3)

### 3.1 Case studies as first fact-check

What are the backgrounds of current data stewards and what training have they received? What can this tell us about the needs of data stewards across the Netherlands?

For this chapter, case studies have been collected to gather initial insights into data stewardship training and education in HEIs and RPOs in the Netherlands. The rationale for gathering case studies is that, although there is some information available about training which current or prospective data stewards attend, there is less information on how various organisations currently support the training needs of their data stewards. For example:

- Do institutions have dedicated training plans for their data stewards?
- Is training provided internally or externally, or as a mixture of the two?
- Who delivers such training?
- What kind of topics are covered by the training programme?
- How do institutions address the needs of longer-term educational needs of the data stewards?

The case studies focus on the data stewards, how they are embedded in the organisation, and the training and education they receive. They are a first fact-check on the current landscape of data stewardship training and education. The perspective of for instance the organisations, training providers or training certifiers is discussed in Chapter 5.

To reflect the diversity of various approaches and provide insight into the suitability of different models depending on the types of institutional setting, there are:

- Two case studies from universities: TU Delft, Radboud University
- Three case studies from university medical centres: University Medical Center Utrecht, Erasmus MC, Maastricht UMC+
- Three case studies from universities of applied sciences: Avans University of Applied Sciences, University of Applied Sciences Leiden, HAN University of Applied Sciences



As example, Figure 3.1 visualises an exemplary data steward at TU Delft, whereas Figure 3.2 summarises the local data stewardship situation. All eight case studies can be found in Annex 2.

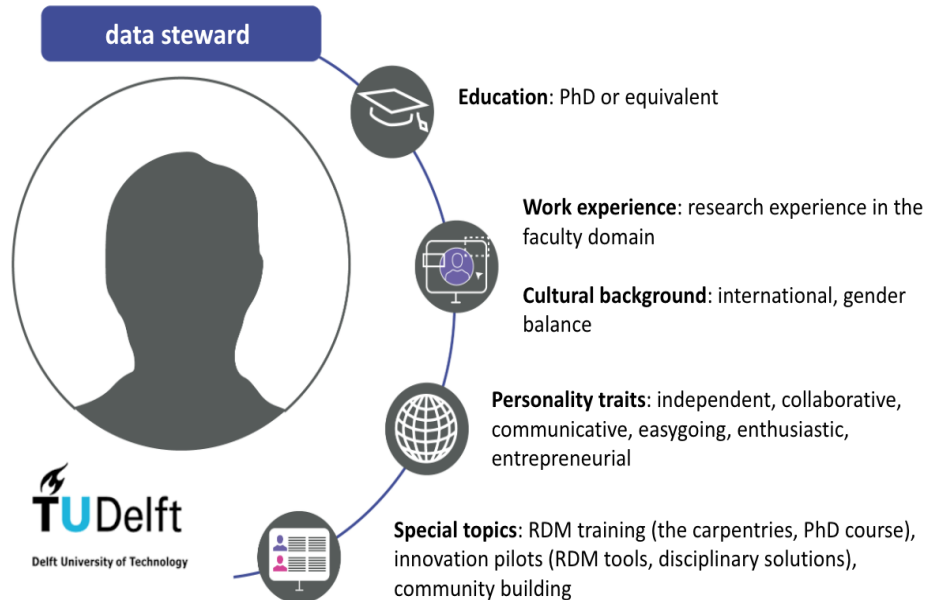


Figure 3.1 Delft University of Technology data steward

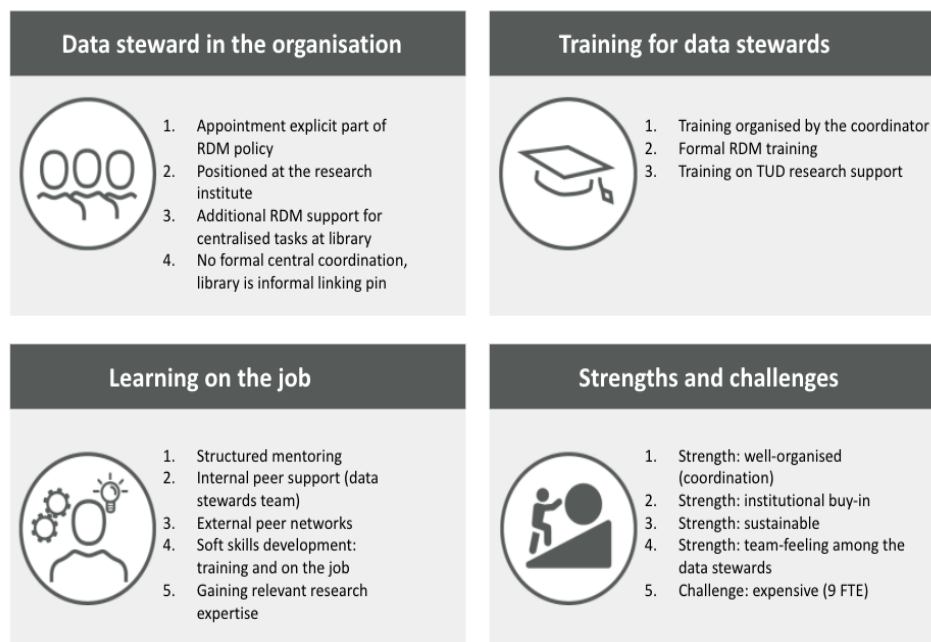


Figure 3.2 Delft University of Technology case study reference card

Together, these case studies form an interesting landscape, which we subsequently analysed with a view to finding challenges as well as lessons that could be drawn; see next sections.

### 3.2 Challenges emerging from the landscape analysis

Contrary to Chapters 4 to 6, the challenges in this chapter were not the basis for the analysis. Instead, the following challenges emerged from the landscape analysis:

Challenge number	Challenge
Chal_Case_1 <i>Formalisation</i>	Lack of formal education and training for data stewards. Also, it is often unclear which skills and personal development are required for data stewardship. <i>This confirms the raison d'être of the current project.</i>
Chal_Case_2 <i>Defining needs</i>	Difficulty of defining training needs of data stewards. Defining them will help trainers and educational experts to develop training programs. <i>This is also addressed in Chapter 6.</i>
Chal_Case_3 <i>Existing training</i>	Lack of insight into existing training for data stewards: what is available, with which content, what knowledge gaps does it address, and how do learners value it? <i>This is also addressed in Chapter 5.</i>
Chal_Case_4 <i>Life-long learning</i>	Lack of support for life-long learning for data stewards.
Chal_Case_5 <i>Strategic vision</i>	A bottom-up implementation of data stewardship, without a strategic vision and corresponding capacity, governance and funding in an organisation, is vulnerable and susceptible to frequently changing priorities.
Chal_Case_6 <i>Awareness</i>	The difficulty of raising awareness about data stewardship among researchers.

Table 3.3 Challenges that emerged from the landscape analysis

### 3.3 Recommendations: lessons learned from the case studies

Both local organisations and umbrella organisations are envisioned to be involved in planning and designing training and education programmes. Based on the challenges, strengths and

lessons learnt from the case studies, the following recommendations are made to the stakeholders:

Recommendations landscape analysis		
<p>Rec_Case_1 Use case studies to plan training</p>	<p>Recommendation Use the case studies to better plan the structure and content of training. To clarify the needs of your organisation's data steward, analyse your local case study. Use the case studies to set up a community survey to further define the needs of data stewardship training and identify gaps, taking into account institutional settings like type of institution, size, research disciplines, available resources.</p>	<p>Stakeholders Local organisations Umbrella organisations</p>
	<p>Reasoning The case studies offer a first look into training needs. Sharing knowledge could help identify shared training elements. In the case studies, for example, Essentials for Data Support was frequently listed as relevant training and could be used as a basis for future training. The survey may provide a complete overview of the Dutch landscape and the (local) training needs.</p>	<p>Challenges addressed Chal_Case_2 Chal_Case_3</p>
<p>Rec_Case_2 Care for your data steward</p>	<p>Recommendation Reflect on the lessons learned from the case studies, particularly on how to take good care of your data steward. Create an open, continuous learning-on-the-job culture, that includes:</p> <ul style="list-style-type: none"> <li>• Soft skills and transferable skills</li> <li>• Networking, collaboration and sharing experiences</li> <li>• Gaining disciplinary research insights</li> <li>• Joining external, (inter)national community activities</li> <li>• Personal and professional development</li> </ul>	<p>Stakeholders Local organisations</p>

	<p>Reasoning</p> <p>A continuous learning environment is pivotal. Data stewards need to continuously develop their skills and competence based on the research context and organisational culture. Social interaction is an important component, just like communication, networking, interviewing and coordination skills. Having skills in understanding RDM challenges, transferring RDM knowledge and collaborating on developing RDM solutions is crucial to do the job well.</p>	<p>Challenges addressed</p> <p>Chal_Case_1 Chal_Case_2 Chal_Case_4</p>
<p>Rec_Case_3</p> <p>Collaborate in training</p>	<p>Recommendation</p> <p>Collaborate with (other) local organisations and umbrella organisations when developing training.</p>	<p>Stakeholders</p> <p>Local organisations Umbrella organisations</p>
	<p>Reasoning</p> <p>It is crucial to share resources and knowledge. Local organisations should define the learning objectives for training for data stewards, based on job profiles of data stewards (Chapter 4). National and other non-local training will be more effective if it meets local training needs. Furthermore, umbrella organisations can help by (co-) funding training and train-the-trainer initiatives, and by connecting people.</p>	<p>Challenges addressed</p> <p>Chal_Case_1 Chal_Case_2 Chal_Case_3</p>
<p>Rec_Case_4</p> <p>Community and networking</p>	<p>Recommendation</p> <p>Encourage the growth of data steward communities and networking activities.</p>	<p>Stakeholders</p> <p>Umbrella organisations</p>
	<p>Reasoning</p> <p>Data stewards need effective networks of peers. Case studies mentioned mentorship as a way to help an organisation and its data stewards to explore the requirements of the job as well as the needs of the trainee itself.</p>	<p>Challenges addressed</p> <p>Chal_Case_2 Chal_Case_4</p>

Rec_Case_5 Coordinated approach	<p>Recommendation</p> <p>Choose a well-organised, coordinated approach to data stewardship, as part of the institutional strategy and accompanied by policies, training and governance.</p>	<p>Stakeholders</p> <p>Local organisations</p> <p>Umbrella organisations</p>
	<p>Reasoning</p> <p>Develop data-related training to support the organisation's strategy or policy. Employing data stewards locally, in proximity to the researcher, and coordinating them centrally in the organisation is essential.</p> <p>Establishing close collaboration in the organisation helps embedding the data stewards and align data management needs of the various stakeholder groups. This includes cooperation with the faculties, with the various support departments, and with executive layers of the organisation. At a national level, a good example are the FAIR data workshops organised by ZonMw, about meeting the ZonMw FAIR requirements in COVID-19 related projects.</p>	<p>Challenges addressed</p> <p>Chal_Case_5</p> <p>Chal_Case_6</p>
Rec_Case_6 Flexibility	<p>Recommendation</p> <p>Flexibility in the job helps to stimulate professional development and offers an opportunity to tailor the job to the needs of the researchers.</p>	<p>Stakeholders</p> <p>Local organisations</p>
	<p>Reasoning</p> <p>Flexibility and a broad range of tasks helps to stimulate passion and enthusiasm of data stewards and contributes to recognition by researchers for the work data stewards do.</p>	<p>Challenges addressed</p> <p>Chal_Case_1</p> <p>Chal_Case_6</p>
Rec_Case_7 Proximity to peers	<p>Recommendation</p> <p>Train and position the data stewards in close proximity to their colleagues and collaborators (PhD candidates, researchers, teachers) helps to address disciplinary needs and differences.</p>	<p>Stakeholders</p> <p>Local organisations</p>

	<p>Reasoning</p> <p>A disciplinary focus helps building relationships with researchers.</p> <p>Data stewards should not work in isolation but need effective networks of domain-specific peers, to integrate with and learn from.</p>	<p>Challenges addressed</p> <p>Chal_Case_6</p>
--	---	--

*Table 3.4 Recommendations based on the landscape analysis*

## Chapter 4: Job profiles for data stewards and research software engineers

This chapter addresses the following topics:

- Data stewardship as an emerging field of expertise (Section 4.1)
- Challenges for job profiling (Section 4.2)
- Basic job profile components (Section 4.3)
- Recommendations for job profiles (Section 4.4)

Because this chapter is based on extensive information in the annexes, we also provide pointers here to the relevant annexes:

- Domain areas, responsibilities and tasks, as well as the competences of a data steward (Annex 3).
- Basic components of the data steward job profile, structured loosely in alignment with the UFO job classification system (Annex 5).
- Three local university data steward job profiles, proposed for the FUWAVAZ job classification system (Annex 6).
- Proposal for a job profile for data stewards in universities of applied sciences (Annex 7)
- Basic components of the research software engineer job profile (Annex 4).

### 4.1 Emerging field of expertise

In this chapter the domain areas, responsibilities and tasks, and the competences of data stewards are sketched. These help HEIs and RPOs to define the data stewardship need in the organisation, create job adverts to recruit new staff, and ease the process of job profiling, remuneration and career development. It is also expected to contribute to uniformity in the data steward function in the context of the existing job evaluation systems. This concerns four groups of professionals:

- Data stewards at universities, who fall under the UFO job classification system.
- Data stewards at university medical centres (UMCs), who fall under the FUWAVAZ job classification system.
- Data stewards at universities of applied sciences (UASs), who fall under the FUWA-HBO and/or HAY job classification system.
- Research software engineers (RSEs) within HEIs and RPOs.

While the main report focuses on professionalising data stewardship, this chapter includes research software engineering as well (although not in the same level of detail). Despite the pronounced differences between the two roles, together, the data steward and research

software engineer are a central part of the above-mentioned Dutch development of establishing local DCCs at universities, UMCs and UASs. This neatly aligns with a recent paper<sup>77</sup> from the RSE-NL community that recommends to “stimulate the formalisation of Research Software Engineer (RSE) and Data Steward roles. These roles should have formal job descriptions and have proper career progression paths”.

## 4.2 Challenges for job profiling

In the development of job profiles for data stewardship and research software engineering functions, the following challenges can be identified:

Challenge number	Challenge
Chal_Job_1 <i>Profiles</i>	Job classification systems lack a data steward and research software engineer profile. Data stewards and research software engineers are positioned in profiles that don't match their responsibilities and tasks. Organisations develop their local data steward and research software engineer profile, which results in a great variety of profiles, between and within organisations. This complicates the recruitment of candidates.
Chal_Job_2 <i>Career tracks</i>	Difficulty of defining clear career tracks for data stewards and research software engineers, including recognition and remuneration for their expertise. This complicates equipping staff with specialised competences, and makes the career prospects for these jobs unattractive, potentially leading recruits away to industry.
Chal_Job_3 <i>Position</i>	In the organisation, lack of attention for the distinctive position of a locally embedded data steward and research software engineer at the faculty or department level, and a centrally positioned generic data steward and research software engineer.

<sup>77</sup> Akhmerov, A., Cruz, M., Drost, N., Hof, C., Knapen, T., Kuzak, M., Martinez-Ortiz, C., Turkyilmaz-van der Velden, Y., van Werkhoven, B. 2020. Raising the Profile of Research Software: Recommendations for Funding Agencies and Research Institutions in the Netherlands. <https://doi.org/10.5281/zenodo.4015242>



<p>Chal_Job_4 <i>Good practices</i></p>	<p>As data stewardship and research software engineering are new fields of expertise, with a mixture of existing and newly to be recruited professionals with various backgrounds and expertise, it is challenging to capture this in a job profile. Even if good practices and example job profiles exist, most organisations are not aware of these examples, as they are not shared among organisations.</p>
<p>Chal_Job_5 <i>Capacity</i></p>	<p>Lack of capacity in data stewardship and research software engineering, due to the demands for FAIR data and software. Formal job profiles will contribute to professionalising data stewardship and research software engineering and thus facilitate the process of recruitment.</p>

Table 4.1 Challenges for job profiles for data stewardship and research software engineering functions

### 4.3 Basic job profile components of a data steward



Figure 4.2 Basic job profile components of a data steward

We bring together the components of a basic job profile for data stewards in one image, point the reader to the annexes that provide full details on these components.

Figure 4.2 contains the basic job profile components of a data steward, against the background of an organisational context and the three overlapping data steward roles - policy, research and infrastructure - as highlighted in Figure 2.5 (Section 2.3). The figure contains eleven core activities for a data steward, which a job profile should reflect. Therefore, they are called basic job profile components. These components consist of the eight domain areas of a data steward (purple) plus two result areas relevant for a data steward (light pink). In addition, a professional data steward is competent in certain soft skills (pink). We list the components first and describe next where to find more information.

1. *Policy and strategy*: design strategies for raising awareness of RDM policies and regulations
2. *Compliance*: advise on institutional compliance with RDM policies and regulations
3. *Facilitating good RDM practices*: advise relevant stakeholders on good practices of management of research data
4. *RDM services*: propose, implement and monitor RDM workflows and practices
5. *Data infrastructure*: identify the requirements for adequate RDM infrastructure and tools
6. *Knowledge management*: determine the adequate level of RDM knowledge and skills
7. *Network and communication*: create and participate in (inter)national RDM networks
8. *Data sharing and publishing*: analyse gaps in support for data sharing and publishing
9. *Coordination of work*: lead, supervise and support less experienced colleagues
10. *Coaching and process improvement*: make proposals for improving work processes at different levels
11. *Soft skills*: this area comprises activities like accuracy and persuasiveness

For a general understanding of what a data steward is and does, Annex 3 provides an overview of the domain areas, responsibilities and tasks, as well as the competences of a data steward. These areas match components 1-8 in Figure 4.2.

The basic components of the data steward job profile can be found in Annex 5. The structure of Annex 5 is loosely based upon the UFO standard, as a result of exploratory, joint efforts of the project team and the UFO working group to get the data steward integrated into the UFO job classification system. The proposed basic profile can be used by organisations and should ideally be integrated into the job classification systems mentioned before. Clearly, the proposed basic profile in Annex 5 takes the eight domain areas into account (in Table Annex 5.1). However, it identifies three additional areas, which match components 9-11 in Figure 4.2. Components 9 and 10 are part of the description of the four function levels Junior, Medior,

Senior, and Expert data steward (Table Annex 5.4). Finally, component 11 - soft skills - is fleshed out in Table Annex 5.2.

While Figure 4.2 summarises the proposed job components of the data steward, we address the function of a research software engineer in a similar manner: first, the RSE domain areas and responsibilities, and second, the RSE job profile (Annex 4).

#### 4.4 Recommendations for job profiling

Job profiling of data stewards and research software engineers involves both local organisations and umbrella organisations. Based on the challenges and output, the following recommendations are made to the stakeholders:

Recommendations job profiles		
Rec_Job_1 Formalise profiles	Recommendation Formalise a data steward and research software engineer function profile in all job classification systems and stimulate convergence between these systems.	Stakeholders Umbrella organisations
	Reasoning Uniform job profiles lead to efficiency saving in drafting job vacancies, justifying remuneration levels and are a basis for adequate training.	Challenges addressed Chal_Job_1
Rec_Job_2 Adopt profiles	Recommendation Stimulate local organisations to adopt these formal profiles.	Stakeholders Local organisations
	Reasoning Awareness raising is important as HR departments might not see the problem or don't have the solution. It eases the recruitment of skilled data steward and research software engineers. The profiles could be used when creating job descriptions and job vacancies.	Challenges addressed Chal_Job_1

Rec_Job_3 Create career perspectives	Recommendation The job profiles should promote professional development and growth.	Stakeholders Local organisations Umbrella organisations
	Reasoning A clear career perspective diminishes a mismatch between job expectations and skill sets, and limits knowledge drains because of the lack of growth perspective.	Challenges addressed Chal_Job_2
Rec_Job_4 Allow diversity of roles and types	Recommendation Organisations and researchers should think about the sort of data steward needed to fulfil their specific needs.	Stakeholders Local organisations
	Reasoning Each project, faculty, and central or local department may need different roles and types of data stewards and research software engineers.	Challenges addressed Chal_Job_3
Rec_Job_5 Adopt good practices	Recommendation Adopt good practices for job descriptions and vacancies as well as make HR departments aware of them.	Stakeholders Local organisations
	Reasoning The advantage of using good practices is that job descriptions can be adjusted locally as needed.	Challenges addressed Chal_Job_4
Rec_Job_6 Secure positions	Recommendation Recognise, reward and secure the position of data stewards and research software engineers at various levels.	Stakeholders Local organisations Umbrella organisations
	Reasoning Data stewards and research software engineers are new roles which need to be sustainably funded.	Challenges addressed Chal_Job_5

*Table 4.3 Recommendations for job profiles for data stewardship and research software engineering functions*

## Chapter 5: Data stewardship training, education and training certification

This chapter addresses the following topics:

- Identification and findability of training and education (Section 5.1)
- Challenges for training, education and certification (Section 5.2)
- Annotating and certification of training (Section 5.3)
- Recommendations (Section 5.4)

### 5.1 Identification and findability of (certified) training and education is key

In order to professionalise data stewardship, an unambiguous overview of competences is needed. This overview should be endorsed by organisations like VSNU, NFU, VH and NVAO (Dutch-Flemish Accreditation Organisation). Such an overview will be the starting point for a specific data steward curriculum to be taught at universities and university medical centres (WO), universities of applied sciences (HBO), perhaps secondary vocational education (MBO) and by training providing organisations. If one can become a data steward after WO, HBO or MBO education, a distinction between those three education types needs to be apparent in the level of work, and the job classification profile.

Several challenges related to data stewardship training and education were raised in Chapter 3, emerging from the landscape analysis. The current chapter focuses on two aspects of training and education:

- Overviews with relevant training opportunities exist and provide valuable resources for the community<sup>78</sup>. However, such overviews can be hard to find. Key to findability is annotation with proper metadata. For this reason, a pilot study was done to annotate a number of data stewardship courses with the NPOS/ELIXIR competency framework for data stewards.
- Certification of data stewardship training, training providers and/or acquired competences is high on the wish list of many organisations and institutions, as it would give a clear indication of the competences that a data steward can acquire or has acquired. Therefore, we made an inventory of existing certification mechanisms for training and education in the Netherlands, both in general sense and specifically for data stewardship.

---

<sup>78</sup> See for instance the annex in Scholtens, S., Jetten, M., Böhmer, J., Staiger, Ch., Slouwerhof, I., Van der Geest, M. & Van Gelder, C.W.G. (2019, October 3). Final report: 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.3474789>

## 5.2 Challenges for training, education and certification

There are a number of data stewardship challenges that directly relate to data stewardship training and education:

Challenge number	Challenge
Chal_Tr&Ed_1 <i>Findability</i>	Lack of findable, adequate education and training for data stewards. Or where they exist, missing clarity on how they relate to the competence development of data stewards.
Chal_Tr&Ed_2 <i>Competences</i>	Lack of agreement on responsibilities, tasks and competences of data stewards, which complicates developing education and training
Chal_Tr&Ed_3 <i>Coordination</i>	Lack of alignment and coordination among and between local, national and international stakeholders on developing education and training related to data stewardship.
Chal_Tr&Ed_4 <i>Certification</i>	Lack of certification for data stewardship related education and training, training providers and trainers.

Table 5.1 Data stewardship challenges relating to training and education

## 5.3 Annotating and certifying training and education

### Inventory of training resources and pilot annotation of courses

Building on earlier overviews, an inventory of training resources for (aspiring) data stewards was made (Annex 8). This non-exhaustive overview contains thirty trainings, mainly in English. Twenty of them are publicly accessible, whereas the others are provided by Dutch universities to their own researchers and require institutional credentials. However, as the latter don't appear to be organisation-specific, they could become steppingstones for national training. In addition, the overview lists nearly twenty international registries, handbooks and other pointers to data-related training materials. Those materials can be used in training and data stewards should be familiar with them.

Next, to see if and how these trainings match with the required competences of data stewards, a pilot study was done by annotating a selection of seven trainings with the data steward competences that are mentioned in Annex 3. The annotation pilot is described in Annex 9. Furthermore, two courses were explored to map to finer-grained learning objectives.

The pilot annotation yielded the following main insights (more specific findings are reported in Annex 9):

- Annotating trainings with a concise set of competences is both feasible and useful. Since familiarity with the trainings is helpful, annotating is ideally left to training providers. Annotation by means of the finer-grained learning objectives requires more effort.
- Finding a common language to annotate training is essential to tackle challenges that stem from lacking alignment.

### Certification for data stewardship

Sufficient education and training for data stewards is lacking. Therefore, it is first and foremost necessary to gain insight into the competences (i.e., knowledge, skills and abilities) that a data steward must have, and based on that, what the basic components should be of a curriculum for a data steward. This in turn would be the framework on which certification can be based. Certification mechanisms for data stewardship are high on the wish list of many stakeholders as certification implies a quality stamp. Certification for data stewards is needed because it proves and fosters consistency in data stewardship competences, also for employers.

We made an inventory of existing certification mechanisms for training and education in the Netherlands, both in general and specifically for data stewardship (Annex 10). We looked at certification for four categories: courses, trainees, trainers, and organisations,

We observe that there is hardly any certification in place. Only one university and two universities of applied sciences provide data stewardship training as part of their curriculum and one UAS offers certification. Once more data stewardship education will be developed and implemented at the HEIs (more specifically in the MSc and BSc programmes), the regular certification mechanisms that are in place in the HEIs can and will be applied. Related to data stewardship training, the training providers listed are not-for-profit organisations (DANS, RDNL, SURF, DTL, NleSc) and commercial providers (Phortos, the Hyve and Micelio). None of them have a certification mechanism in place but all do hand out a certificate of attendance.

From our inventory it is clear that certification for data stewardship is still in its early days. Certification schemes can only be developed after next steps are taken in the establishment of dedicated data stewardship curricula that are endorsed by all stakeholders. Certification of trainers, organisations and trainees will follow once these dedicated curricula are in place, since these will form the basis on which certification schemes for trainees, trainers and organisations can be developed.

In addition, certification will not be a national effort, but needs to be done in alignment with similar activities in Europe and beyond. For this reason, it is good to stay tuned with developments outside the Netherlands and align next steps in accordance with emerging standards.

#### 5.4 Recommendations

A combination of formal, curriculum-based education and training is envisioned, spanning both universities, university medical centres, and universities of applied science. Training programmes are essential to fill the knowledge gap before a curriculum is in place, and to allow for life-long learning.

The following recommendations concern data stewardship training and education. Please note that some recommendations are addressed to an implementation team - not in place yet -, which is recommended to support the umbrella organisations in implementing the various recommendations.

Recommendations training and education		
Rec_Tr&Ed_1 Standardise metadata for training	<p>Recommendation</p> <p>Develop and maintain an easy-to-use list of tags based on the domain areas (Annex 3). Integrate the tag list in the recommended data stewardship skills tool.</p>	<p>Stakeholders</p> <p>Implementation team</p>
	<p>Reasoning</p> <p>Ideally the annotation tags have unambiguous and persistent definitions, which can be used in automated processes.</p>	<p>Challenges addressed</p> <p>Chal_Tr&amp;Ed_1 Chal_Tr&amp;Ed_3</p>
Rec_Tr&Ed_2 Develop training annotation process	<p>Recommendation</p> <p>Stimulate training providers to tag their data-related courses, starting with the trainings in Annex 8. Develop guidelines for training content developers. Periodically look for gaps in the tagged training provision.</p>	<p>Stakeholders</p> <p>Implementation team</p>
	<p>Reasoning</p> <p>Dynamically tracking training opportunities is time-consuming. It is more efficient if training providers themselves tag their training provision.</p>	<p>Challenges addressed</p> <p>Chal_Tr&amp;Ed_1</p>



	Gaps imply that more training providers should be encouraged to tag their trainings or there is a need for new training.	
Rec_Tr&Ed_3 Create curated resource	Recommendation Turn the inventory of trainings (Annex 8) into a curated community resource (see also the FAIRsFAIR report). Establish a community platform for sharing such resources and for collaboration.	Stakeholders Implementation team
	Reasoning Common guidelines facilitate exchange and reuse of FAIR training materials. A community platform can aggregate open training materials; trainers, their profiles, expertise and certificates; and training events	Challenges addressed Chal_Tr&Ed_1 Chal_Tr&Ed_3
Rec_Tr&Ed_5 Align with international certification initiatives	Recommendation Monitor international certification developments and act on what emerges, e.g., recommend national endorsement of a standard.	Stakeholders Implementation team
	Reasoning Certification of training, training providers, organisations, and/or students is both sensitive - who is the certifying authority? - and of international interest <sup>79</sup> .	Challenges addressed Chal_Tr&Ed_4
Rec_Tr&Ed_6 Identify certification provider	Recommendation Arrange that a neutral organisation holds the certification schema for trainings, with certifying bodies performing the actual certification based on that schema.	Stakeholders Implementation team
	Reasoning E.g., Edustandard <sup>80</sup> could maintain a certification scheme.	Challenges addressed Chal_Tr&Ed_4

Table 5.2 Recommendations relating to training and education

<sup>79</sup> Many EOSC-related organisations work towards certification, e.g., the Community of Practice of training coordinators (<https://www.openaire.eu/cop-training>) and the EOSC Working Group on skills and training (<https://www.eoscsecretariat.eu/working-groups/skills-training-working-group>). Dutch training experts are involved in both.

<sup>80</sup> <https://www.edustandaard.nl/>

## Chapter 6: Design of a data steward skills tool

This chapter addresses the following topics:

- A tool that connects job profiles, competences, and training opportunities (Section 6.1)
- Challenges for a data steward skills tool (Section 6.2)
- Basic features of such a tool (Section 6.3)
- Recommendations for a data steward skills tool (Section 6.4)

### 6.1 Connecting job profiles, competences, and training

Through our research for this project, it has become apparent that many individuals and organisations are asking the same questions:

- How are other organisations defining the role of their data stewards?
- What skills and roles are needed now, and in the future?
- What training is available for data stewards?

As highlighted in previous challenges in this report, everyone is currently relying on serendipity and on national communication networks to share ideas and information and to find answers to their questions, and there is an absence of any (inter)national coordination of knowledge. With the growing population of data stewards, it is becoming increasingly difficult to keep track of all opportunities and examples of best practice.

It seems possible to provide answers to these questions by building a data steward skills tool on a national, web-based platform: a single point of reference for up-to-date information on agreed competences, job profiles, and training opportunities, and allowing for (self)assessment of data stewards and identification of career development options.

In this chapter, we explore two international web-based tools to present competences and training opportunities, along with possibilities for (self)assessment and career planning. Because of the rapid development of data stewardship in the Netherlands, an investment in such a tool would benefit many stakeholders. It would help with the planning for, recruiting, training, and development of data stewards, and would benefit employers of data stewards, potential and current data stewards themselves, and also those organisations involved in developing, offering or funding training.

## 6.2 Challenges for a data steward skills tool

In the development of a data steward skills tool, the following challenge can be identified:

Challenge number	Challenge
Chal_Tool_1 <i>Single point of reference (tool)</i>	<p>A single point of reference (tool) is needed for up-to-date information on competences, profiles, training, and allowing for (self)assessment and identification of career development, whilst during 2021 there is rapid recruitment of new data stewards across the Netherlands. <i>This confirms the raison d'être of the current project.</i></p> <p>Such a tool has to meet the needs of the data stewardship community and its stakeholders. It will need to be useful for the various types of data stewards, for employers looking at recruitment, for employers and employees wanting to make an assessment of skills and development needs, and for training providers wanting to develop new training opportunities.</p>

*Table 6.1 Challenge for the development of data steward skills tool*

This single challenge summarises several challenges mentioned in other chapters, for instance:

- There exist many different types of data steward and it would be useful to be able to analyse and compare these roles across, or within, institutions to clarify responsibilities and to ensure that staff teams contain complementary skills (Chal\_Job\_3, Chal\_Tr&Ed\_2).
- Individual data stewards have no benchmarks against which they can assess their own skills and knowledge. Furthermore, they have very varied backgrounds and therefore have very different development needs (Chal\_Case\_1, Chal\_Case\_2, Chal\_Tr&Ed\_1).
- Employers have to rely on previous examples of job vacancies to describe the competences for new job profiles and assess potential candidates (Chal\_Job\_1).
- Data stewards have no clarity in potential career progression either within their employing institution, or in a different role elsewhere (Chal\_Job\_2).
- Data stewards and employers often rely on serendipity to discover training opportunities (Chal\_Case\_3).
- Data stewards and their employing institutions frequently rely on national networks to share ideas and information, but with the growing population of data stewards it is becoming increasingly difficult to keep track of all opportunities and examples of best practice (Chal\_Case\_4, Chal\_Job\_4).

### 6.3 Basic features of a data steward skills tools

As a first step, we discussed the type of tool needed to bring together the competences and the training needs defined earlier in the project, to present them in a useful way to data stewards, their local organisations and other relevant stakeholders. This resulted in the exploration of two existing tools focused on data steward competences and training:

- Competency Hub<sup>81</sup>, which is used by various health and bioinformatics groups.
- CILIP Professional Knowledge and Skills Base (PKSB)<sup>82</sup> used by the library association in the UK.

Both of these tools provide a way to produce a profile of a person or job whereby the level of skills or knowledge are rated against a graded level of a range of competences and go on to link to training opportunities to develop those knowledge or skills. Images from these two tools can be found in Annex 11.

In March 2020, the project group contacted both tool managers to ask if they would be interested in working with us to develop a new version of their tool that would include data stewardship skills. CILIP indicated that although they were due to start a funded project to revise the tool and had interest in our view on how data stewardship does (or does not) align with the revisions that will likely be proposed for the PKSB tool, they could not commit to work on our project at this time.

The Competency Hub was, however, interested to expand their tool for our use case and discuss adaptations needed. Based on the matrices of the ZonMw data stewardship project (see Section 2.2), and in collaboration with ELXIR, as a pilot, content was added to the Competency Hub tool: expertise areas, responsibilities, tasks, KSAs, and learning objectives (see Annex 11).<sup>83</sup> This exercise shows that the Competency Hub can very well deliver in the short term. Obviously, decisions about the sophistication of the tool would depend on funding available.

Independently of existing tools, ideas have been discussed about how users might expect to use a tool to assess their competences and find out about training opportunities. That might help to decide about how a tool could be designed or adapted. Five different data steward personas have been identified, to show how stewards would benefit from the tool and what possible pathways they might want to take to move through such a data steward skills tool.

---

<sup>81</sup> <https://competency.ebi.ac.uk/framework>

<sup>82</sup> <https://www.cilip.org.uk/page/PKSB>

<sup>83</sup> The NPOS/ELIXIR Competency Framework, see <https://competency.ebi.ac.uk>

An example data steward persona and corresponding pathway is presented here (Figures 6.2 and 6.3). All personas and pathways can be found in Annex 12.

Annex 13 gives an overview of other features of a data steward skills tool, whether a new tool is built from scratch or an existing tool is adapted to our needs. This includes an indication of the broader audience that may benefit from such a tool.

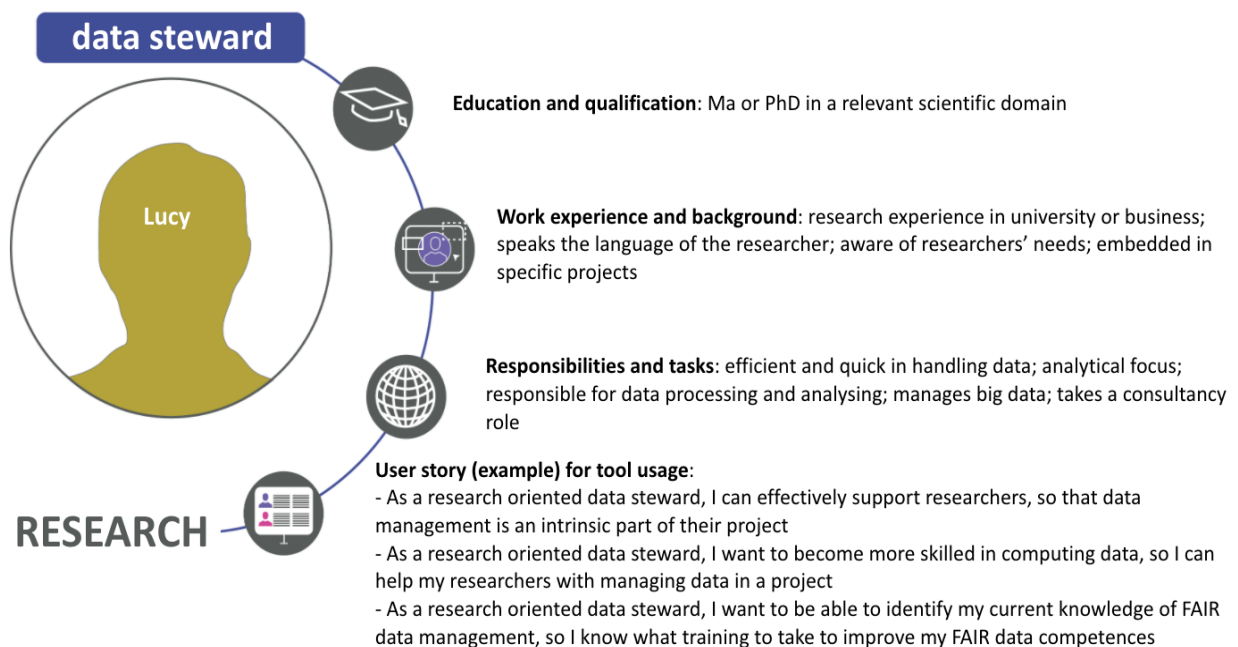


Figure 6.2 Research data steward persona (individual perspective)

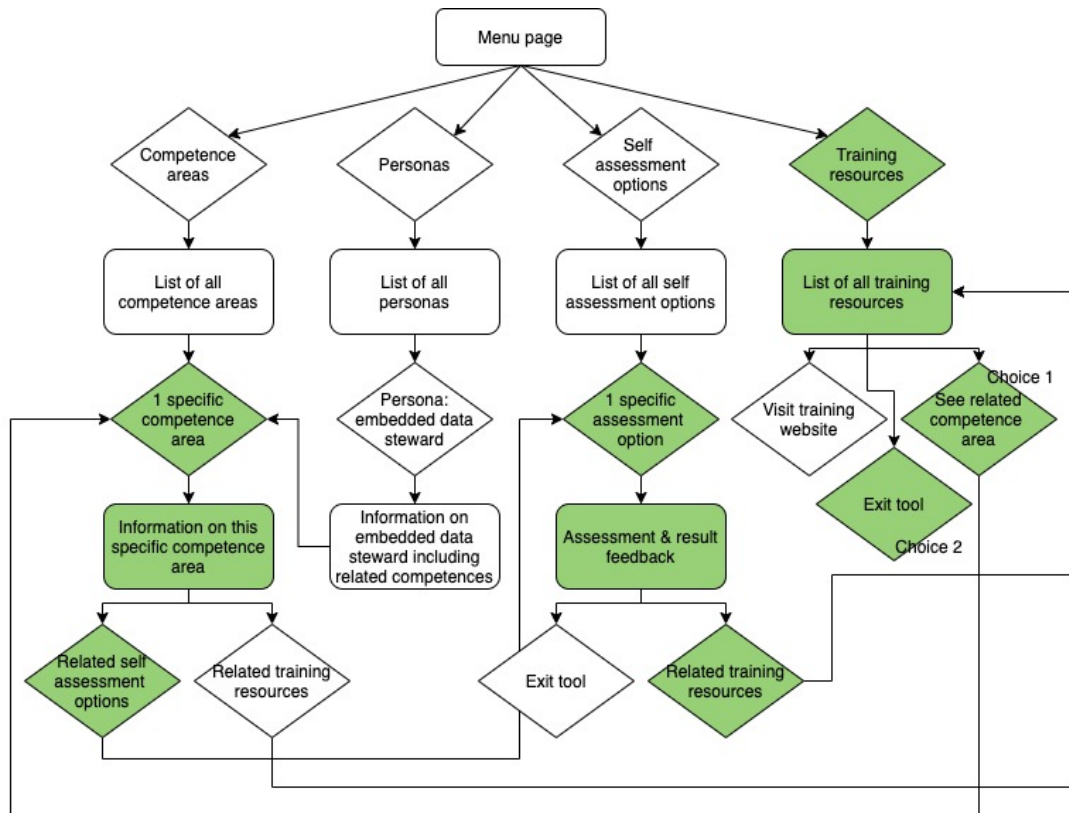


Figure Annex 6.3 Path 4, training resources, example learner journey of a research data steward

#### 6.4 Recommendations for a data steward skills tool

The further development of a data steward skills tool involves mainly umbrella organisations and/or efforts of an implementation team. Based on the challenges and output, the following recommendations are made to these stakeholders.

Please note that some recommendations are addressed to an implementation team - not in place yet -, which is recommended to support the umbrella organisations in implementing the various recommendations.

Recommendations data steward skills tool		
Rec_Tool_1 Competency Hub integration	<p>Recommendation</p> <p>As soon as possible, and as far as the current tool setup allows it, add the project's competences, personas, and training opportunities to the Competency Hub</p>	<p>Stakeholders</p> <p>Umbrella organisations Implementation team</p>
	<p>Reasoning</p> <p>The Competency Hub is already available and to use at no cost, and therefore provides a quick solution that meets the immediate needs of the community. Content may quickly be added.</p>	<p>Challenges addressed</p> <p>Chal_Tool_1</p>
Rec_Tool_2 Committee of stakeholders for development process	<p>Recommendation</p> <p>NPOS board carries out a stakeholder analysis and sets up a committee of stakeholders to monitor the development process and whether the needs are met.</p>	<p>Stakeholders</p> <p>Umbrella organisations Implementation team</p>
	<p>Reasoning</p> <p>The committee of stakeholders decides whether further technical development of an existing tool is needed and possible. With our current knowledge, we suggest using the Competency Hub, as collaboration and willingness to adjust the tool to new requirements is already secured. However, stakeholders should be involved, and their needs should be clear before completing and launching a final tool.</p>	<p>Challenges addressed</p> <p>Chal_Tool_1</p>
Rec_Tool_3 Working group for content	<p>Recommendation</p> <p>The committee of stakeholders sets up an implementation team to deliver the content for the tool.</p>	<p>Stakeholders</p> <p>Umbrella organisations Implementation team</p>

	<p>Reasoning</p> <p>Setting up an implementation team guarantees sufficient time investment and involvement of all stakeholders. The content for the tool should include the current report's output and recommendations on competences, training and education of the current report, and the learning paths should be evaluated, adjusted if necessary and built into the tool. Use an agile development approach: design and develop one learning path, test it with end users, and adapt it. Add the other paths stepwise.</p>	<p>Challenges addressed</p> <p>Chal_Tool_1</p>
<p>Rec_Tool_4</p> <p>Potential owner inventory</p>	<p>Recommendation</p> <p>The committee of stakeholders approaches potential 'owners' of the tool to agree where decisions about the tool will lie. This tool-owner should ideally be an existing organisation or consortium.</p>	<p>Stakeholders</p> <p>Umbrella organisations</p> <p>Implementation team</p>
	<p>Reasoning</p> <p>To make sure the tool stays up to date, both from a content-related and technical perspective, it should be clear who owns, maintains, further develops and finances the tool. A national skills tool requires an explicit business model, identifying at least what the tool provides, for whom, and how maintenance and funding are arranged. These owners will also be able to answer questions regarding the tool from the data stewardship community and collaborate with similar international initiatives.</p>	<p>Challenges addressed</p> <p>Chal_Tool_1</p>

Table 6.4 Recommendations for the development of a data steward skills tool



## Chapter 7: Towards implementing FAIR data stewardship in the Netherlands

This chapter addresses the following topics:

- The role of data stewardship in the transition to FAIR data (Section 7.1)
- Recommendations from chapters 3-6, summarised (Section 7.2)
- Recommendations for specific stakeholders (Section 7.3)

### 7.1 The role of data stewardship in the transition to FAIR data

A roadmap for the implementation of professionalising data stewardship in the Netherlands should be positioned in the broader context of the NPOS FAIR data programme, as a joint effort of the NPOS-E and NPOS-F projects (see Section 1.3). The ambition of the NPOS FAIR data programme line is to achieve the necessary facilities and other preconditions for the optimal use and reuse of research data in the Netherlands, including a consistent system for FAIR research data, sustainable storage, and implementing standards for interoperability. This ambition contains a roadmap towards 100% FAIR metadata in the coming years, including a substantial part of the data itself being FAIR and, as much as possible, open.

Professionalising data stewardship is a crucial element in the transition to FAIR data. Setting ambitions with respect to FAIR data, in turn, gives rise to a high demand for professional data stewardship, both in capacity or number of data stewards and in their competences, that is, their knowledge, skills and abilities. It also questions the organisational position of data stewards, and the balance between, on the one hand, the embedded and more research-oriented data stewards and, on the other hand, the generic and more policy- and/or infrastructure-oriented data steward. It must be noted that, although data stewards are primarily based at research institutes, there is a common interest for all stakeholders, including policy makers and research-funding organisations, to ensure a sustainable position for data stewards. This implies a shared approach for recognising and rewarding data stewards and shared responsibility for financing their position. This, again, generates a demand for certified curricula to train current and future data stewards. Consequently, the stakeholders of the NPOS-F project should act now. By initiating and committing to this project, they raised expectations in the research data community.

The FAIR data ambitions are, in yet another way, influenced by professionalising data stewardship. In the Netherlands, the development of data stewardship as an emerging field of

expertise was recently further stimulated by the NWO call<sup>84</sup> to establish local digital competence centres (DCCs) in universities, university medical centres and universities of applied sciences, in which the appointment of data stewards has a central role. Consequently, an important reason for a FAIR data stewardship roadmap is facilitating these future DCCs at universities, university medical centres and universities of applied sciences in creating sufficient and adequate data stewardship capacity. A first implementation step is already taken in the context of the ZonMw funded COVID-19 programme, in which with help of GOFAIR, Health-RI and DTL, around 100 projects with their project leads and data stewards are trained to create metadata schemes to make COVID-19 data FAIR<sup>85</sup>. This example shows that FAIR data stewardship, though still in development, exists and will influence the data stewardship needs of the future Digital Competence Centres. The ZonMw FAIR COVID-19 programme also demonstrates that data stewardship developments are taking place in non-academic, private research organisations as well.

Furthermore, there is an abundance of international initiatives around the professional development of data stewards. Examples include high-level ambitions in the context of the European Open Science Cloud, with its Working Group on skills and training<sup>86</sup>, as well as the RDA Interest Group on professionalising data stewardship, which is a bottom-up initiative recently endorsed in the Research Data Alliance<sup>87</sup>. Dutch data stewardship should continue to be part of this movement. The data landscape in the Netherlands is rich but fragmented<sup>88,89</sup> and stands to benefit hugely from a concerted approach to professionalising data stewards.

## 7.2 Summary of the recommendations

As raised in the preamble, to professionalise data stewardship, it is necessary to address the following questions: what is a data steward, in what context does the data steward operate, and how do we recognise and reward the position of a data steward? Each of the preceding four chapters covered an aspect of this data stewardship landscape and drafted tangible recommendations that will help stakeholders in the Netherlands to take the next steps in professionalising data stewardship, i.e., to hire, train, and educate data stewards, while also establishing clear job profiles and career paths.

---

<sup>84</sup> <https://www.nwo.nl/en/calls/local-digital-competence-centres>

<sup>85</sup> <https://www.health-ri.nl/workshops-delivering-fair-metadata-covid-19-data-portal>

<sup>86</sup> <https://www.eoscsecretariat.eu/working-groups/skills-training-working-group>

<sup>87</sup> <https://rd-alliance.org/groups/professionalising-data-stewardship-ig>

<sup>88</sup> De Vries, M., Kok, R., Bouwhuis, M., Schipper, P. 2020. NPOS Eindrapportage – Verkenning en optimaliseren nationaal datalandschap. <https://www.openscience.nl/projecten/project-e-verkenning-en-optimaliseren-nationaal-datalandschap>

<sup>89</sup> EOSC Synergy Landscape Report The Netherlands. <https://doi.org/10.17026/dans-2by-ereu>

The NPOS-F recommendations will be summarised from two different perspectives. Here, the general recommendations per topic are indicated. Based on the recommendations, a detailed next steps list is given in Section 7.3 for each of the stakeholders.

The recommendations per topic are the following:

*Landscape analysis of current data stewardship practices* (Chapter 3)

- Analyse local data stewardship case studies to clarify the needs of the data stewards of your organisation.
- Develop a coordinated approach to data stewardship, which includes care for the data steward.
- Collaborate locally and nationally in organising training.

*Job profiles and careers for data stewards* (Chapter 4)

- Formalise the data steward and research software engineer function profiles in the job classification systems and stimulate local organisations to adopt these.
- Local organisations should start building up data stewardship capacity and secure these positions in their organisation.
- Make sure that job profiles reflect the required competences, and are sensitive to diversity in background and expertise, as well as to differences in roles, types and positioning in an organisation.

*Education and training of data stewards* (Chapter 5)

- Standardise metadata for data stewardship training and education, based on the defined data steward competences, and use these metadata tags to create curated training resources.
- Develop a dedicated, if possible certified, curriculum to train current and future data stewards, to be able to meet the required data steward expertise level and capacity.

*Data steward skills tool* (Chapter 6)

- Create a data steward skills tool as a single point of reference for data steward competences, training and education.
- Add the competences, personas, and training opportunities presented in this report to the existing Competency Hub tool.
- Organise a committee of stakeholders and an implementation team for, among others, further tool development (see also Section 7.3).

### 7.3 NPOS-F recommendations for specific stakeholders

What are the foreseen steps in a FAIR data stewardship roadmap for the stakeholders represented in the NPOS-F project?

A first action on the roadmap to converge the actions and initiatives of the various stakeholders in the data stewardship landscape and to stimulate the further development of professionalising FAIR data stewardship, could be the installation of a temporary **implementation team** that - together with the NPOS **committee of stakeholders** - plans, consolidates and harmonises these actions and initiatives. The recommendations of this project often concern joint actions of local and umbrella stakeholders, which all will benefit from harmonising activities in setting joint next steps. This implementation team should work like a project, with explicit goals, an end date, and a budget. Identifying dependencies among the recommendations and prioritising recommendations should be the first step. Furthermore, the implementation team should periodically evaluate the progress of all recommendations, together with the committee of stakeholders.

#### *The NPOS steering committee*

- Make professionalising data stewardship a core element of the national strategy for FAIR data, including capacity and governance.
- Monitor and evaluate the progress of the implementation of this report's recommendations for professionalising data stewardship, in the context of the NPOS FAIR data programme and as part of the FAIR data stewardship roadmap.
- Initiate the building of the proposed data steward skills tool as a single point of reference for data steward competences, training and education, via the installation of an NPOS committee of stakeholders and the implementation team.

#### *Universities, university medical centres, universities of applied sciences, and their board members, deans and HR managers, including libraries and research/RDM/data support desks ("local organisations")*

- Recognise and reward data stewards by adopting formal job profiles, and secure their position, either embedded in a project - as a research data steward - or generically in a central support service, as a policy or infrastructure data steward.
- Stimulate continuous professional development and enable the data steward to adjust the job to the researchers' needs.
- Collaborate with other RPOs and umbrella organisations in developing a certified data steward curriculum and encourage the growth of data steward communities and networking activities.

- Reconsider the composition of research groups for the transition to Open Science by including data stewards, and additionally, train PhD candidates so that they are able to pursue a data stewardship career (research-oriented data stewards or data-oriented researchers)<sup>90</sup>. Indirectly, this also contributes to solving data steward capacity issues.
- Use the Plan-Do-Check-Act cycle to implement professionalising data stewardship, based on the recommendations of Chapters 3 through 6 and the local situation. The plan should be specific and include a budget for training and/or education. Ideally, institutional and departmental plans are reflected in individual development plans, linked to relevant data-related job profiles. Periodically the organisation should check if and how the plan is carried out, and whether there are reasons for revision, such as changing needs for professionalisation. Organisations could also share their experiences to further improve professionalisation.

*VSNU, NFU and VH and similar representative organisations (“umbrella organisations”)*

- Make professionalising data stewardship, including capacity and governance, a core element of the strategy for Open Science of RPOs.
- Stimulate the sharing of data steward function profiles in local organisations and the formalisation of these profiles in the job classification systems. Local HR managers play an important role, as representatives in the job classification system maintenance and development teams.
- Stimulate and facilitate the development of a data steward curriculum.

*Research-funding organisations, such as ZonMw and NWO*

- Develop a common approach to ensure sufficient data stewardship expertise in grant proposals and granted projects.
- Stimulate, recognise and reward the data steward as an essential project member in a granted project, and ensure sufficient budget for data stewardship in granted projects.
- Develop criteria for recognition and rewarding of data stewardship skills and activities in the output of granted projects.

---

<sup>90</sup> The 2020 PNN PhD survey also recommends to train PhD candidates to prepare them for other-than-research jobs <https://hetpnn.nl/wp-content/uploads/2020/08/PNN-PhD-survey-report-Criteria-Recognition-and-rewards-Open-science-Career.pdf>. Furthermore, this report’s case studies (Chapter 3) also show that data stewards often start their career as a researcher. Moreover, having research experience is an important element of the data steward job profile (Chapter 4).

*Representatives of the researcher communities, such as PNN, the networking organisation for PhD candidates, and the local Open Science communities*

- Stimulate more PhD candidates to take up an active data stewardship role in their research group to contribute to the transition to Open Science. Gaining data stewardship competences also provides extra career opportunities for this group of professionals<sup>91</sup>.
- Demonstrate to local and umbrella stakeholders that the function of the data steward - particularly the embedded, research-oriented data steward - first, should be situated in the primary research process and second, should be recognised and rewarded when it leads to creating research output.

*Service-providing, networking and training organisations, such as DTL, SURF, LCRDM, Health-RI, and RDNL*

- Collaborate with RPOs and umbrella organisations in further developing training and education for current and future data stewards. Seek national and international alignment on future certification of training and education.
- Facilitate local and national communities and networking activities, such as the Data Stewards Interest Group<sup>92</sup>, LCRDM pool of experts<sup>93</sup> and Health-RI Data Stewardship Community<sup>94</sup>.

---

<sup>91</sup> This is in line with the position paper position paper “Room for everyone’s talent: towards a new balance in the recognition and rewards for academics” <https://vsnu.nl/recognitionandrewards/wp-content/uploads/2019/11/Position-paper-Room-for-everyone%E2%80%99s-talent.pdf>

<sup>92</sup> <https://www.dtls.nl/community/interest-groups/data-stewards-interest-group>

<sup>93</sup> <https://www.lcrdm.nl/en/pool-of-experts-rdm-mailing-list>

<sup>94</sup> <https://www.health-ri.nl/health-ri-communities>

## Annex 1: Glossary and abbreviations

Term	Explanation
Capacity	An aspect of organisations, e.g., “data stewardship capacity” can be understood as “the number of capable data stewards” in an organisation.
Competence	Knowledge, skill and abilities relating to a topic e.g., ‘workflow set-up and management’ (based on EOSCPilot glossary <sup>95</sup> ).
Competence area	Nine competence areas for the data stewards can be distinguished: policy/strategy, compliance, alignment with FAIR, services, infrastructure, knowledge management, network, data archiving, and soft skills (detailed in Annex 3).
Data steward	A person responsible for keeping the quality, integrity, and access arrangements of data and metadata in a manner that is consistent with applicable law, institutional policy, and individual permissions. Data stewardship implies professional and careful treatment of data throughout all stages of a research process. A data steward aims at guaranteeing that data is appropriately treated at all stages of the research cycle (i.e., design, collection, processing, analysis, preservation, data sharing and reuse). (Based on LCRDM glossary <sup>96</sup> . See Section 2.1 for related definitions).
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.
Education	Education is a process of systematic teaching and learning. The term refers to formal curricula at universities, university medical centres, and universities of applied sciences. See also “Training”.

<sup>95</sup> Whyte, A. et al. (2019). EOSCPilot D7.5 Strategy for Sustainable Development of Skills and Capabilities. <https://eoscipilot.eu/sites/default/files/eoscipilot-d7.5-v1.1.pdf>

<sup>96</sup> <https://www.lcrdm.nl/en/glossary>

Higher-education institution (HEI)	Higher-education institutions or HEIs include traditional universities and profession-oriented institutions, which are called the universities of applied sciences (based on this source <sup>97</sup> ).
Implementation team	This team doesn't exist yet. Part of the recommendations is to install a team that should (help to) coordinate cross-organisational implementation of the other recommendations.
Local organisation	A local research-performing or higher-education institution, as opposed to "umbrella organisation".
Professionalisation	All education and training activities that increase a data steward's competence.
Research data management	This report uses "research data management" (RDM) as interchangeable with data stewardship.
Research software engineer	A research software engineer (RSE) combines professional software engineering expertise with an intimate understanding of research (Society of Research Software Engineering <sup>98</sup> ).
Soft skills	People's abilities to communicate with each other and work well together (Cambridge Dictionary <sup>99</sup> ).
Training	Training is a process of point-of-need teaching and learning, typically focusing on direct usability in the working place. See also "Education".
Umbrella organisation	In this project we distinguish three umbrella or membership organisations: the VSNU representing universities, the VH representing universities of applied sciences, and the NFU representing university medical centres. Umbrella organisations are stakeholders that can endorse strategic or political decisions. See contrasting term "local organisation".

Table Annex 1.1 Glossary

<sup>97</sup> <https://www.igi-global.com/chapter/service-oriented-architecture-higher-education/16788>

<sup>98</sup> <https://society-rse.org>

<sup>99</sup> <https://dictionary.cambridge.org/dictionary/english/soft-skills>



Term	Explanation
BKO	University Teaching Qualification (Dutch: Basiskwalificatie Onderwijs)
BROK	Foundational course in regulation and organisation for clinical researchers (Dutch: Basiscursus Regelgeving en Organisatie voor Klinisch onderzoekers)
DAC	Data Access Committee
DCC	Digital Competence Centre
DMP	Data Management Plan
DPIA	Data Protection Impact Assessment
ELIXIR	Intergovernmental organisation that brings together life science resources from across Europe ( <a href="https://elixir-europe.org">https://elixir-europe.org</a> )
EOSC	European Open Science Cloud ( <a href="https://www.eosc.eu">https://www.eosc.eu</a> )
FAIR	Findable, Accessible, Interoperable, Reusable
FTE	Full-time equivalent
FUWA-HBO	Job classification system in Dutch universities of applied sciences
FUWAVAZ	Job classification system in Dutch university hospitals
GCP	Good Clinical Practice
GDPR	General Data Privacy Regulation (Dutch: Algemene Verordening Gegevensbescherming)
HANDS	Handbook for Adequate Natural Data Stewardship
HAY	Job classification system in Dutch universities of applied sciences
HBO	Higher education (Dutch: Hoger Beroepsonderwijs)
HEI	Higher-Education Institutions
HR	Human Resources

KNAW	Royal Netherlands Academy of Arts and Sciences (Dutch: Koninklijke Nederlandse Akademie van Wetenschappen) ( <a href="https://knav.nl/en">https://knav.nl/en</a> )
KSA	Knowledge, skills and abilities
LCRDM	National Coordination point Research Data Management (Dutch: Landelijk Coördinatiepunt Research Data Management) ( <a href="http://www.lcrdm.nl/en">www.lcrdm.nl/en</a> )
MBO	Secondary vocational education (Dutch: Middelbaar Beroepsonderwijs)
NFU	National association of collaborating university hospitals ( <a href="http://www.nfu.nl">www.nfu.nl</a> )
NGWI	Netherlands Code of Conduct for Research Integrity (Dutch: Nederlandse Gedragscode voor Wetenschappelijke Integriteit)
NWO	The Dutch Research Council ( <a href="http://www.nwo.nl">www.nwo.nl</a> )
NPOS	National Platform Open Science ( <a href="http://www.openscience.nl">www.openscience.nl</a> )
NVAO	Dutch-Flemish Accreditation Organisation ( <a href="https://www.nvao.net">https://www.nvao.net</a> )
ORCID	Open Researcher and Contributor ID
RDA	Research Data Alliance ( <a href="https://rd-alliance.org">https://rd-alliance.org</a> )
RDM	Research Data Management
RDNL	Research Data Netherlands ( <a href="https://researchdata.nl">https://researchdata.nl</a> )
RPO	Research-Performing Organisation
RSE	Research Software Engineer
TO2	Collaborating Applied Research Organisations (Dutch: Samenwerkende Toegepast Onderzoek Organisaties) ( <a href="https://www.to2-federatie.nl">https://www.to2-federatie.nl</a> )
UAS	University of Applied Sciences (Dutch: hogeschool)
UB	University library (Dutch: Universiteitsbibliotheek)
UFO	Job classification system in Dutch universities (Dutch: Universitair Functieordeningssysteem)

UMC	University Medical Center
VH	Netherlands Association of Universities of Applied Sciences (Dutch: Vereniging Hogescholen <a href="https://www.vereniginghogescholen.nl">https://www.vereniginghogescholen.nl</a> )
VSNU	National association of collaborating universities ( <a href="http://www.vsnu.nl">www.vsnu.nl</a> )
WMO	Medical Research Involving Human Subjects Act (Dutch: Wet medisch-wetenschappelijk onderzoek met mensen)
WO	University-level education (Dutch: Wetenschappelijk Onderwijs)
ZonMw	The Netherlands Organisation for Health Research and Development ( <a href="https://www.zonmw.nl">https://www.zonmw.nl</a> )

*Table Annex 1.2 Abbreviations*

## Annex 2: Case studies

### Case study 1 Delft University of Technology (TUD)

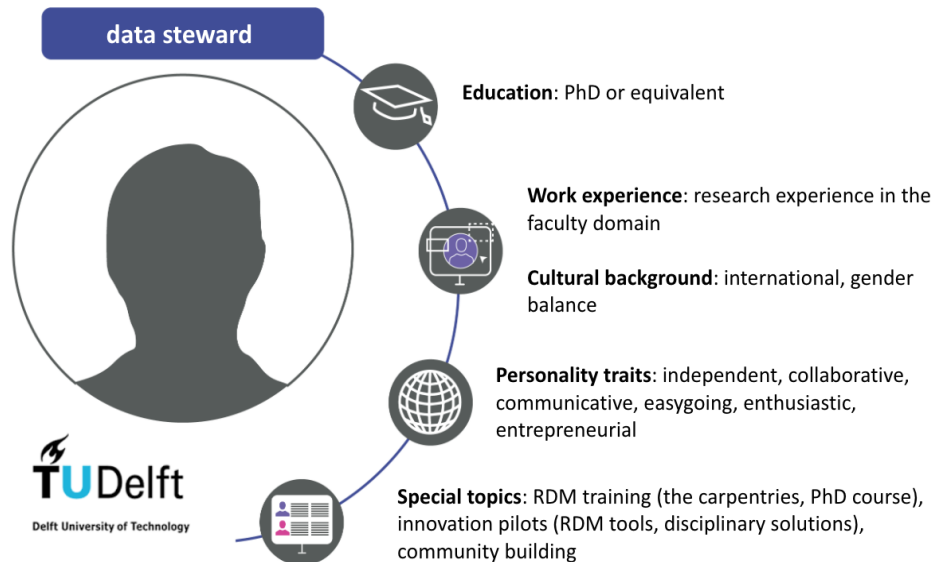


Figure Annex 2.1 Delft University of Technology data steward



Figure Annex 2.2 Delft University of Technology case study reference card

### *Organisation and funding*

I am one of eight data stewards at TU Delft (TUD). We have one data steward at each TUD faculty, I am therefore part of a team. Data stewards are appointed as full-time, permanent members of staff and are managed by faculty executive secretaries. Our work is coordinated by the data stewardship coordinator, a full-time, permanent employee of the library.

### *Training*

When I started, I received external training focused on data management, as well as internal training on research support at TUD. Training was arranged by the coordinator.

### *RDM training*

- *Data management crash course*  
I took the data management crash course Essentials 4 Data Support<sup>100</sup>. I learnt “the basic knowledge and skills (essentials) to enable a data supporter to take the first steps towards supporting researchers in storing, managing, archiving and sharing their research data”.
- *Software and data carpentry workshops*  
The Carpentries training<sup>101</sup> was useful to get practical skills on working with code and data. I joined the software carpentry workshop and subsequently became a certified carpentry instructor. Since then, I often act as a helper during carpentry workshops and frequently help organise and teach disciplinary data carpentry workshops.
- *Thematic training*  
I participated in various thematic workshops, e.g., training on GDPR, data de-identification, ethics, FAIR principles, qualitative data management, reproducibility, code management.

### *TUD research services*

- *Intro to 4TU.ResearchData*  
This training covered the benefits of archiving and the data deposition workflow. It also highlighted available support services.
- *Data management planning*
- *LIBguides about how to search data*  
This training explained the key stakeholders involved in data management planning, e.g., the privacy team, ethics committee, ICT. It also introduced funders’ data policies. In addition, I learnt about the workflow for DMP support, including the institutional DMP tool, and the organisation of DMP requests.

---

<sup>100</sup> <https://datasupport.researchdata.nl/en>

<sup>101</sup> <https://carpentries.org>

- *ICT support*  
This session introduced storage solutions available at TUD, as well as security measures recommended to researchers.
- *Open Access services*  
This training provided an overview of the TUD Open Access publishing policy and tips where to find relevant information and expertise on Open Access publishing.
- *Other training*  
Additional sessions were organised on request from the team. For example, we had a session on GitLab, version control, persistent identifiers (such as ORCiDs), ethics, academic integrity, patenting and IP, training development.

#### *Education outside of training*

To be effective in my role, I constantly need to learn new skills to support researchers in this fast-evolving environment. These less formal and less structured learning efforts are also an important part of my education.

#### *Mentoring*

My faculty line manager also acts as my mentor and helps me embed within my faculty. Additional mentorship is offered by the coordinator, who has regular meetings with each one of us.

#### *Internal peer support*

Crucial part of my education is also learning from other data stewards. We work as a team. Collaboration and team building is helped by:

- Weekly team meetings: we update each other about our work, share problems and challenges
- Slack workspace: we use it to quickly exchange messages and to ask each other for support. Saves a lot of emails!
- Shared drives: we share all our documents with the team
- Weekly drinks: we meet for regular after work drinks to informally chat about non-work matters and socialise. This really helps with trust building.

#### *External peer networks*

In addition to internal networks, we also exchange practice with other colleagues externally. Many of us are members of the DTL Data Stewards Interest Group (DSIG)<sup>102</sup>, various LCRDM

---

<sup>102</sup> <https://www.dtls.nl/community/interest-groups/data-stewards-interest-group>



task groups<sup>103</sup> and the Research Data Alliance (RDA)<sup>104</sup>.

#### *Soft skills*

Soft skills are essential in my daily job. I gained these skills through a combination of formal training organised by the coordinator (e.g., presentation skills, theatre and performance skills, running collaborative and engaging meetings), as well as informal training, mostly learnt on the job (e.g., chairing meetings, organising workshops, sharing feedback, or through collaborative development of publications, surveys, workflows, opinion pieces).

#### *Research expertise*

I also had to gain expertise in the research area of my faculty. I achieved this primarily through developing networks with researchers at my faculty (in particular, data champions), participating in seminars and attending relevant conferences.

---

<sup>103</sup> <https://www.lcrdm.nl/en/lcrdm-task-groups>

<sup>104</sup> <https://www.rd-alliance.org>

## Case study 2 Radboud University (RU)

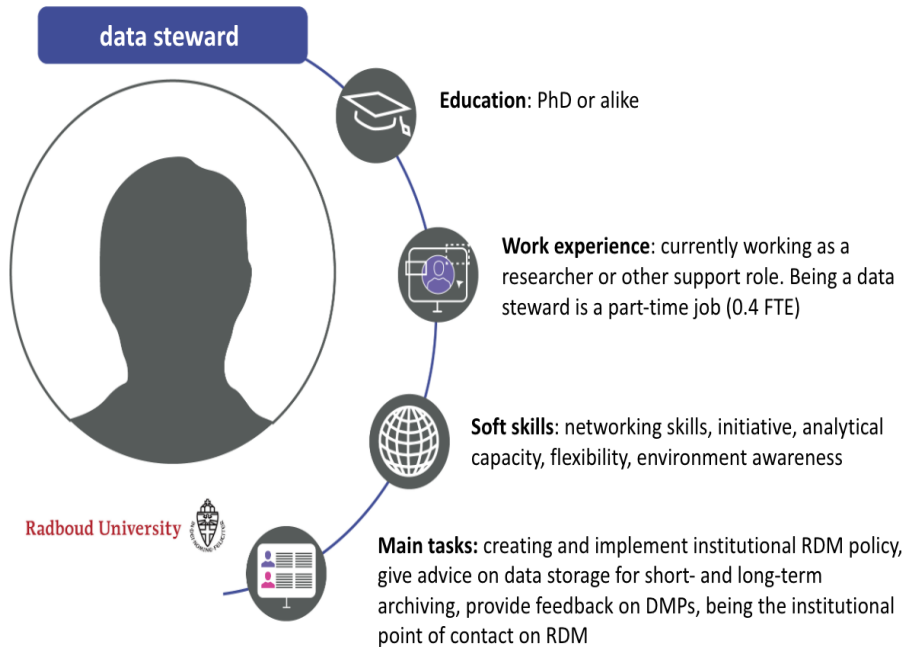


Figure Annex 2.3 Radboud University data steward

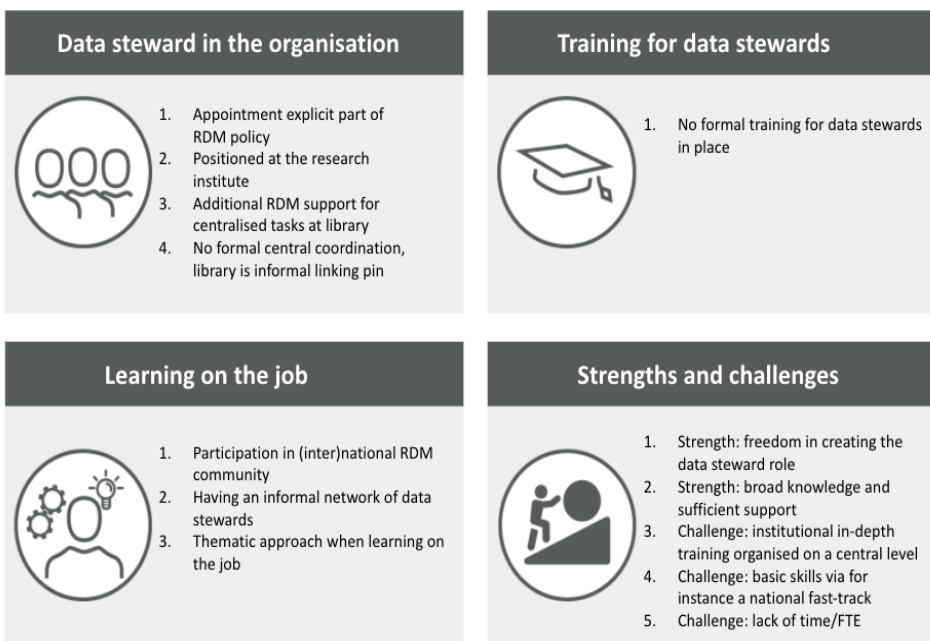


Figure Annex 2.4 Radboud University case study reference card



*Profile: data steward in the organisation*

In 2017, I was appointed as data steward at one of the research institutes at Radboud University. My appointment was a result of Radboud University's RDM policy<sup>105</sup>, which among others states that the generic RDM policy was to be supplemented by the research institutes. To push this effort, and as part of the recent strategic vision of the RDM steering committee, appointing a data steward has been a main element of the research data management policy at the research institutes.

Since then, all sixteen research institutes have appointed a data steward, who have their offices at the institute. My job as data steward is part of the daily research service office of the faculty, and research data management formally falls under the responsibility of the research director. However, this differs between institutes, as it is the institutes that have appointed data stewards and positioned them in the organisation.

The background of data stewards differs for each research institute. I have completed a PhD, just like many of my data steward colleagues. I'm a part time data steward, in addition to in my case my main job at the university library (RDM support), just like most of my data steward colleagues, who are for instance (almost full time) lab manager, research support officer, but also often researchers. The RDM steering committee is reviewing the current system of data stewards at Radboud University to strengthen the position of the data steward.

Formally, the data stewards aren't centrally coordinated, but the university library's RDM coordinator is informally the link between the data stewards, the university library, and other university support services in the field of research data management, data stewardship and research support.

*Training for data stewards*

At Radboud University, there is no training specific for data stewards that I can enroll in. RDM and Open Science training for PhDs and other researchers are currently offered by the RDM support team of the university library. I am free to join to broaden my knowledge on research data management. However, the information provided in this training is tailored for researchers and not specific for me as a data steward.

I was informed by the university library of the existence of (inter)national training, such as the Essentials 4 Data Support, Mantra<sup>106</sup> and the Carpentries, but this isn't particularly stimulated nor centrally coordinated at Radboud University. Moreover, the limited capacity of my

---

<sup>105</sup> <https://www.ru.nl/rdm/vm/policy-documents>

<sup>106</sup> <https://mantra.edina.ac.uk>

employment makes it complicated to spend the necessary time on these training.

### *Learning on the job*

As a consequence of the lack of training, I'm mainly self-trained. This is difficult since I'm the only data steward in the institute and faculty and I have no other data stewards in my direct environment that I can learn from or ask for help. On the other hand, as a data steward, I have a broad direct network with data steward colleagues in the university as well as with data-related functions in my institute and faculty (such as privacy officers, legal officers, policy makers, ICT staff, ethical committees and data curators), whose expertise areas converge somewhat with mine, as being a data steward includes very diverse tasks. This is also a big part of learning on the job. It also helps that we have a central RDM support team at the university library.

I am part of the LCRDM pool of experts<sup>107</sup>, in which we share information on a national level, and I subscribed to a number of (virtual) communities (such as DTL Data Steward Interest Group), mailing lists (such as the one from LCRDM<sup>108</sup> and JISC<sup>109</sup>), newsletters and webinars to stay updated and ask the advice of the community.

An important part of my network are the facilities organised by the university library. Over the last few years, the university library organised monthly face-to-face meetings with individual data stewards, the university library's RDM coordinator and a research information manager from the ICT department. In 2019, this was changed to bi-monthly meetings with all data stewards, the university library's RDM support officers, the research information manager from the ICT department, and the RDM policy officer and secretary of the RDM steering committee. This really helps me find my way in the RDM landscape and implement the RDM policy in my research institute.

Right now, task groups are being organised on various generic topics, in the context of these Radboud University data steward meetings. This way, I work on updating my knowledge and skills together with other data stewards by sharing and discovering new information. Examples of task groups that are now being formed are data rights, data statements in PhD theses, engaging researchers, and data steward roles.

### *Strengths*

As I am the only data steward in my research institute, I have a great freedom in creating my own function. I can decide which tasks I find most important and will perform myself, and for

---

<sup>107</sup> <https://www.lcrdm.nl/en/pool-of-experts-rdm-mailing-list>

<sup>108</sup> <https://www.lcrdm.nl/en/pool-of-experts-rdm-mailing-list>

<sup>109</sup> <https://www.jiscmail.ac.uk/cgi-bin/webadmin?a0=research-dataman>

which tasks I ask the support of the RDM team in the university library. This way, I can ensure that the data steward profile fits in well with the discipline specific procedures and data life cycle of my research institute and its researchers. In line with this, I have the chance to choose myself in which competence areas I want to develop myself as a data steward, which offers great freedom for my career path.

Because of the widely varying group of data stewards at Radboud University, I have the possibility to meet with and learn about data stewardship and RDM in its broadest sense, including disciplinary approaches. Additionally, the support of the university library's RDM team and the information manager gives me many valuable insights in data stewardship. In my opinion, the data stewardship model at Radboud University provides a broad basic knowledge with sufficient support options if I don't have the expertise (yet) myself.

### *Challenges*

In order to grow as a data steward, I welcome specific training for data stewards, but I prefer them to be organised by my university, to reassure them that they cover what I need as Radboud University wants me to gain knowledge according to their data policy. This could be based on for instance the shared responsibilities and tasks of data stewards at Radboud University.

In my opinion, the training should be provided by the central RDM support team of the university library, or by a training organisation. For each specific, by the Radboud University assigned responsibility or set of tasks, I prefer the possibility of attending a specific session, to really deepen my knowledge on each of these relevant competence areas. This could be a face-to-face training or an e-learning module.

For new data stewards, I consider a fast-track training of great added value. This way, new data stewards are provided with a basic skill set to start their job. I favour a fast-track course offered on a national level, for instance based on the Essentials 4 Data Support training, Mantra or the Carpentries. Then, a training focused on the Radboud University assigned responsibilities and tasks I described earlier, would be a perfect institutional addition and would broaden the RDM knowledge, but again: tailored to Radboud University's data policy.

Another challenge I encounter is the lack of time I have to perform my job as a data steward. I only work part-time as a data steward (0.2 FTE, will be extended to 0.4 FTE in the near future). Since the research institutes are free to formulate the position of the data steward, my tasks are limited to the amount of time I have. However, in my opinion the data steward could play a bigger and more significant role in the institute when there is more time for the job.

### Case study 3 University Medical Center Utrecht (UMCU)

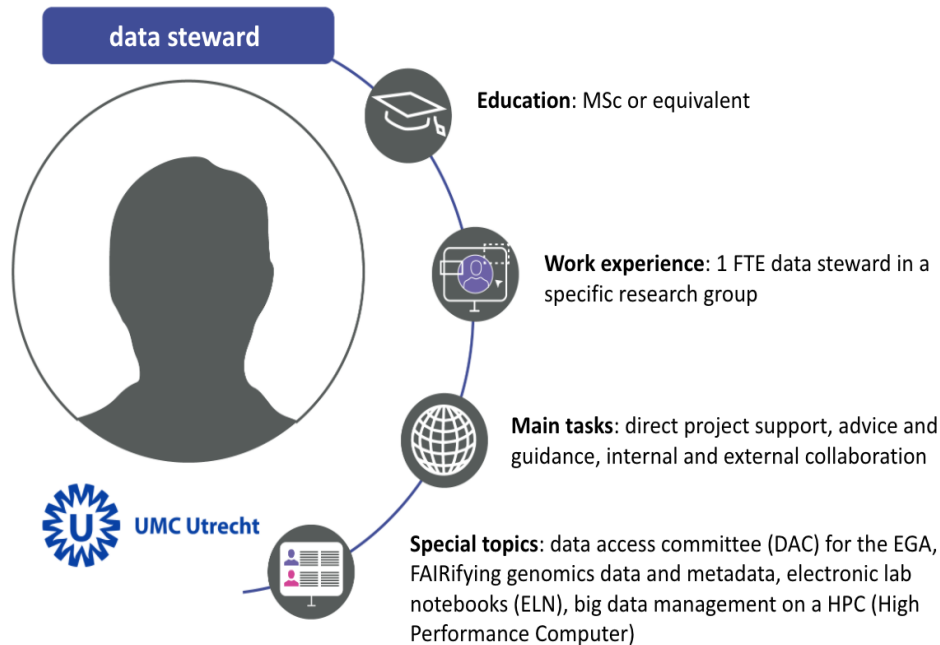


Figure Annex 2.5 UMC Utrecht data steward



Figure Annex 2.6 UMC Utrecht case study reference card

### *Personal narrative*

Since 2018, I am a data steward in the research group of Jeroen de Ridder in the Center for Molecular Medicine (CMM) at the UMC Utrecht. The genetics research institute CMM is part of the division Laboratoria, Apotheek en Biomedische Genetica (dLAB).

Based on my experience of having been part of a research group and operating on project level I would consider myself a departmental data steward.

My main focus lies on the research project that I support, but I also provide data stewardship services for the CMM, as well as the wider Utrecht bioinformatics community in the Utrecht Science Park - in the shape of the Utrecht Bioinformatics Center (UBC)<sup>110</sup>.

### *Organisational model*

From what I have perceived, the element that mostly distinguishes my data stewardship position from other data steward employments is funding of my position. My data stewardship services are part of the Utrecht Bioinformatics Expertise Core (UBEC)<sup>111</sup> of the UMC Utrecht, which is a facility that provides bioinformatic centred support and services. Based on my daily work experiences I chose to split my work focus in three major topics, in addition to operational work: i) project support, ii) collaboration, iii) advice and guidance. Due to the nature of being a project-focused data steward, the main work is concentrated on direct RDM support for a variety of research projects. Collaborating is a big part of data stewardship, therefore institutional, external, national and international participation in working groups and initiatives plays a big role. Lastly, offering general and tailored advice is a big responsibility and can be done via a variety of channels. For me, the work with a HPC (high performance computer) and the data access committee (DAC) are special topics from the bioinformatics and biomedical research domain I support.

### *How is the data steward embedded in the organisation?*

The UMC Utrecht contains three different parts: the hospital, another one for education and training, and the research part. In the areas overlapping hospital and research, data managers are in place to support the intersection of clinical and scientific work. The showcased data steward in this case is located in the genetics research division in the research part of the UMC Utrecht.

The research data management and data stewardship activities of the UMCU are coordinated and executed by the central research ICT. However, research data specific support services

---

<sup>110</sup> <https://ubc.uu.nl>

<sup>111</sup> <http://ubec.nl>

and a dedicated data archive are under development; the library services of the University Utrecht (UU) on campus are available for UMC Utrecht members as well.

There are divisional data managers in place to manage and foster data management related efforts on that organisational level. The highlighted data steward is part of a specific research group and operates on departmental - even institute specific level. The 1 FTE position is funded via project-based money streams.

#### *Training for data stewards*

Introduction and further education in the domain specific knowledge (i.e., bioinformatics) is available and encouraged by the research group leader. These mainly consist of online courses, or available literature.

Project management, soft-skill improvements, medical related certificates (i.e., BROK®: a basic course about regulations and organisation for WMO-project research<sup>112</sup>) are available via the UMCU internal training and education services.

#### *Learning on the job*

- The ins and outs of file formats, data types and volume, data protection needs of the domain specific data.
- Research and data protection compliance: informed consent and broad consent, WMO and non-WMO research, GDPR, METC (Medical Research Ethics Committee), etc.
- Data Access Committee (DAC) work for the dedicated domain specific archive: European Phenome and Genome Archive (EGA)
- Collaborations in the Utrecht Science Park.

#### *Challenges*

- Working individually on data stewardship on departmental level.
- Implementing data stewardship bottom up with limited resources and developing central support.
- Albeit being directly available for the researchers, awareness raising and general intro into RDM and data stewardship are needed first, before establishing domain best-practises.
- Improving and supporting RDM for wet-lab and dry-lab (i.e., computational work) is challenging.

---

<sup>112</sup> <https://nfu-ebrok.nl>



### *Strengths*

- Focus on research groups and research line specific topics and develop solutions and recommendations.
- Great freedom in structuring and planning workload and goals.
- Opportunity to find national and international collaborators to shape recommendations and best practises.
- Opportunity to find a solution for FAIR data for a specific data type/ research line.

## Case study 4 Erasmus MC

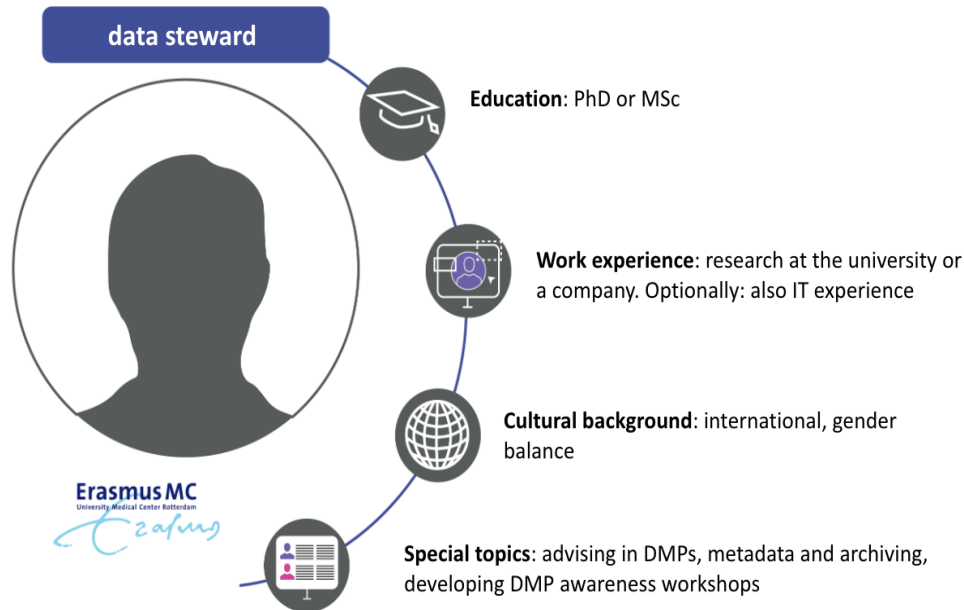


Figure Annex 2.7 Erasmus MC data steward

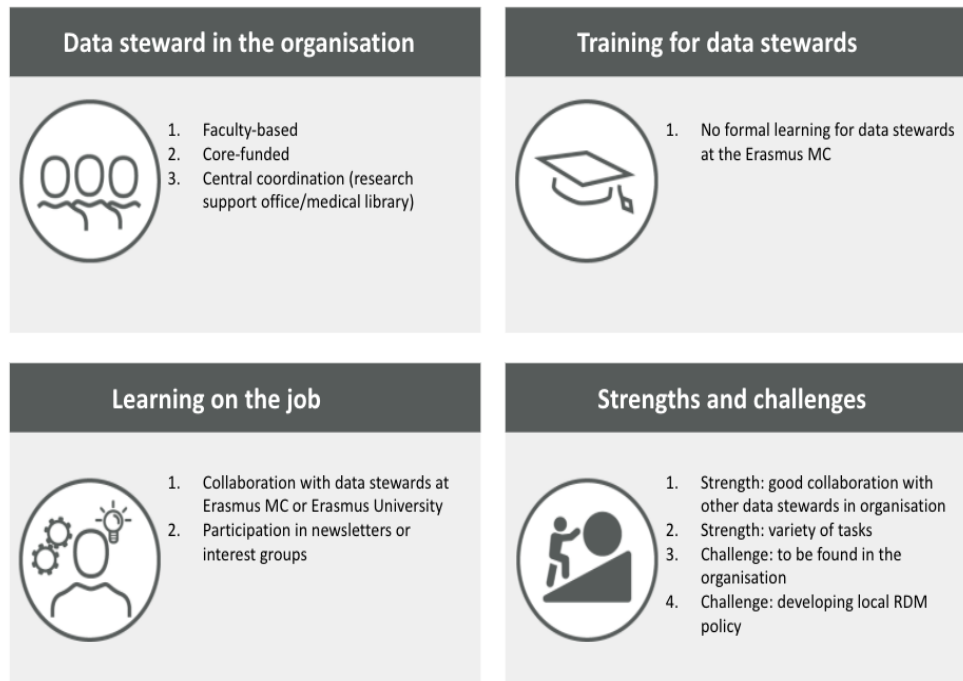


Figure Annex 2.8 Erasmus MC case study reference card



### *Profile: data steward in the organisation*

Like other university medical centres, the Erasmus MC has three core tasks: patient care, education and research. The employees with data steward-related tasks are all active in the research core. The data stewards are either working on a broad organisational level or are involved in specific, large research projects. In the latter situation, these “data managers” are the guardians of proper data management from the start until the end of the project. Most research projects do not have their own data steward, but they can consult the data stewards that have been appointed centrally in the organisation. The research support office has one data steward who collaborates with two information specialists from the medical library. I am one of the information specialists of the medical library with data stewardship-related tasks (0.2 fte) since April 2018. I mostly consult researchers about RDM while they are writing their DMP or when they want to deposit their data after their manuscript has been accepted for publication. Meanwhile, we started our awareness workshops on how to write a DMP.

### *Training for data stewards*

The Erasmus MC and Erasmus University (EUR) do not offer training for data stewards. However, we are encouraged to participate in external courses. I attended the Essentials 4 Data Support from RDNL, which formed a foundation for my RDM knowledge and how to implement this during support. Last year, I followed the three-day Helix Academy course FAIR Data Stewardship<sup>113</sup>. This year I started with the Delivering Research Data Management Services MOOC<sup>114</sup>, but could not complete it due to lack of time. Online courses are a good initiative, but it seems that I am more motivated to block/prioritise my agenda for group courses.

### *Learning on the job*

I learnt the most on our own work floor and the active collaboration between the medical library and the research support office. We subscribed to several newsletters or groups, such as LCRDM, Health-RI, DTL Data Stewards Interest Group (DSIG), and other biomedical information specialists. We also have regular meetings with data stewards from the university library.

### *Challenges*

- Making researchers aware of data steward assistance. Luckily, they become more aware since it is obligatory for some grants to consult a data steward while writing a DMP.
- Make the organisation itself (at a higher level) aware of the importance of data stewards. I do not have much time to work as a data steward, which makes it difficult to innovate.

---

<sup>113</sup> <https://helisacademy.com/en/data-analysis-stewardship>

<sup>114</sup> <https://www.futurelearn.com/courses/delivering-research-data-management-services>

Hopefully we can get more FTE and also experienced persons in the future. This project may help to achieve that.

- Not always being able to refer to RDM policy. The Dutch UMCs have HANDS<sup>115</sup>, but the Erasmus MC is still developing its own specific policy or guidelines. That makes it challenging to consult researchers.
- Collaboration with the university library. The university library represents all other faculties of Erasmus University. All these faculties together have the same size as the Erasmus medical library. The type of research is also different and therefore we cannot use the exact same facilities.

#### *Strengths*

- Strong collaboration with the well-known research support office and the ability to connect with the correct people in the organisation.
- There are just a few data stewards in the Erasmus MC to collaborate with. That makes it very easy to communicate with each other.
- Since there are just a few RDM related positions there is a variety in tasks for each person (could be a challenge too).
- A lot of freedom to design and execute your tasks.

---

<sup>115</sup> <https://www.health-ri.nl/services/hands>

### Case study 5 Maastricht UMC (UM/MUMC+)

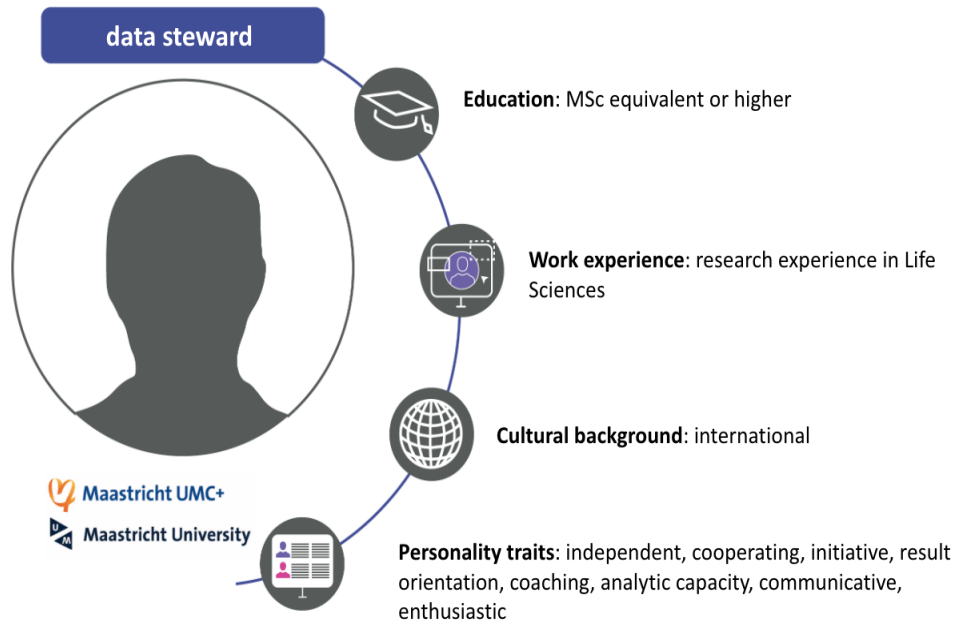


Figure Annex 2.9 Maastricht UMC data steward

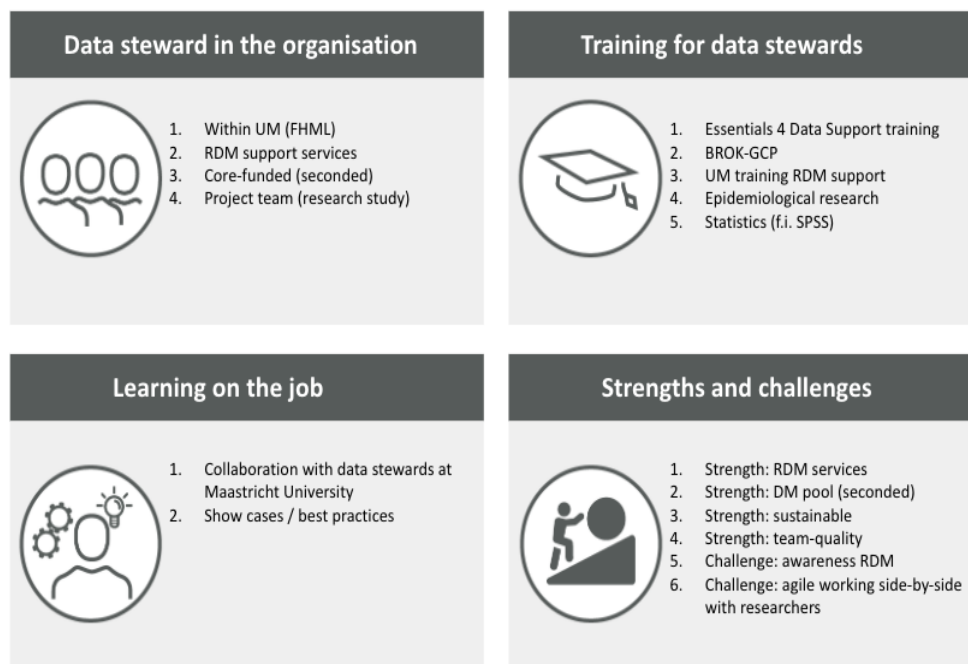


Figure Annex 2.10 Maastricht UMC case study reference card

### *Organisation and funding*

I am appointed at MEMIC, the centre for data and information management of the Faculty Health Medicine and Life Sciences (FHML) of Maastricht University (UM). I am seconded as a data steward to a large population-based cohort study and linked to six other data stewards at DataHub (as a linking pin). Our faculty and the academic hospital Maastricht together form Maastricht UMC+ (MUMC+). Our target group are research projects of FHML/MUMC+. The study I'm working on focuses on etiological issues where very different data is generated (questionnaires, imaging, biobank, etc.).

I advise researchers, PhDs and research support staff on management of research data in accordance with relevant regulations and the FAIR data principles. I have to oversee research data workflows (from data collection to storage) and use of (meta)data standards. My colleagues at MEMIC have different functions like data manager, research software engineer, data support employees, information manager, privacy officer and quality manager.

### *Education*

I started working at UM 30 years ago as a research employee, so I have extensive experience in scientific research. I have a master's degree in Mental Health. When I started at MEMIC in 2015, I received external training focused on research data management, good clinical practice (BROK® and GCP<sup>116</sup>) as well as internal training at our UB (library) on research support at UM.

### *RDM training*

- RDNL course: data management course Essential 4 Data Support. Focus on supporting researchers in collecting, storing, managing, archiving and sharing their research data.
- BROK-GCP: basic Course in Regulations and Organisation for Clinical Research.
- Library UM: an introduction to research data management specific for PhDs. How do you ensure that your data remain accessible and usable in the future, for yourself and others?
- Epidemiological research: design and interpretation (VUmc).
- Statistics: SPSS and others
- Leergang zelfbewust leiderschap: leadership training/personal leadership. This is important because I have to coordinate data managers (which are our embedded data stewards).

### *UM research services*

- CDDI (front runner DCC). MEMIC is one of the attendees of the Community for Data-

---

<sup>116</sup> <https://gcpcentral.com>

- Driven Insights (CDDI)<sup>117</sup>.
- Training Library UM and IDS (Institute for Data Science)  
Training in research data management, DMP, privacy, FAIR etc. I learnt general about DMP support, including the institutional DMP tool, and the organisation of DMP requests.
  - Others  
A lot of different courses and training. For example, SPSS, Ldot, Discover, Questionnaires (Castor, Qualtrics), Snowmed, Competence training etc.

#### *Other training*

To be effective in my role, I constantly need to learn new skills to support researchers in this fast-evolving environment. In the field of research data management, I am skilled in maintaining data security, data integrity and data privacy and I am able to help researchers to make their data FAIR.

#### *Mentoring*

The head of the department MEMIC acts as my mentor. Being mentored helps me to feel embedded in my faculty.

#### *Internal peer support*

Crucial part of my education as a data steward is learning from other data stewards, both from the faculty and the library (UB). Via collaboration and weekly meetings, we share problems and together tackle challenges.

#### *External peer networks*

In addition to internal networks, we also exchange practice with other colleagues externally. I'm also a member of the LCRDM pool of experts.

#### *Soft skills*

Soft skills are essential in my daily job. It is important to be able to understand researchers and PhDs. The most important soft skills are work independently, being cooperative, taking initiative, being result oriented, coaching, analytic capacities, being communicative and being enthusiastic.

#### *Research expertise*

I also have a lot of expertise in the research area of my faculty. I achieved this primarily through my job as a research employee and through networking with colleagues and researchers at my

---

<sup>117</sup> <https://library.maastrichtuniversity.nl/research-support/rdm/cddi>

faculty, active participation in seminars and attending relevant conferences.

#### *Challenges*

- Raise awareness of good RDM
- Close collaboration with “chain partners” of RDM support; hospital and university
- Continuing to learn from best practices
- Obtaining enough data stewards (generic and embedded) in future
- Working side-by-side with researchers

#### *Strengths*

- Strong collaboration with “chain partners” (is also a challenge)
- Broad pool of data stewards with different knowledge and skill levels
- Can build on years of experience with scientific research
- Close cooperation with GDPR-team (privacy)

## Case study 6 Avans University of Applied Sciences

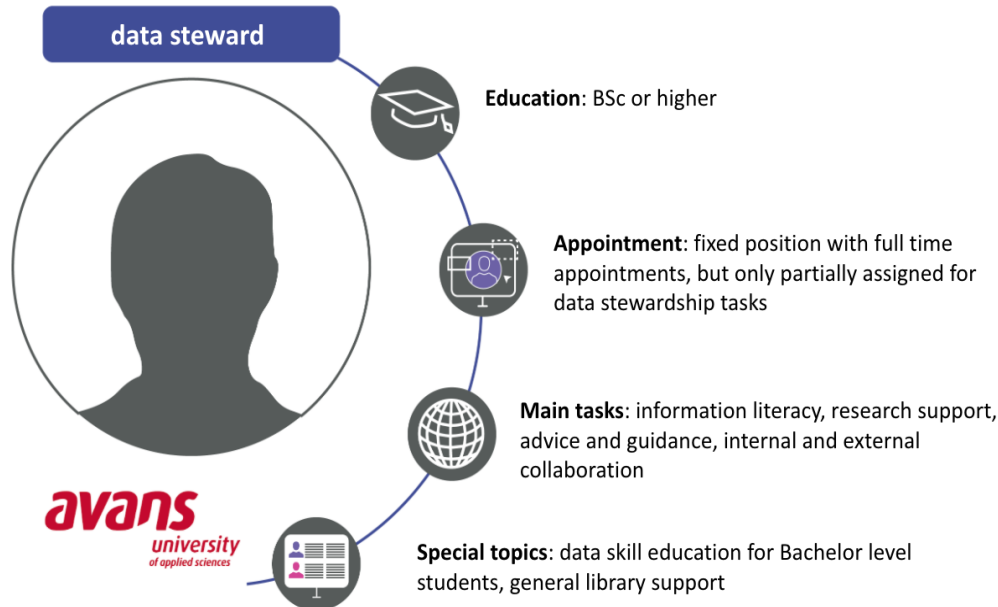


Figure Annex 2.11 Avans University of Applied Sciences data steward



Figure Annex 2.12 Avans University of Applied Sciences case study reference card

### *Organisation*

Avans started the Learning Innovation Centre (LIC)<sup>118</sup> to organise and promote education innovation. This learning centre aims to include new learning technologies in our daily work. Data stewardship is one of the construction directions. Such a learning innovation centre is a unique department among universities of applied sciences.

At the moment, no official data stewards are hired yet, but several staff members of LIC are coordinating data stewardship related activities. LIC members are fully hired with fixed contracts. The main focus of Avans is education, thus current data stewardship training is mainly given to teachers via individual academies. In the coming five years, the goal of LIC is to provide more support to research groups (Dutch: Lectoraten).

### *Training*

LIC members who are linked to data intensive subjects took part in the Essentials 4 Data Support training.

Further internal and/or external training in data stewardship is being planned. Several trainings have been taken by individuals to prepare for future roles in data stewardship.

### *LIC current general activities*

- Library (Xplora)
- Learning innovation in data skills
- Information literacy: provide LIBguides and provide training
- Digital and non-digital content
- Help academies with research skills and critical thinking for students

### *Challenges*

- LIC has just started in research support, and since it is a unique organisation, there is little experience to exchange with peer schools
- Implementing data stewardship bottom up with lack of resources and central support.
- There is no clear definition of the data stewardship role, and which tasks should be assigned to data stewards
- Support RDM for wet-lab and dry-lab (i.e., computational work) for the research groups.
- Lack of a helicopter view of available training, resources and requirements within and outside of the organisation

---

<sup>118</sup> <https://lic.avans.nl>





### *Strengths*

- Have a steady workforce (LIC) to tackle existing problems.
- Great freedom in structuring and planning workload and goals.
- Opportunity to find national and international collaborators to shape training recommendations and best practices
- Enough academies are willing to participate in data related innovations

## Case study 7 University of Applied Sciences Leiden

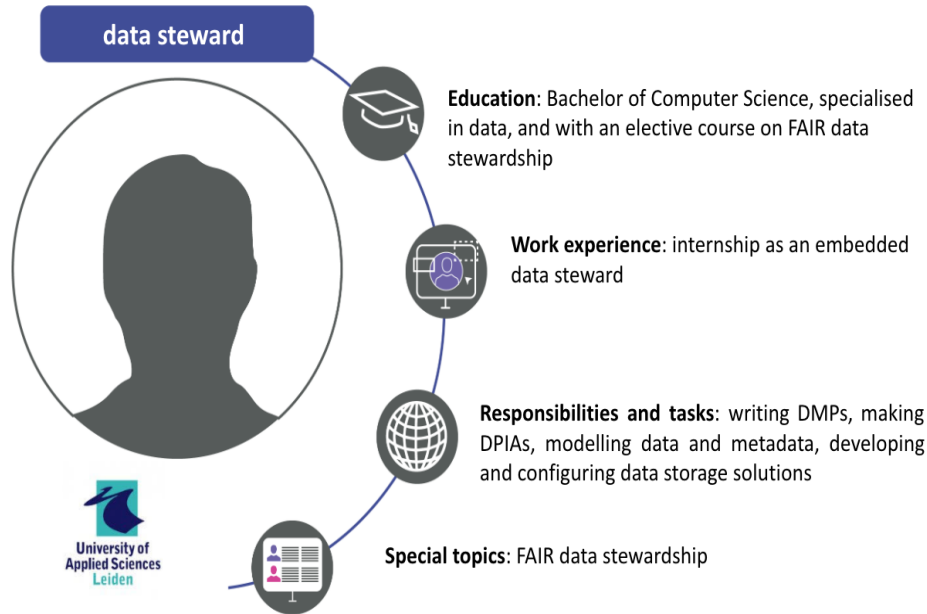


Figure Annex 2.13 University of Applied Sciences Leiden data steward



Figure Annex 2.14 University of Applied Sciences Leiden case study reference card

### *Organisation*

The Hogeschool Leiden offers Bachelor and Master programmes, minors, (post-graduate) courses and performs applied research in a variety of domains ranging from health care and social work to economics, and from technology to education. There are five faculties: Health, Social Work & Applied Psychology, Education, Management & Business, and Science & Technology.

### *Data stewardship*

Data stewardship in applied research at the Hogeschool Leiden is starting to take root. Currently there is one embedded data steward in an internship position at the Leiden Centre for Applied Bioscience (LCAB). Also, the Library and Policy & Advice are in the process of formalising data stewardship in the organisation.

### *Training in data stewardship*

The faculty of Science & Technology offers a Bachelor of Computer Science programme. One of the specialisations is Business Data Management. Furthermore, Business Data Management students can follow the elective course Introduction to FAIR Data Stewardship. This course has recently been developed in cooperation with the GO FAIR Foundation.

Students that take this path have general ICT knowledge and skills such as network technology and programming, and more specific skills in data storage, data engineering, data analytics, data management and FAIR data stewardship. These students can further develop themselves by doing a data steward internship.

Currently the first student is doing such an internship. He is the data steward intern mentioned above at the LCAB. In the interview below he talks about his work as a data steward and about his educational background.

### *What are your tasks and responsibilities as data steward at the LCAB?*

I am working on a research project about DNA tests. I have to write the data management plan for this project. Also, because there is patient data involved, I am making privacy impact assessments. Finally, I also have to design and configure the data storage solutions. When the project starts and the data is collected, I expect that my work might get more technical, for example developing ETL processes (Extraction, Transformation and Load).

### *Which courses that you followed during your studies helped you with these tasks and responsibilities?*

The courses Data Management and Introduction to FAIR Data Stewardship helped me the most, especially with writing the data management plan and with data storage solutions.



Business Process Modelling helped me with the privacy impact assessments. The courses about data architecture such as Data warehousing also helped me with the data storage solutions. In a more general sense, most of the Computer Science and especially the Business Data Management courses were very useful.

*What knowledge and skills do/did you need to perform your job as a data steward that you have not learned at the Hogeschool Leiden? How did you fill the gap?*

Especially for the privacy impact assessments I had to know quite a bit about laws and regulations that were not taught in the computer science program. For example, about retention periods and licences. Luckily, I have also done a minor in law. Furthermore, online, national courses such as Essentials 4 Data Support also helped me. Finally, it is sometimes hard for me to interpret the data because I don't have a background in genetics. I solve this by talking with the researchers.

## Case study 8 HAN University of Applied Sciences

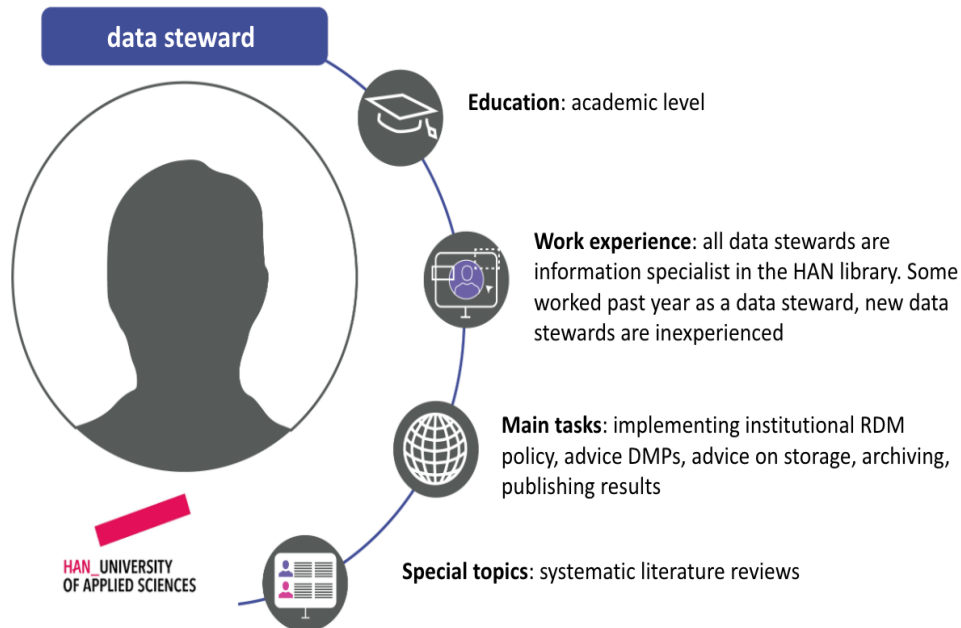


Figure Annex 2.15 HAN University of Applied Sciences data steward



Figure Annex 2.16 HAN University of Applied Sciences case study reference card

### *The organisation*

HAN has a campus at two locations, in Nijmegen and Arnhem. Both are home to different academies. Since the beginning of 2020, education and research have been organised in close proximity in an academy in order to bring practice-oriented research more into contact with education and vice versa. In the field of practice-oriented research, three focal points have been established within HAN: Health, Smart Region and Sustainable Energy & Environment.

Research support has recently been brought to the attention of researchers and lecturers. Before, it was already possible to ask questions via a mailbox, managed by information specialists with the role of data steward. At the same time, information about research support was available on part of the site of the HAN Study Centres. This has recently been adapted to the current working method of the data stewards.

Since the end of May 2020, an advisory report on the professionalisation of research support has been published. This includes proposals to change the name for research support to HAN Research Services. A year ago, 1 FTE was already allocated for data management tasks. Recently, these hours have been prolonged to the end of 2021, with extra hours for improving and expanding the services for practice-oriented research.

### *The role of data steward*

I worked in the field of data management, together with three colleagues. We are appointed as information specialists for a specific domain, having data management as an additional task. In addition to knowledge of data management, I have knowledge of the use of research and survey tools. My colleagues have different expertise, for example systematic literature review. Recently, however, I have changed jobs, so I'm no longer a data steward anymore.

### *Training*

I have completed a higher professional education (Library and Information Studies). This is the basis for my work as an information specialist. Over the years, I have gained a lot of experience in information services. These activities have similarities with the activities and skills of a data steward. During the pre-master's programme and later the master's programme in Educational Sciences that I followed, conducting and analysing (quantitative and qualitative) research was discussed and I gained experience in conducting research.

I gained knowledge about data management in the RDNL course Essentials 4 Data Support. It forms a proper basis for supporting researchers in dealing with data during their research. In my job, providing training to researchers was only a small part of the job. In the future, if this becomes a more important part of the activities, the HAN Academy offers the opportunity to obtain a Basic Didactic Competence. This is a nationally recognised course within higher

professional education, which allows you to teach at a university of applied sciences.

### *Learning on the job*

In addition to formal education and training, most of my knowledge was acquired informally, on the job. For instance: a question from a researcher had to be sorted out before I could answer it. As a team, we also observed gaps in knowledge on the use of systems, and consequently training about Microsoft Teams, Topdesk and SURF Research Drive were provided.

### *Strengths*

- Research support is still under development.
- HAN stakeholders within research such as policy makers, researchers and lecturers, see the added value of the data stewards. Close collaboration ensures that the workflow of the research cycle works more efficiently.

### *Challenges*

- HAN Research Support is still under development. It is difficult to shape the position of data steward.
- The data stewards have only received basic training, which is somewhat supplemented with data stewardship webinars offered nationally.
- The extension in FTE is only for one year, after which new funds should be allocated. This makes the future of the data stewards uncertain.
- In addition to the tasks of the data steward, the data stewards currently also perform tasks that are aimed at education (information specialist). This sometimes creates conflicts between the activities.

## Annex 3: Domain areas, responsibilities and tasks of a data steward

For a general understanding of what a data steward is and does, this annex provides an overview of the domain areas, responsibilities and tasks of a data steward. It is based on the ZonMw data stewardship project with its three distinctive data steward roles (data steward policy, data steward research and data steward infrastructure), but for this annex combined for one, overarching data steward. The original, detailed descriptions can be found in the ZonMw project's final report<sup>119</sup> and the accompanying matrices<sup>120</sup>. The tasks can further be divided into competences, i.e., knowledge, skills and abilities (KSAs) for data stewards, which also can be found in the ZonMw project's final report and matrices.

### **Area 1: policy/strategy**

Development, implementation and monitoring of research data management policy and strategy for the research institute

#### *Description*

A data steward is responsible for the advice on, alignment with and/or development, implementation and monitoring of a research data management (RDM) policy and strategy for the organisation and its researchers and students. A data steward supports the full data life cycle, FAIR data and Open Science. A data steward aligns with the relevant stakeholders and makes sure to comply with financial and legal constraints. Data stewards include the requirements of an adequate data infrastructure and tool landscape that fits with the policy and strategy.

#### *In bullet points, a data steward*

- Gives sound advice on RDM policy and strategy
- Is responsible for the availability of an up-to-date institutional data management plan (DMP)
- Is responsible for advice and awareness of RDM policy, FAIR principles and Open Science
- Is responsible for advice on an adequate research data infrastructure and proper tools
- Advises policy officers (on a strategic and tactical level)

---

<sup>119</sup> Scholtens, S., Jetten, M., Böhmer, J., Staiger, Ch., Slouwerhof, I., Van der Geest, M. & Van Gelder, C.W.G. (2019, October 3). Final report: 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.3474789>

<sup>120</sup> Scholtens, S., Anbeek, P., Böhmer, J., Brullemans-Spansier, M., Van der Geest, M., Jetten, M. ... Van Gelder, C.W.G. (2019). Function and competencies matrices for three types of data stewards 'policy', 'research' and 'infrastructure' (Version 2.1). Zenodo. <https://doi.org/10.5281/zenodo.3239079>



*Main tasks and activities of a data steward*

- Monitors the institute's RDM policy
- Advises the institute's management to advance
- Explores new needs, opportunities and trends in RDM
- Develops (together with others) DMP templates
- Implements RDM as a regular aspect of doing research
- Translates the RDM policy to data infrastructure and tools requirements
- Creates an overview of available data infrastructure and tools of the institute
- Development and operationalisation of products and services in the RDM domain
- Translates international developments into policies and practices at the university
- Leads the development and implementation of the faculty's data management policy
- Liaison function at policy level
- Connects with stakeholders in the research domain and acts as sparring partner for leading scientists and board of directors
- Coordination of the innovatory information agenda as part of the complete research life cycle
- Safeguards the data architecture
- Optimising RDM processes
- Securing provenance and audit trail

**Area 2: compliance**

Compliance to relevant scientific, legal, and ethical standards

*Description*

A data steward provides advice to the organisation and its researchers and students, and is responsible for compliance with relevant codes of conduct, legislation and field specific standards. This includes the compliance of the data infrastructure and tool landscape.

*In bullet points, a data steward*

- Is familiar with, and works according to, 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), Good Clinical Practice (GCP) and other laws and regulation. Furthermore, a data steward continuously aligns with legal and ethical standards
- Is responsible for advice on compliance before, during and after a research project
- Is familiar with relevant legislation and field specific standards
- Is responsible for advice on the data infrastructure and tooling within legislation

#### *Main tasks and activities of a data steward*

- 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
- Consults with privacy and security officers
- 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)
- Monitors and supervises the execution of a project or data collection in line with the DMP and relevant codes of conduct and legislation, including ethical, legal and social issues (ELSI)
- Identifies gaps in policy and takes action if needed
- Train researchers and research support employees on compliance requirements,
- 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
- Advises researchers and research support staff about how to deal with privacy sensitive data in accordance with the (GDPR) guidelines
- Monitoring up-time and security breaches of servers and services

#### **Area 3: alignment with FAIR data principles**

Alignment to the FAIR data principles and the principles of Open Science

#### *Description*

A data steward is responsible for facilitating and supporting the organisation and its researchers and students in FAIR data and FAIR software, as well as aligning the RDM policy, DMP, software management plans and data infrastructure and tool landscape to the FAIR data principles, FAIR software principles and the principles of Open Science.

#### *In bullet points, a data steward*

- Is responsible for alignment with the FAIR data principles and the principles of Open Science
- Facilitates and supports FAIR data and FAIR software
- Is responsible for advice on an adequate research data infrastructure and proper tools
- Is responsible for alignment of the data infrastructure and tool landscape with the FAIR principles

*Main tasks and activities of a data steward*

- 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 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
- Engages researchers in developing metadata schemes and documentation standards to improve FAIR data and software
- 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
- Advise researchers on careful management of research data with the FAIR principles,
- Ensuring that research data from various domains is adapted in accordance with FAIR principles

**Area 4: services**

Availability of adequate support on research data management, in staff or services

*Description*

A data steward is responsible for delivering sufficient and adequate support on RDM in the organisation and its researchers and students. This includes support regarding the data infrastructure and tool landscape.

*In bullet points, a data steward*

- Is responsible for advice for the availability of adequate support on RDM for the researchers, PhDs and research support staff of the institute
- Is responsible for identifying the requirements of adequate data infrastructure for RDM for researchers
- Provides training in the field of RDM
- Analyses data management needs of a researcher/research group
- Proposes, implements and monitors workflows to improve RDM practices
- Oversees and assists in executing data collection, description, cleaning, merging, licensing, sharing and use of metadata standards

*Main tasks and activities of a data steward*

- Advises board, 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
- Provides training in the field of RDM
- Analyse data management needs of a researcher and/or research group
- Propose, implement and monitor workflows to improve RDM practice
- Oversee and assist in executing data collection, description, cleaning, merging, licensing, sharing and use of metadata standards
- Provides guidance and instruction on discovery, acquisition and use/reuse of data
- Solves (together with others) practical RDM problems
- Advises and supports researchers on data infrastructure and tools
- Advises, develops and takes care (standard) ICT solutions for recurring issues and questions are implemented
- Provides guidance and instruction on the use of data infrastructure and tools
- Support researchers, PhDs, and students to store, (re)use and analyse research data and information
- Advise researchers on careful management of research data and about procedures and technical aspects that are important for the quality of (meta)data
- Takes the lead in developing further service provisions at research group, discipline or faculty level
- Contributes to scientific articles and helps to draft subsidy applications
- Performing statistical analyses
- Processing data
- Building scripts to assist data control and data cleaning
- Correct data extraction
- Data capturing/data harvesting
- Data enrichment/linking datasets
- Reusing and developing (machine learning) algorithms
- Visualising data
- Working on complex data(processing) issues, providing input to improve processes or systems
- Developing new workflows for ingestion, aggregation and export of data

## **Area 5: infrastructure**

Availability of adequate data infrastructure for research data management

### *Description*

A data steward is responsible for identifying the needs of the organisation and its researchers and students regarding data infrastructure and tools for RDM, in accordance with the DMP and the RDM policy. This supports the full data life cycle, FAIR data and Open Science.

### *In bullet points, a data steward*

- Is 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
- Monitors the department's or project group's needs including supporting adequate access, in accordance with the DMP
- Identifies 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

### *Main tasks and activities of a data steward*

- Initiates and supervises requests and acquires data infrastructures and tools for RDM within the institute
- Supervises monitoring of the need, use and availability of data infrastructures and tools
- Advises the management of the institute on data infrastructures and tools
- Is aware of and advises the institute on relevant (inter)national data infrastructures
- Monitors the needs regarding data infrastructure and tools for RDM within the department, project or data collection
- Supports access to data infrastructure and tools for RDM
- Sets requirements for data infrastructure and tools for RDM
- 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
- Monitors a process, system or the data architecture to optimise it
- Building databases
- Caring for structured and secure data storage
- Securing provenance and audit trail

- Advising on the improvement of the data architecture
- Working on complex data (processing) issues, providing input to improve processes or systems
- Developing new workflows for ingestion, aggregation and export of data
- Finding optimal solutions for use of existing data infrastructure
- Extending existing data infrastructure

### **Area 6: knowledge management**

Adequate level of knowledge and skills on research data management within the institute, department or project

#### *Description*

A data steward is responsible for determining the adequate level of knowledge and skills on RDM in the organisation and its researchers and students. This includes knowledge and skills on RDM related data infrastructure and tools.

#### *In bullet points, a data steward*

- Is 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

#### *Main tasks and activities of a data steward*

- 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
- Monitors RDM skills in the department or project
- Identifies knowledge and skill gaps of the DMP, and takes action if needed
- Initiates and provides training on RDM, tailored to the needs of the researchers and research support staff of the institute
- Creates awareness on RDM among researchers and research support staff
- Explains the added value of RDM
- Introduces to new employees the institute's RDM
- 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, department and researchers and explains the added value to RDM
- Advise researchers and research support staff on careful management of research data

and about procedures and technical aspects that are important for the quality of (meta)data

- Develop or give training courses that relate to the RDM field
- Is proactive in knowledge dissemination, for example, by organising events
- As spokesman for the faculty has a mission to convince researchers of the added value of good research data management

### **Area 7: network**

Obtaining and maintaining a network of aligned expertise areas and relevant departments and organisations inside and outside the institute, department or project

#### *Description*

A data steward is responsible for obtaining and maintaining a network of aligned expertise areas and stakeholders inside and outside the organisation.

#### *In bullet points, a data steward*

- Is 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

#### *Main tasks and activities of a data steward*

- Within the organisation, builds and maintains a network of researchers interested in RDM
- Refers researchers to other RDM related facilities and services (legal, financial and/or operational), inside and outside the institute, the department or project
- Liaises with (technical) experts inside and outside the institute, the department or project
- Maintains a network with RDM related colleagues and other relevant departments and organisations
- Connects data support people with each other
- Refers researchers and research support staff to RDM related data infrastructure and tools, inside and outside the institute
- Take the initiative to establish contacts with discussion partners from both research or specialist domain groups and colleagues at faculty or discipline level in order to consult on subjects from the data management field
- Deploys his or her knowledge of (inter)national developments in the field of data management and Open Science
- Works to build up an (inter)national network in the RDM field

## **Area 8: data archiving**

Adequate support and data infrastructure for FAIR and long-term archiving of data of the institute, department or project

### *Description*

A data steward is responsible for identifying the requirements of adequate support and data infrastructure for FAIR and long-term archiving of data of the organisation and by its researchers and students.

### *In bullet points, a data steward*

- Is responsible for identifying the requirements of adequate support and data infrastructure for FAIR and long-term archiving of data of the institute, 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

### *Main tasks and activities of a data steward*

- 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, department, project
- Assesses whether internal and/or external data storage and archiving facilities meet the applicable requirements
- 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
- 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
- Support researchers and students to store, use/reuse and analyse research data and information
- Caring for structured and secure data storage after a research project



## Annex 4: Components of the research software engineer job profile

### Domain areas and responsibilities of a research software engineer

The following domain areas and responsibilities are based on the Netherlands eScience Center job profiles (Dutch: functiehuis) and the King's College London list of roles for research software engineers. At the time of writing this report, the Netherlands eScience Center was in the process of updating its job structure. The information in this section is based on the previously existing job structure. Readers who are interested in an updated version of the job structure may contact the eScience Center directly.

#### **Area 1: research implementation**

##### *Description*

A research software engineer is responsible for applying advanced ICT based on scientific questions within a well-defined context.

##### *In bullet points, a research software engineer*

- Is responsible for producing code, technical frameworks, user interfaces, and systems for modelling, analysis, storage, presentation, and simulation of research-intensive problems (and dissemination of results)
- Is responsible for applying advanced ICT based on scientific questions within a well-defined context
- Is responsible for working with colleagues to produce code, technical frameworks, user interfaces, and systems for modelling, analysis, storage, presentation, and simulation of research-intensive problems (and dissemination of results)

#### **Area 2: research analysis**

##### *Description*

A research software engineer is responsible for analysing research questions and turning them into practical software solutions.

##### *In bullet points, a research software engineer*

- Translates scientific questions into software solutions, defining requirements of technical solutions and overseeing the design and development of software
- Is responsible for translating scientific questions into software solutions

- Responsible for deploying existing domain knowledge, or rapidly accumulating more, to understand the computational algorithms, requirements and interfaces involved in a research programming project
- Is responsible for working with colleagues to define requirements for research-related technical solutions
- Is responsible for overseeing the design and development of several significant software code bases
- Is responsible for developing and/or documenting methodologies, applications and software that can be applied within a specific area of research
- Is responsible for advising on the technical feasibility of projects within a technically complex environment

### **Area 3: software project management**

#### *Description*

A research software engineer is responsible for managing all aspects of complex software projects, from the inception of the project to the delivery of a final software product.

#### *In bullet points, a research software engineer*

- Translates scientific questions of a complex nature into plans for extensive software projects, including costing, reporting and timing of the project
- Is responsible for translating scientific questions into project plans
- Is responsible for the design and delivery of technical solutions, and their integration into wider technical frameworks and strategies
- Is responsible for development and contributing towards the reporting on projects.
- Is responsible for team/project deadlines to be met
- Is responsible for providing itemised cost estimates for technical solutions suitable for inclusion in funding bids
- Is responsible for leading research projects of a large and complex nature, ensuring that the development and use of eScience is coordinated and carried out within the predetermined time, goals, quality and costs
- Is responsible for reporting on progress and results, and deciding on the direction of the project
- Is responsible for establishing non-grant funded sources for one or more code bases (e.g., service cost models, industrial partnerships)

#### **Area 4: knowledge sharing**

##### *Description*

A research software engineer is responsible for sharing knowledge in all relevant aspects of eResearch.

##### *In bullet points, a research software engineer*

- Is responsible for sharing knowledge, acting as technical and scientific coach, contributing to training initiatives and providing eResearch support
- Is responsible for sharing knowledge internally and externally by giving presentations, workshops and training courses and by speaking at national and international conferences
- Is responsible for using the expertise of colleagues in the development (or reuse) of methodologies, software and applications
- Is responsible to act as a technical and/or scientific coach for colleagues
- Is responsible for contributing to training initiatives organised by eResearch or their local faculty or department, including introductory programming and software courses
- Is responsible for providing online and face to face support, and associated documentation, for staff and students using software built or supported by eResearch

#### **Area 5: system, software, and data maintenance and support**

##### *Description*

A research software engineer is responsible for ensuring that eResearch systems and tools are functioning and performing as expected.

##### *In bullet points, a research software engineer*

- Monitors and maintains software and data infrastructures
- Is responsible for monitoring eResearch systems and tools, and patch/upgrade as required to ensure security and performance
- Is responsible for producing technical and end user documentation to aid the use, support, and maintenance of eResearch systems and tools

#### **Components of the research software engineer job profile**

Following the domain areas and responsibilities, we propose the basic components of the research software engineer job profile. The structure is loosely based upon the Netherlands

eScience Center and Kings College London<sup>121</sup>, and other publicly available job profiles<sup>122</sup>. The proposed basic profile could be used by organisations, and should ideally be integrated, in a form that they see fit, into the aforementioned job classification systems.

### Profile title

Research software engineer (RSE)

### Profile objective

To ensure that the research work can be conducted efficiently, and applying advanced ICT in the most suitable fashion, an RSE:

- Translates research questions into software solutions
- Analyses the technological needs of the research
- Oversees the implementation of necessary software components in collaboration with relevant stakeholders
- Develops and delivers training
- Monitors and maintains software and data infrastructures

### Profile context

Collaborates in conducting ICT intensive research with the following professionals:

- Research project team leader
- Principal investigator
- Other support service departments

### Profile result areas

Core activity	Result
Research implementation	<p><i>Translate research questions into ICT solutions</i></p> <ul style="list-style-type: none"> <li>• Develop ICT solutions for data intensive research</li> <li>• Participate in (inter)national developments of good research software practices</li> <li>• Advise on ICT infrastructure and tools</li> <li>• Contribute to other relevant policies and regulations (strategic and tactical level)</li> </ul>

<sup>121</sup> <https://www.kdl.kcl.ac.uk/blog/rse-career-development>

<sup>122</sup> <https://github.com/RSE-leaders/evidence-bank>

Research analysis	<p><i>Analysing and advising on suitable ICT solutions for conducting data intensive research</i></p> <ul style="list-style-type: none"> <li>• Defining requirements of technical solutions</li> <li>• Advising on technical feasibility of research projects</li> <li>• Developing and documenting methodologies for data intensive research</li> </ul>
Software project management	<p><i>Manage complex research software projects to deliver a final software product</i></p> <ul style="list-style-type: none"> <li>• Develop plans for extensive software projects</li> <li>• Develop technical solutions that integrate into wider technical frameworks</li> <li>• Ensuring research software projects are carried out within predetermined time, goals, quality and costs</li> <li>• Advise in budgeting of technical solutions for inclusion in funding bids</li> </ul>
Knowledge sharing	<p><i>Responsible for sharing knowledge on ICT skills</i></p> <ul style="list-style-type: none"> <li>• Contributing to training activities in ICT skills</li> <li>• Acting as technical and/or scientific coach to colleagues</li> <li>• Providing support to staff and students on using research software</li> </ul>
System, software, and data maintenance and support	<p><i>Ensure eResearch systems and tools are functioning and performing as expected</i></p> <ul style="list-style-type: none"> <li>• Monitoring and maintaining software and data infrastructures</li> <li>• Producing technical and end-user documentation of eResearch systems and tools</li> </ul>

*Table Annex 4.1 RSE profile result areas*

### **Profile soft skill competences**

The soft-skill competences for research software engineers are similar to those of data stewards, which are shown in Table Annex 5.2.

## Annex 5: Components of the data steward job profile

This annex proposes the basic components of the data steward job profile. The structure is loosely based upon the UFO standard, as a result of exploratory, joint efforts of the project team and the UFO working group to get the data steward integrated into the UFO job classification system. The proposed basic profile could be used by organisations, and should ideally be integrated, in a form that they see fit, into the aforementioned job classification systems.

The content is based on the data steward domain areas, responsibilities and tasks, and the competences of the ZonMw data stewardship project (see Annex 3), as well as on the LCRDM data stewardship project. Additionally, comparisons were made with the UFO job profiles Specialist scientific information and Developer ICT and the skills listed by the terms4FAIRskills project<sup>123</sup>.

### **Profile title**

Data steward

### **Profile objective**

To ensure that the research work can be conducted efficiently, and that the quality and interoperability of the data is sufficient for reuse, a data steward:

- Analyses the needs for research data management
- Identifies research data management support gaps
- Develops training and services
- Participates in latest developments on research data management locally, nationally and internationally
- Acts as a liaison with leads of other research services
- Advises policy makers and those implementing research data management ICT infrastructure
- Oversees research data workflows (from collection to publishing) and use of (meta)data standards

### **Profile context**

Reports to and/or receives hierarchical guidelines from one or more of the following professionals:

- Executive board
- Dean

---

<sup>123</sup> <https://terms4fairskills.github.io>

- Policy director
- Full professor
- Associate professor
- Head of department
- Service department director
- Research project leader
- Research project manager
- Principal investigator
- Other data support staff

### Profile result areas

Core activity	Result
Policy and strategy	<p><i>Design strategies for raising awareness of RDM policies and regulations</i></p> <ul style="list-style-type: none"> <li>• Identify gaps in research data infrastructure and tools provision and develop roadmaps to address these gaps</li> <li>• Participate in (inter)national developments of RDM policies and practices</li> <li>• Translate (inter)national developments into RDM policies and practices at the local level</li> <li>• Provide input to development of RDM policy and strategy</li> <li>• Raise awareness of RDM policy</li> <li>• Advise on adequate research data infrastructure and tools</li> <li>• Contribute to other relevant policies and regulations (strategic and tactical level)</li> </ul>
Compliance	<p><i>Advise on institutional compliance with RDM policies and regulations</i></p> <ul style="list-style-type: none"> <li>• Monitor compliance with RDM policies and regulations</li> <li>• Advise on compliance before, during and after a research project</li> <li>• Advise on how to deal with privacy sensitive data in accordance with appropriate guidelines</li> <li>• Work according to relevant laws and regulations</li> </ul>

Facilitating good RDM practices	<p><i>Advise relevant stakeholders on good practices of management of research data</i></p> <ul style="list-style-type: none"> <li>• Responsible for alignment of practices, infrastructure and tools with (inter)national standards and research needs/setting</li> <li>• Advise researchers and research support staff on careful management of research data in accordance with relevant regulations</li> <li>• Assist researchers in planning data management and writing data management plans</li> <li>• Assist researchers in making their data findable, accessible, interoperable and reusable</li> </ul>
RDM services	<p><i>Propose, implement and monitor RDM workflows and practices</i></p> <ul style="list-style-type: none"> <li>• Analyse the RDM skill needs, identify gaps and respond to these gaps</li> <li>• Develop and provide RDM training</li> <li>• Advise on and provide adequate RDM support</li> <li>• Collaborate with other departments involved in RDM services</li> <li>• Advise and assist researchers on careful management of research data and about procedures and technical aspects that are important for the quality and interoperability of (meta)data</li> </ul>
Data infrastructure	<p><i>Identify the requirements for adequate RDM infrastructure and tools</i></p> <ul style="list-style-type: none"> <li>• Care for structured and secure data storage in accordance with (inter)national principles</li> <li>• Model data structures and define and implement database needs</li> <li>• Work on complex data(processing) issues and provide input to improve processes or systems</li> </ul>
Knowledge management	<p><i>Determine the adequate level of RDM knowledge and skills</i></p> <ul style="list-style-type: none"> <li>• Determine the adequate level of RDM knowledge and skills in order to comply with relevant standards and policies and to facilitate good RDM practices</li> </ul>
Network and communication	<p><i>Create and participate in (inter)national RDM networks</i></p> <ul style="list-style-type: none"> <li>• Create, sustain and participate in a network of aligned RDM expertise locally, nationally and internationally</li> </ul>



	<ul style="list-style-type: none"> <li>• Liaise with diverse stakeholder groups and aligns RDM goals</li> </ul>
Data sharing and publishing	<p><i>Analyse gaps in support for data sharing and publishing</i></p> <ul style="list-style-type: none"> <li>• Analyse gaps in support for data sharing and publishing and develop strategies to address these gaps</li> <li>• Identify the various requirements for sharing, publishing and long-term archiving of data, within or outside an organisation in accordance with (inter)national principles</li> <li>• Assist in modelling and publishing data in accordance with (inter)national principles for findability, accessibility, interoperability and reusability</li> </ul>

Table Annex 5.1 Data steward profile result areas

### Profile soft skill competences

For data stewards, we consider the following soft skills competences essential. These are based on and a selection of the UFO job classification system competences.

Competence	Description
<i>Cluster: analysing and devising</i>	
Analytical capacity	Analysing situations or information and deciding what is of major and what is of lesser importance. Seeing interrelationships and getting to the core of the matter
Environment awareness	Showing that he or she is well informed about social, political and job-related developments. Using this knowledge effectively for the benefit of his or her own job or organisation
<i>Cluster: communicating and influencing</i>	
Presenting	Presenting ideas and information clearly, taking the target group into account
Persuasiveness	Succeeding in convincing others about ideas and plans
Networking skills	Creating and maintaining contacts both within and outside the organisation

Cooperating	Contributing with other persons or groups to a joint result, even when he or she has no personal interest in doing so
<i>Cluster: realising and evaluating</i>	
Accuracy	Acting precisely, carefully and impeccably when performing work
Initiative	Identifying problems and opportunities and taking appropriate action on own initiative
<i>Cluster: managing and supervising</i>	
Coaching	Assisting and encouraging staff or students so that they make the best possible use of their personal and professional qualities and develop their talents further in ways that benefit their careers
Managing for results	Directing and steering staff or a project group in order to achieve the desired objectives and results
<i>Cluster: personal effectiveness</i>	
Flexibility	Adapting to and capitalising on different people and new or changing circumstances
Integrity	Acting and complying with existing values, principles and rules to the best of their knowledge. Can be challenged about this and challenges others about it too

*Table Annex 5.2 Data steward profile soft skills competences*

### **Profile function level and salary scale**

The data steward job profile typically is a support staff function, and has four levels, i.e., junior, mediator, senior and expert, depending on education, experience and responsibilities. Consistent with the LCRDM data stewardship report, the data steward job profile ranges from salary scale 9 to 12 (according to the UFO classification scaling). The data steward function level and salary scale should match an employee's research experience, competences and educational background (such as a Bachelor, Master or PhD). As PhD graduates experienced in improving scientific data workflows are considered good candidates for a data steward job, it is important to make the data steward function attractive for this group of professionals, including recognition and remuneration for their expertise.

The following table gives an indication of the differences between the four data steward levels.

Level	Description
Junior <i>Scale 9</i>	<ul style="list-style-type: none"> <li>• Follows latest local, national and international RDM developments</li> <li>• Works within applicable regulations and supports creation of policies, so that the research work can be conducted efficiently</li> <li>• Ensures that the quality and interoperability of the data and software is sufficient for reuse</li> <li>• Hands-on: assists in implementing good data management practices across the entire research lifecycle</li> </ul>
Medior <i>Scale 10</i>	<ul style="list-style-type: none"> <li>• Further develops existing training and services</li> <li>• Participates in latest local, national and international RDM developments</li> <li>• Monitors researchers' needs on RDM tools and services</li> <li>• Contributes to policy development, so that the research work can be conducted efficiently and ensures that the quality and interoperability of the data is sufficient for reuse</li> <li>• Advises junior data stewards on various RDM aspects and advises researchers when needed</li> </ul>
Senior <i>Scale 11</i>	<ul style="list-style-type: none"> <li>• Analyses the needs for RDM</li> <li>• Identifies RDM support gaps</li> <li>• Develops innovative training and services</li> <li>• Participates in local, national and international developments on RDM</li> <li>• Acts as a liaison with leads of other research services</li> <li>• Advises policy makers</li> <li>• Advises on the implementation of RDM infrastructure so that the research work can be conducted efficiently and ensures that the quality and interoperability of the data is sufficient for reuse</li> <li>• Oversees research data workflows (from collection to publishing) and the use of (meta)data standards</li> </ul>
Expert <i>Scale 12</i>	<ul style="list-style-type: none"> <li>• Leads local, national and international developments on RDM</li> <li>• Ensures good RDM in the own organisation</li> <li>• Affects RDM policy in the own organisation</li> </ul>

*Table Annex 5.3 Differences between the four data steward levels*

Taking into account the detailed tasks of a data steward, this results in the following differences between the four data steward levels. This fine-grained analysis of the data steward job profile results in two extra core activities and results, namely *coordination of work* (i.e., lead, supervise and support less experienced colleagues) and *coaching and process improvement* (i.e., make proposals for improving work processes at different levels).

Function level <i>Scale</i>	1. Expert 12	2. Senior 11	3. Medior 10	4. Junior 9
Core activities and results				
<i>General</i>	Is competent in 2. Senior tasks	Is competent in 3. Medior tasks	Is competent in 4. Junior tasks	
<i>Policy and strategy</i>	Lead development of RDM policy and strategy	Design strategies for raising awareness of RDM policies and regulations  Identify gaps in data infrastructure and tools provision and develop roadmaps to address these gaps  Provide input to other relevant policies and regulations (strategic and tactical level)	Provide input to development of RDM policy and strategy  Raise awareness of RDM policy  Advise on adequate research data infrastructure and tools  Provide input to other relevant policies and regulations (practical level)  Translate (inter)national	Raise awareness of RDM policy  Contribute to translating (inter)national developments into RDM policies and practices at the local level

		Participate in (inter)national developments of RDM policies and practices	developments into RDM policies and practices at the local level	
<i>Compliance</i>	Responsible for institutional compliance framework with RDM policies and regulations	Advise on institutional compliance with RDM policies and regulations	Monitor compliance with RDM policies and regulations in research projects	<p>Advise on compliance before, during and after a research project</p> <p>Advise on how to deal with privacy sensitive data in accordance with the appropriate guidelines</p> <p>Work according to relevant laws and regulations</p>
<i>Facilitating good RDM practices</i>	<p>Advise on development of RDM facilities (infrastructure etc.) to comply with (inter)national standards and disciplinary demands</p> <p>Contribute to and lead the</p>	<p>Responsible for alignment of practices, infrastructure and tools with (inter)national standards and research needs/setting</p> <p>Participate in and provide advice on the</p>	Advise researchers and research support staff on careful management of research data according to (inter)national standards and research needs/setting	<p>Contribute to reproducibility and transparency (in relation to code of conduct for scientific integrity)</p> <p>Assist researchers in planning data management</p>

	development of international data and metadata standards	development of domain specific data and metadata standards that enable good RDM workflows		and writing data management plans  Assist researchers in various domains in making their data FAIR
<i>RDM services</i>	Lead innovation project of RDM service and training  Responsible for the RDM workflow overall	Analyse the RDM skill needs, identify gaps and respond to these needs and gaps by developing new and innovative training  Propose, implement and monitor workflows to improve RDM practice  Oversee execution of data collection, description, cleaning, modelling, merging, licensing, sharing, publishing and use of	Further develop existing training in the field of RDM  Advise on adequate RDM support  Assist in execution of data collection, description, cleaning, modelling, merging, licensing, sharing, publishing and use of (meta)data standards  Collaborate with other departments involved in RDM services	Provide training in the field of RDM  Assist in executing data collection, description, cleaning, modelling, merging, licensing, sharing, publishing and use of (meta)data standards  Advise researchers on careful management of research data and about procedures and technical aspects that are important for the

		(meta)data standards		<p>quality and interoperability of (meta)data</p> <p>Assist in making data actionable for computational methods relevant to the data</p> <p>Assist in discovering other people's data</p>
<i>Data infrastructure</i>	Strive to ensure compatibility of various RDM infrastructure elements used locally and (inter)nationally	Responsible for identifying the requirements of adequate data infrastructure and tools for RDM	<p>Monitor a process, system or the data architecture to optimise it</p> <p>Care for structured and secure data storage in accordance with (inter)national principles</p> <p>Model data structures and define database needs</p>	<p>Implement databases</p> <p>Work on complex data (processing) issues, providing input to improve processes or systems</p>
<i>Knowledge management</i>	Evaluate institutional RDM	Determine the adequate level	Determine the adequate level	Understand researchers and

	knowledge and skill level, propose RDM knowledge management roadmap or strategy	of RDM knowledge and skills in order to comply with (inter)national standards and policies	of RDM knowledge and skills in order to comply with the institute's RDM policy	research support staff's knowledge and skills regarding RDM, procedures and technical aspects that contribute to the quality of (meta)data
<i>Network and communication</i>	Create and sustain a network of aligned expertise with regard to RDM locally, nationally and internationally	Liaise with diverse stakeholder groups and align RDM goals	Participate in networks inside and outside the organisation	Participate in networks inside the organisation
<i>Data sharing and publishing</i>	Develop strategies to address gaps	Analyse gaps in support for data sharing and publishing	Identify the various requirements for sharing data, within or outside an organisation in accordance with (inter)national principles  Identify the various requirements for publishing data in accordance with	Assist in modelling and publishing data in accordance with (inter)national principles for findability, interoperability, accessibility and reusability



			(inter)national principles  Identify the various requirements for long-term archiving of data in accordance with (inter)national standards	
<i>Coordination of work and coaching</i>	Lead, supervise and support a number of less experienced colleagues	Coordinate the work of staff affiliated with a department or a project	Coordinate and/or organise selected elements of RDM services (e.g., training) and coordinates their execution	Provide professional support to colleagues
<i>Process improvement</i>	Make detailed proposals for improving work processes at strategic levels	Make detailed proposals for improving work processes at operational levels	Make proposals for improving own work processes	Identify improvement possibilities in own tasks

*Table Annex 5.4 Fine-grained analysis of the four data steward levels*

### Summary of core activities and results

For clarity reasons, we end this annex with a summary of the 11 relevant core activities and results for a data steward:

1. *Policy and strategy*: design strategies for raising awareness of RDM policies and regulations
2. *Compliance*: advise on institutional compliance with RDM policies and regulations
3. *Facilitating good RDM practices*: advise relevant stakeholders on good practices of management of research data

4. *RDM services*: propose, implement and monitor RDM workflows and practices
5. *Data infrastructure*: identify the requirements for adequate RDM infrastructure and tools
6. *Knowledge management*: determine the adequate level of RDM knowledge and skills
7. *Network and communication*: create and participate in (inter)national RDM networks
8. *Data sharing and publishing*: analyse gaps in support for data sharing and publishing
9. *Coordination of work*: lead, supervise and support less experienced colleagues
10. *Coaching and process improvement*: make proposals for improving work processes at different levels
11. *Soft skills*: this area comprises activities like accuracy and persuasiveness

## Annex 6: Good practices

In this annex, we present a good practice of three data steward function profiles as proposed for the FUWVAZ job classification system by a local organisation, the Radboudumc (translations into English is ours). These may inspire local and umbrella organisations when formulating job descriptions and job vacancies.

### Job description: data steward A

<p><b>Radboudumc</b> <i>Version 1.0, December 2019</i></p>
<p><b>Job description</b> <i>Job title:</i> Data steward A <i>Job family:</i> Data management <i>Job classification:</i> Scale 9</p>
<p><b>1. Purpose of the position</b></p>
<p>Monitoring the quality and FAIR aspects (findable, accessible, interoperable, reusable) of data within a research programme / project in relation to the research object and related research objects and the total research programme in order to make data reproducible, (re)usable and (long-term) available for (further) scientific research.</p>
<p><b>2. Result areas and responsibilities</b></p>
<p><i>a. Research data management</i></p> <ul style="list-style-type: none"> <li>• Supports researchers in drawing up a data management plan for the research programme or project</li> <li>• Develops templates for data collection</li> <li>• Promotes and supervises the collection of data and RDM in accordance with the RDM policy in the field</li> <li>• Identifies bottlenecks and gaps in RDM policy and takes action to resolve them</li> <li>• Advises, supports and issues guidelines to researchers regarding the findability, accessibility, usability and reusability (FAIR) of data within the research project or program</li> </ul>

*b. Infrastructure*

- Identifies the needs regarding the digital infrastructure for research data management within the project or program
- Ensures the availability of resources to promote the accessibility of the digital infrastructure
- Follows developments in the field of tools and infrastructure in the field of research data management
- Monitors security and compliance with GDPR and other relevant legislation

*c. Knowledge development regarding RDM*

- Monitors the level of knowledge regarding RDM and identifies shortcomings in knowledge and skills
- Is responsible for the development and availability of training courses for researchers
- Promotes awareness with regard to research data management and the added value that structured RDM has

*d. Data storage and archiving*

- Examines the policy within the project or programme regarding internal and external data collection and data storage
- Tests the actual data collection and storage within the project or programme against the formulated policy
- Determines whether the internal and external storage meets the applicable requirements

**3. Scope and authorisation**

- The officer is accountable to the line manager with regard to the quality of the work and the usefulness of advice regarding RDM and FAIR
- The most important frameworks are formed by the policy in the field of data research management within the field
- The officer makes decisions in supporting scientific researchers in drawing up a data management plan for the research programme / project, monitoring the level of knowledge regarding RDM and identifying shortcomings in knowledge and skills and determining whether the internal and external storage is adequate to the applicable requirements

<b>4. Contacts</b>
<ul style="list-style-type: none"> <li>• With researchers on how to apply research data management to exchange and coordinate information</li> <li>• With administrators and developers of the resources that promote the accessibility of digital infrastructure to tune</li> <li>• With RDM experts outside the project or programme to exchange information and knowledge</li> </ul>
<b>5. Expertise (knowledge and skills)</b>
<ul style="list-style-type: none"> <li>• Knowledge of the Radboudumc organisation</li> <li>• Knowledge of research data management methods and techniques within the scientific field</li> <li>• Skills in researching and testing used RDM methods</li> <li>• Skill in developing templates for RDM</li> <li>• Skills in applying and developing data FAIRification tools and semantic modelling of data</li> <li>• Advisory skills</li> </ul>

Table Annex 6.1 Good practice: data steward A

Job description: data steward B

<p><b>Radboudumc</b> Version 1.0, December 2019</p>
<p><b>Job description</b> <i>Job title:</i> Data steward B <i>Job family:</i> Data management <i>Job classification:</i> Scale 10</p>
<p><b>1. Purpose of the position</b></p> <p>Monitoring the quality and FAIR aspects (findable, accessible, interoperable, reusable) of data within related scientific research programs within a research institute in relation to the research discipline and related disciplines in order to meet the requirements and standards for research data management by Radboudumc posed</p>

## 2. Result areas and responsibilities

### *a. Research data management*

- Translates the research data management policy within the scientific research institute into research programs
- Advises managers of research programs on the collection and use of data and the FAIR aspects
- Advises, supports and issues guidelines to researchers regarding the findability, accessibility, usability and reusability (FAIR) of data within the research project / program
- Examines trends and developments in the field of RDM for relevance to the research programs
- Ensures that the RDM policy within the programs meets the standards set for quality and applicable codes of conduct, the FAIR standards as well as the legal and ethical standards
- Monitors the compliance of the RDM policy within the research programs with GDPR and other relevant laws and regulations

### *b. Infrastructure*

- Identifies the needs regarding the digital infrastructure for research data management of the research programs
- Ensures the availability of resources to promote the accessibility of the digital infrastructure
- Follows developments of tools and infrastructure in the field of research data management
- Monitors security and compliance with GDPR and other relevant legislation

### *c. Knowledge development regarding RDM*

- Monitors the level of knowledge regarding RDM and identifies shortcomings in knowledge and skills
- Is responsible for the development and availability of training courses for researchers
- Promotes awareness with regard to RDM and the added value that structured RDM has

### *d. Data storage and archiving*

- Examines the policy within the research programs regarding internal and external data collection and data storage

<ul style="list-style-type: none"> <li>• Tests the actual data collection and storage within programs against the formulated policy</li> <li>• Determines whether the internal and external storage meets the applicable requirements</li> </ul>
<p><b>3. Scope and authorisation</b></p>
<ul style="list-style-type: none"> <li>• The officer is accountable to the manager with regard to the usability of the translated RDM policy</li> <li>• The most important frameworks are formed by Radboudumc's RDM policy in the field of data research management as well as applicable laws and regulations</li> <li>• The officer makes decisions when translating Radboudumc's RDM policy for research programs</li> </ul>
<p><b>4. Contacts</b></p>
<ul style="list-style-type: none"> <li>• With managers of the research programs on the applicable RDM policy to coordinate</li> <li>• With administrators and developers of the resources that promote the accessibility of digital infrastructure to tune</li> <li>• With RDM experts outside the programme and the research institute to exchange information and knowledge</li> </ul>
<p><b>5. Expertise (knowledge and skills)</b></p>
<ul style="list-style-type: none"> <li>• Knowledge of the Radboudumc organisation</li> <li>• Knowledge of RDM methods and techniques within the scientific field</li> <li>• Skills in applying and developing data FAIRification tools and semantic modelling of data</li> <li>• Skills in researching and testing used RDM methods</li> <li>• Skills in translating policy for RDM</li> <li>• Advisory skills</li> </ul>

*Table Annex 6.2 Good practice: data steward B*

Job description: data steward C

## Radboudumc

Version 1.0, December 2019

Job description

*Job title:* Data steward C

*Job family:* Data management

*Job classification:* Scale 11

### 1. Purpose of the position

Monitoring the quality and FAIR aspects (findable, accessible, interoperable, reusable) of data within a scientific research institute in relation to the research field and related disciplines and other research institutes in order to meet the requirements and standards that apply to research data management from the scientific field as well as by Radboudumc.

### 2. Result areas and responsibilities

#### *a. Research data management*

- Develops policy regarding research data management within the scientific research institute
- Advises the management of the scientific institute or management within the field
- Identifies trends and developments in the field of RDM in relation to the scientific discipline
- Ensures that the RDM policy within the institute meets the standards set for quality and applicable codes of conduct, the FAIR standards as well as the legal and ethical standards
- Monitors the compliance of the RDM policy within the research institute with GDPR legislation and regulations
- Advises management and researchers on the application of the RDM policy within the research institute

#### *b. Infrastructure*

- Makes an inventory of the needs concerning the digital infrastructure for research data management within the research institute / scientific discipline
- Ensures the availability of resources to promote the accessibility of the digital infrastructure



- Follows developments of tools and infrastructure in the field of research data management

*c. Knowledge development regarding RDM*

- Monitors the level of knowledge regarding RDM and identifies shortcomings in knowledge and skills
- Is responsible for the development and availability of training courses for researchers
- Promotes awareness with regard to research data management and the added value that structured RDM has

*d. Data storage and archiving*

- Examines the policy within the scientific institute regarding internal and external data collection and data storage
- Tests the actual data collection and storage within the scientific institute against the formulated policy
- Determines whether the internal and external storage meets the applicable requirements

**3. Scope and authorisation**

- The officer is accountable to the manager with regard to the usability of the developed RDM policy
- The most important frameworks are formed by the policy objectives of Radboudumc in the field of data research management as well as the applicable laws and regulations
- The officer makes decisions in the development and implementation of RDM policy for the research institute

**4. Contacts**

- With the management of the research institute about developing policy to create and coordinate support
- With administrators and developers of the resources that promote the accessibility of digital infrastructure to tune
- With RDM experts outside the research institute to exchange information and knowledge

### 5. Expertise (knowledge and skills)

- Knowledge of the Radboudumc organisation
- Broad theoretical knowledge of research data management methods and techniques within the scientific field
- Skills in researching and testing used RDM methods
- Skills in developing policy for RDM
- Advisory skills

Table Annex 6.3 Good practice: data steward C

Similar to the previous annex, in the following table the three job profiles are summarised, placed next to each other and compared.

Function level <i>Scale</i> Function elements	1. Senior <i>11</i>	2. Medior <i>10</i>	3. Junior <i>9</i>
1. Goal of the function	Monitor quality and FAIR aspects of data in the research institute	Monitor quality and FAIR aspects of data in a research program	Monitor quality and FAIR aspects of data in a research programme / project
2a. Tasks and responsibilities: RDM	<p>Develop RDM policy</p> <p>Advisor of management at the institute regarding FAIR and policy</p> <p>Identify RDM trends and developments in relation to the research area</p> <p>Ensure that RDM policy within institute is GDPR compliant</p>	<p>Implement RDM policy</p> <p>Advisor of managers of research programs regarding FAIR</p> <p>Investigate RDM trends in relation to the research programs</p> <p>Ensure that RDM policy within research programme is GDPR</p>	<p>Identify and solve issues related to RDM policy</p> <p>Advisor of researchers regarding FAIR</p> <p>Develop templates for gathering of data</p> <p>Advise researchers when they are writing a DMP</p>

		compliant	Monitors RDM safety and GDPR compliance
2b. Tasks and responsibilities: infrastructure	<p>Make an inventory of digital infrastructure needs for RDM within the institute</p> <p>Ensure availability of digital infrastructures</p> <p>Follow developments in tools and infrastructure</p>	<p>Make an inventory of digital infrastructure needs for RDM within the research programs</p> <p>Ensure availability of digital infrastructures</p> <p>Follow developments in tools and infrastructure</p>	<p>Make an inventory of digital infrastructure needs for RDM within the project or program</p> <p>Ensure availability of digital infrastructures</p> <p>Follow developments in tools and infrastructure</p>
2c. Tasks and responsibilities: development of RDM knowledge	<p>Monitor RDM knowledge level and identify gaps</p> <p>Ensure training for researchers is developed and remains available</p> <p>Create RDM awareness</p>	<p>Monitor RDM knowledge level and identify gaps</p> <p>Ensure training for researchers is developed and remains available</p> <p>Create RDM awareness</p>	<p>Monitor RDM knowledge level and identify gaps</p> <p>Ensures training for researchers is developed and remains available</p> <p>Create RDM awareness</p>
3. Data storage and archiving	<p>Investigate policy within institute regarding data gathering and storage</p> <p>Assess whether data gathering and storage within</p>	<p>Investigate policy within programme regarding data gathering and storage</p> <p>Assess whether data gathering and storage within</p>	<p>Investigate policy within project / programme regarding data gathering and storage</p> <p>Assess whether data gathering and storage within</p>

	institute comply to policy	programme comply to policy	program/project comply to policy
4. Authorisations	Make decisions to develop and implement RDM policy in the research institute	Make decisions to translate RDM policy of the UMC to research programs	Make decisions to assist researchers
5. Contacts	Management of institute to develop policy  ICT department for digital infrastructure  RDM experts inside and outside the institute	Managers of research programs  ICT department for digital infrastructure  RDM experts outside the programme and institute	Researchers  ICT department for digital infrastructure  RDM experts outside the project or programme
6. Knowledge and skills	RMD knowledge related to the research area  Skills to investigate and assess RDM methods used at the UMC  Policy developing skills  Consultancy skills	RMD knowledge related to the research area  Skills to investigate and assess RDM methods used at the UMC  Skills to apply and develop FAIRification tools  Policy translation skills  Consultancy skills	RMD knowledge related to the research area  Skills to investigate and assess RDM methods used at the UMC  Skills to apply and develop FAIRification tools  Skills to develop RDM templates  Consultancy skills

*Table Annex 6.4 Good practice: summary and comparison of the different data stewards*

## Annex 7: UAS data steward job profile

### Challenges

More and more practice-oriented research is being conducted within universities of applied sciences (UASs). Such research aims at answering issues in society with the intention to make knowledge flow directly back to professional practice and education (Andriessen, 2014)<sup>124</sup>. In contrast to research universities, research at UAS has only been part of the activities for about 15 years and education is still the core activity.

The aim of creating a data steward job profile is to increase the quality of research data management (RDM) and thus to comply with the national legislation and regulations of the Netherlands Code of Conduct for Scientific Integrity (NGWI, 2018)<sup>125</sup> and the requirements of grant providers such as ZonMw. In addition, this job profile ensures that HR departments of higher professional education institutions are provided with clear guidance to position the data steward job in their organisation and to recruit employees with the right competences. It also ensures that the work of data stewards is recognised and appreciated. The job description provides attractive terms of employment so that the UASs can compete with the research universities, institutes and the commercial business community, which are also fully engaged in fulfilling the position of the data steward.

There are a number of challenges for universities of applied sciences when it comes to the position of data steward:

1. More requirements for practice-oriented research
2. National cooperation between universities of applied sciences
3. Uncertainty about the position of data stewards within higher professional education

#### *1. More requirements for practice-oriented research*

More and more demands are made on practice-oriented research. These requirements in the field of research and research support follow from the NGWI, the General Data Protection Regulation (GDPR), the FAIR principles for research data (Findable, Accessible, Interoperable and Reusable), Open Science and the Branch Protocol Quality Assurance Research 2016-

---

<sup>124</sup> Andriessen, D. (2014). *Praktisch relevant en methodisch grondig; dimensies van onderzoek in het hbo*. Utrecht. Hogeschool Utrecht.

<sup>125</sup> Nederlandse gedragscode wetenschappelijke integriteit (2018). <https://doi.org/10.17026/dans-2cj-nvwu>

2022 (2015)<sup>126</sup>. Due to these developments, it is currently required that RDM is implemented in a structured and professional manner so that data is and remains available for future investigations and integrity issues. In addition, data management plans must be drawn up in consultation with a data steward, according to the format of grant providers. A project can only start once the DMP has been approved and the support of a data steward is one of the requirements for approval.

Various universities of applied sciences have started adopting RDM, often after the NGWI was adopted in 2018. Data stewards with expertise in data management are needed to support RDM in accordance with the guidelines of the NGWI and the subsidy providers. A researcher must be supported in the field of RDM. The quality requirements of data management, making data FAIR and writing a data management plan have increased. The researcher has too little time and expertise in the field of RDM to be able to meet these conditions independently. National and international developments in the field of data and RDM are fast and a researcher cannot be expected to keep up with these developments in addition to his or her research and teaching tasks. It also concerns specific expertise that a researcher often uses only to a limited extent to gain routine.

At the same time, practice-oriented research within universities of applied sciences has increased, both in quality and quantity. More project applications have been accepted by subsidy providers and these projects must be supported properly. Since 2015, the subsidies awarded have quadrupled (figures according to the subsidy departments of UAS Avans, HAN and Saxion). This increase has shown that the activities cannot be added to an existing position, but that an independent position of data steward is needed to support all activities related to RDM.

The role of data steward is already performed within universities of applied sciences via existing positions, often in combination with other tasks. Given the complexity in combination with the quality requirements for RDM and the ever-increasing number of studies as mentioned above, it is therefore necessary to create an independent job profile specifically for data steward. Currently, there is a momentum in this area at universities of applied sciences. For example, an employee at Saxion has been recently released completely to focus on data steward tasks. However, this has not yet been laid down in an actual data steward position. Recently,

---

<sup>126</sup> Vereniging Hogescholen (2019). [Brancherapport praktijkgericht onderzoek 2018](https://www.vereniginghogescholen.nl/system/knowledge_base/attachments/files/000/001/082/original/Brancherapport_Praktijkgericht_Onderzoek_2018_-_definitieve_webversie.pdf?1573652096): Feiten, cijfers en ontwikkelingen. Accessed on 23 September 2020 via [https://www.vereniginghogescholen.nl/system/knowledge\\_base/attachments/files/000/001/082/original/Brancherapport\\_Praktijkgericht\\_Onderzoek\\_2018\\_-\\_definitieve\\_webversie.pdf?1573652096](https://www.vereniginghogescholen.nl/system/knowledge_base/attachments/files/000/001/082/original/Brancherapport_Praktijkgericht_Onderzoek_2018_-_definitieve_webversie.pdf?1573652096)

Rotterdam University of Applied Sciences placed the following vacancy for data stewards with associated activities and job requirements<sup>127</sup> (translation made by the project).

Data stewards (1.6-2.0 FTE)

Rotterdam University of Applied Sciences, Rotterdam

Ondersteuning Kenniscentra en Expertisecentra (OKC) is immediately looking for data stewards. Rotterdam University of Applied Sciences focuses its research in six knowledge centres and two Centres of Expertise (COE). Several lecturers and (teacher) researchers who conduct practice-oriented research are active within these knowledge centres and COEs. OKC is currently developing a range of services to support the knowledge centres and COEs. During the preparation, implementation and completion of research projects, these services provide advice and support in obtaining funding, legal issues and data management.

#### *Position*

As a data steward, you support researchers with project and grant applications by writing a data paragraph and data management plan with which we follow laws and regulations and comply with the FAIR data principles. You advise researchers on the careful handling of research data, and on procedures and technical aspects that are important for the quality of the datasets. You approach researchers proactively and transfer your knowledge, which demonstrates the added value of data management. You work together with the consultants of the library who support research within HR for the publication of articles and the use/reuse of research data and information. You are involved in the entire research process from project idea, during research up to and including completion.

The team in which you work consists of data stewards and advisors from the library. You work partly from the OKC back office, but mainly at the knowledge centres and COEs on various locations and in continuous collaboration with your team members.

#### *Tasks and responsibilities*

- Supports writing data paragraphs and data management plans per project
- Advises in the field of data management and systems
- Ensures the authorisation of the digital work environment of research projects and the use of the prescribed digital infrastructures
- Organises the internal and external archiving of the data in accordance with the FAIR

<sup>127</sup> Hogeschool Rotterdam (2020, 20 juli). Data steward. Accessed on 21 July 2020 via <https://www.linkedin.com/jobs/view/1959886342/?refId=498bfd5-53fd-415f-8831-553690fb1ade>

principles

- Monitors the implementation and execution of the GDPR guidelines within the projects
- Monitors the data management processes and suggests improvements
- Develops and provides training in research data management
- Maintains internal and external networks in the field of RDM
- Follows training regarding the RDM profile

#### *Profile*

We are looking for an enthusiastic colleague who is able to bring together and connect different disciplines:

- UAS or university level working and thinking
- Affinity with data, research methods and ICT
- Knowledge of RDM and GDPR guidelines
- Personal competences important for this position are planning and organising, persuasiveness, network skills and quality-orientation
- Able to work independently

#### *Offer*

Rotterdam University of Applied Sciences offers you a temporary employment contract for one year with an option to extend (D4 CAO-HBO) in scale 9. The salary is a minimum of € 2,845.48 and a maximum of € 3,920.50 gross per month for full-time employment (excluding 8% holiday pay and 8.3% year-end bonus). The size of the appointment can be adjusted in further consultation and is between 0.8 and 1 FTE. Our secondary employment conditions include excellent training opportunities and an attractive pension scheme. To be able to work at Rotterdam University of Applied Sciences, you need a Certificate of Good Conduct (Dutch: VOG).

#### *Information*

For more substantive information about this position and the selection procedure, please contact [...]. Closing date: 30 August 2020

## *2. National cooperation between universities of applied sciences*

In order to take research within universities of applied sciences to a higher level, the Reinforcement Agenda for Practice-oriented Research was set up in 2018<sup>128</sup>. Fourteen

<sup>128</sup> Surf. Versterkingsagenda praktijkgericht onderzoek. Accessed on 23 September 2020 via <https://www.surf.nl/versterkingsagenda-praktijkgericht-onderzoek>



universities of applied sciences have joined this Reinforcement Agenda. They work together to make use of joint knowledge when setting up and organising research support.

A concrete result of this collaboration is the National Integrated Research Support Model, LIOM (Surf, 2020)<sup>129</sup>. This model offers tools for universities of applied sciences to get started with research support in their own organisation and to adapt it to their own situation.

Another result of the Reinforcement Agenda is the project to develop a national digital competence centre (DCCs) for UASs. "The goal is an integrated RDM expertise and service offering that optimally meets the actual needs of researchers in practice-oriented research" (Core Group Reinforcement Agenda Practice-oriented Research, 2020)<sup>130</sup>. This means that the DCC will become a national knowledge and advice centre for universities of applied sciences to reach the Open Science ambition. Topics such as FAIR data, Open Data, big data, data stewardship, training and ICT infrastructure will be supported. The contacts in the DCCs will create a data steward network. A uniformed job description is essential in order to share knowledge and collaborate, for example on joint research projects.

Furthermore, a national, unambiguous description of the position of data steward at universities of applied sciences is necessary to enable mobility between universities of applied sciences. Uniformity between UAS will stimulate employee exchanges and mobility.

### *3. Uncertainty about the position of data stewards in higher professional education*

As described earlier, there is still no general job profile for the data steward within higher professional education institutions. As a result, there is still a lot of uncertainty at universities of applied sciences about the position, activities and expectations of a data steward. Currently, the task of data steward in higher professional education is often fulfilled by information specialists from the library or the information centre because they have a good basis for data management from their expertise in information management and knowledge management. The duties of a data steward are often seen as a role in an existing position. However, data steward tasks can no longer be combined with other tasks and require specific knowledge and unconditional employability. It is therefore important to record these tasks in a separate job profile, because existing job profiles do not describe the specific tasks. They do not correspond to the work that is carried out and the value thereof.

---

<sup>129</sup> Surf (2020). Landelijk integraal onderzoeksondersteuningmodel voor hogescholen. Accessed on 24 September 2020 via [https://www.surf.nl/files/2020-03/landelijk-integraal-onderzoeksondersteuningmodel-hogescholen\\_0.pdf](https://www.surf.nl/files/2020-03/landelijk-integraal-onderzoeksondersteuningmodel-hogescholen_0.pdf)


<sup>130</sup> Kerngroep Versterkingsagenda Praktijkgericht onderzoek (2020). Aanvraag voor financiering van een landelijk Digital Competence Center voor het praktijkgericht onderzoek.

For universities of applied sciences, a job profile should provide clarity about the knowledge areas, competences and skills of a data steward. This makes it clear under which conditions the applicant should be able to perform the position and which specific trainings and experiences are required. This improves insight into professionalisation opportunities, assessment accuracies and appropriate rewards.

### Job description of a data steward in higher professional education

Each university of applied sciences in the Netherlands is free to use its own job classification system. That is why the activities and competences in the job profile (see below) are described as exhaustively as possible so that they are suitable for different sizes and the institutions can make their own choice, in line with the organisational culture.

<i>Profile title</i>	Data steward
<i>Profile objective</i>	<ul style="list-style-type: none"> <li>• The data steward is the expert in the field of research data management and advises and supports the researcher</li> <li>• The ranking of the position depends on the weight of the responsibility and tasks of the position</li> <li>• The data steward is the linking pin between researchers, research services, faculty members, external parties, and specialists from other departments etc.</li> <li>• Depending on the position ranking, the data steward develops or gives trainings, participates in latest developments on RDM, related to local, national and international developments, knows about the FAIR principles, knows about and works within applicable law and regulations and delivers support and advice for policy makers</li> <li>• This way the data steward facilitates that research is properly conducted</li> </ul>
<i>Profile context</i>	The data steward reports to and/or receives hierarchical guidelines from the head of the library or head of (a department similar to) research services. Additionally, the data steward has intensive contacts with various internal departments as the task requires

<p><i>Organisati on</i></p>				
<p><i>Profile result areas</i></p>	<p>Core activity Result</p>	<p>Data steward 1 (policy tasks)</p>	<p>Data steward 2 (coordinati ng tasks)</p>	<p>Data steward 3 (executive tasks)</p>
<p><i>Policy and strategy</i></p>				
<p>Provides input for the development of RDM policy and strategy</p>		<p>✓</p>		
<p>Responsible for knowledge of RDM policy, FAIR principles, Open</p>		<p>✓</p>		
<p>Science and regulations towards policymakers</p>		<p>✓</p>	<p>✓</p>	<p>✓</p>
<p>Contributes to advice for researchers on adequate research data infrastructure and tools</p>		<p>✓</p>		
<p>Contributes to input on other relevant policies and regulations (strategic and tactical level)</p>		<p>✓</p>		

Translates (inter)national developments into RDM policies and practices at the university of applied sciences	✓		
<i>Compliance</i>			
Advices on compliance on RDM requirements of funders according to the Netherlands Code of Conduct of Academic Practice	✓	✓	✓
Able to recognise potential problems concerning compliance issues before, during and after a research project on subject areas such as: intellectual property rights and ownership of the data in agreements made with external parties (such as consortia agreements) and refer to expert when necessary	✓	✓	
Advices about privacy sensitive data in accordance with the appropriate guidelines, like the GDPR. Or can refer researchers to the GDPR/privacy expert and ethical committee	✓	✓	
Is familiar with, advises on and works according to relevant laws and regulations, including legal and ethical standards	✓	✓	✓
<i>Facilitating good RDM practices</i>			
Advices researchers and research support staff on careful management of research data according to (inter)national standards	✓	✓	✓
Advices to maintain reproducibility and transparency (according to the	✓	✓	✓

	Netherlands Code of Conduct for Academic Practice)			
	Advises researchers in planning data management	✓	✓	
	Supports researchers writing data management plans	✓	✓	
	Advises researcher in making their data findable, accessible, interoperable and reusable (FAIR principles)	✓	✓	✓
<i>RDM services</i>				
	Advises on adequate RDM support to stakeholders and policy makers in the institution	✓		
	Advises on data collection, storage, licensing, sharing, publishing and use of (meta)data standards	✓	✓	✓
	Collaborates with other departments involved in research services	✓	✓	
	Advises researchers on careful management of research data concerning version control, structuring data etc.	✓	✓	
	Advises on access rights to (sensitive) data	✓	✓	
	Advises about procedure and technical aspects for the quality of (meta)data according to the FAIR principles	✓	✓	✓
	Assists in discovering other researcher's data	✓	✓	✓

<i>Data infrastructure</i>			
Cares for structured and secure data storage in accordance with (inter)national principles	✓		
Provides input to improve processes, systems and tools	✓	✓	✓
Advises on using secure infrastructure, compliant with laws and regulations	✓	✓	✓
<i>Knowledge management</i>			
Determines the adequate level of RDM knowledge and skills of researchers in order to comply with the institute's RDM policy	✓		
Develops adequate training for researchers	✓		
Determines the adequate level of RDM knowledge and skills of the data stewards in order to comply with the institute's RDM policy	✓		
Keeps knowledge and skills up to date according to (inter)national standards	✓	✓	✓
Informs researchers about RDM, Open Science, FAIR principles and other relevant information	✓	✓	✓
<i>Network and communication</i>			
Is the linking pin between researchers, research services, colleagues at faculty, external parties, specialists from other departments etc.	✓		

	Takes the initiative to establish and support contacts between partners internal and external in order to consult on subjects from the data management field	✓		
	Maintaining internal and external contacts	✓		
	Participates in networks inside and outside the organisation	✓		
	Promotes RDM, FAIR principles and Open Science	✓	✓	✓
<i>Data sharing and publishing</i>				
	Identifies the various requirements for sharing data and software, within or outside an organisation for example licensing and repositories	✓	✓	
	Identifies the various requirements for FAIR and long-term archiving of data and software	✓	✓	
	Advises in publishing data, software and research output in accordance with principles of Findability, Accessibility Interoperability and Reusability	✓	✓	✓
	Advises on adequate metadata standards	✓	✓	

<p><i>Soft skills</i></p>	<p><u>Analysing and devising</u></p> <ul style="list-style-type: none"> <li>• <i>Analytical capacity</i>: analysing situations or information and deciding what is of major and what is of lesser importance. Seeing interrelationships and getting to the core of the matter</li> <li>• <i>Problem analysis</i>: an investigation of the causes of an incident. This is done to identify and make improvements to processes and procedures</li> <li>• <i>Knowledge awareness</i>: using knowledge effectively for the benefit of the job and organisation. In case of knowledge deficit: being able to gain and apply new knowledge</li> <li>• <i>Growth mindset</i>: believing talents can be developed through new insights, good strategies and input from others</li> </ul> <p><u>Communicating and influencing</u></p> <ul style="list-style-type: none"> <li>• <i>Presenting</i>: presenting ideas and information clearly, taking the target group into account</li> <li>• <i>Organisational sensitivity</i>: recognising the influence and consequences that one's own decisions or actions have on the organisation and acting accordingly</li> <li>• <i>Persuasiveness</i>: succeeding in convincing others about ideas and plans</li> <li>• <i>Networking skills</i>: creating and maintaining contacts both within and outside the organisation</li> <li>• <i>Client oriented</i>: focusing on helping customers meet their goals.</li> <li>• <i>Team player</i>: contributing with other persons or groups to a joint result, even when there is no personal interest in doing so</li> <li>• <i>Question articulation</i>: giving shape or expression to a question</li> <li>• <i>Training in data information literacy</i>: designing and implementing programs to teach data management and information skills</li> </ul> <p><u>Realising and evaluating</u></p> <ul style="list-style-type: none"> <li>• <i>Accuracy</i>: acting precisely, carefully and impeccably when performing work</li> <li>• <i>Planning and organising</i>: overseeing activities, setting goals and priorities, plan activities, time and resources</li> <li>• <i>Initiative</i>: identifying problems and opportunities and taking appropriate action on own initiative</li> </ul>
---------------------------	--



	<ul style="list-style-type: none"> <li>• <i>Decisive</i>: having the power or quality of making an individual decision</li> </ul> <p><u>Managing and supervising</u></p> <ul style="list-style-type: none"> <li>• <i>Managing for results</i>: directing and steering staff or a project group in order to achieve the desired objectives and results</li> <li>• <i>Coaching</i>: assisting and encouraging staff or students so that they make the best possible use of their personal and professional qualities and develop their talents further in ways that benefit their careers</li> </ul> <p><u>Personal effectiveness</u></p> <ul style="list-style-type: none"> <li>• <i>Flexibility</i>: adapting to and capitalising on different people and new or changing circumstances</li> <li>• <i>Integrity</i>: acting and complying with existing values, principles and rules to the best of their knowledge. Can be challenged about this and challenges others about it too</li> <li>• <i>Curiosity</i>: desire to find out about (new) things</li> </ul>
<i>Contacts</i>	<ul style="list-style-type: none"> <li>• With all expertise services that are part of the Research Services and the researchers, lecturers and project employees and/or supporters as shown in the figure above.</li> <li>• Contacts within national relevant networks</li> <li>• With data stewards from other universities of applied sciences and the DCC HBO</li> </ul>
<i>Knowledge and experience</i>	<ul style="list-style-type: none"> <li>• Knowledge of research (research knowledge)</li> <li>• Knowledge of and skills to use (meta)data standards</li> <li>• Skills in collecting data and literature</li> <li>• Proficient in the Dutch and English language</li> <li>• Data information literacy</li> </ul>

Table Annex 7.1 UASs job profile activities and competences

## Recommendations

### 1. Create a national description for the position of data steward

The position of data steward cannot be combined with other tasks and roles. It is a specific expertise, and the tasks must be performed with focus and sufficient time. The amount of

research and the growth of datasets is rapidly increasing and the RDM requirements are getting stricter. Researchers and project managers require high-quality and fast advice to be able to continue with their research (application). The allocation of subsidy funds depends on the presence and advice of a data steward. In short, the tasks of a data steward are so important in practice-oriented research that a separate position is justified.

### *2. Include the job description in the local job framework and in the national job classification systems*

The number of subsidised research projects has increased about four times. The datasets from research are constantly growing. This ensures that data steward support can no longer be performed in combination with other activities. The requirements and conditions of subsidy providers have also changed. The activities are also carried out for the entire organisation, including professorships. It is therefore necessary to have the position of data steward included by the HR department in the job framework of a local educational institution.

The inclusion of the position of data steward in the job framework of the universities of applied sciences is necessary in order to jointly professionalise practice-oriented research and research support. If the job data steward is included as uniformed as possible in the job framework of the universities of applied sciences, this will ensure that all institutions fulfil the job in the same way. This will stimulate knowledge exchange between universities of applied sciences.

### *3. Make the position of the data steward in higher professional education interesting and challenging*

Trained data stewards are still scarce, in the Netherlands and Europe. There is a lack of experts with data stewardship skills. Recognising this field of expertise will promote good research practice. The report of the High Level Expert Group on the European Open Science Cloud (European Union, 2016)<sup>131</sup> by the European Commission indicates that in the short term 500.000 data experts are needed for research data management in academic Europe (in Verheul et al., 2019)<sup>132</sup>.

---

<sup>131</sup> European Union, (2016). Realising the European Open Science Cloud First report and recommendations of the Commission High Level Expert Group on the European Open Science Cloud. <http://doi.org/10.2777/940154>

<sup>132</sup> Verheul, I., Imming, M., Ringersma, J., Mordant, A., Van der Ploeg, J., & Pronk, M. (2019). Datastewardship op de kaart: Een verkenning van taken en rollen in Nederlandse onderzoeksinstituten. Zenodo. <http://doi.org/10.5281/zenodo.2642066>

In order to attract the right people for the position of data steward in higher professional education and to retain talent, it is necessary to make the profession attractive. Sufficient challenge could be created by offering career opportunities (different rankings) and placing the position in an attractive salary scale that recognises the context of the work, the research scope and offers room to be a decision-making partner at a level (minimum salary scale 9 for junior data steward, increasing according to experience and seniority to scale 11, according to the CAO-HBO profile).

#### *4. Provide sufficient training and education opportunities*

The field of data stewardship is developing rapidly both nationally and internationally. In order to keep up with the topic and to maintain quality, it is essential to provide sufficient training and education opportunities for data stewards. In addition to the tasks described earlier, the data steward must have the opportunity to familiarise himself/herself with the position and to receive regular training.

#### *5. Place data stewards in the organisation in the library or information centre*

Within the organisational structure of a university of applied sciences, it is preferred to place the data stewards in a team of research services as part of the library or information centre. Historically, libraries and information centres are ideally suited to provide the right support throughout the entire research lifecycle, for example metadata (describing data), making data accessible and thus findable, and collecting Open Data and literature. A UAS library is experienced in supporting educational activities broadly of all study programs. An information specialist who is already used to building networks within academies and can establish contacts with researchers in the same way.

Research services should be seen as a network organisation that can use the expertise of other stakeholders such as subsidy management, ethics committee, security officer, quality assurance agency, etc. (see the figure above).

#### *6. Provide a data analyst and research software engineer within Research Services*

In order to be able to perform all complex tasks within Research Services, a data analyst and a research software engineer will also have to be employed, in addition to a data steward. The data analyst supports the researchers in processing and analysing the data, and a research software engineer supports the researchers with the development of software and the application of adequate technologies. These tasks require different knowledge and competences than those of a data steward.



The collaboration of several disciplines within Research Services will give the advantage of broad support to specific activities that are rapidly developing in different research areas. While the data steward advises the researchers throughout the entire research cycle, the data analyst and research software engineer will support the researchers during the processing of research data. The data analyst and software engineers should have in-depth knowledge in their fields and are able to cooperate with the research topics, whether or not embedded, so that the expertise does not have to be purchased externally. Due to the increase in data-intensive research, the demand for the expertise of a data analyst and a research software engineer will increase and their deployment will also have to be facilitated. It is essential to distinguish the roles of data steward, data analyst and research software engineer.

## Annex 8: RDM trainings

### Trainings about research data management

Taking stock of training resources for (aspiring) data stewards has some tradition, both in the Netherlands and abroad. The current overview benefits from several other initiatives<sup>133</sup>. For the purpose of this project, we used the following criteria:

- Recency: trainings about research data management (RDM) that were developed or substantially revised after the introduction of the FAIR principles (2014) and the implementation of the General Data Protection Regulation (GDPR).
- Accessibility: preferably open training, although we include some references to illustrate that several universities provide trainings for affiliated staff.
- Focus: preferably stand-alone trainings. However, we include a short list of good-quality materials that could be used in blended learning with a proper pedagogical embedding.
- No informal learning: informal learning, for example through networking, is essential, but not the goal of this exercise.

The following trainings, listed by provider in alphabetical order, address major parts of what data stewards should know. Some focus on specific research domains. Trainings take the form of slide decks, videos, websites with assignments, and Massive Open Online Course (MOOC).

- CESSDA Data Management Expert Guide: <https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide> (targets social science researchers)
- DARIAH training: <https://campus.dariah.eu/source/events> (targets arts and humanities researchers)
- Data Carpentry: domain-specific workshops for self-guided learning (or training): <https://datacarpentry.org/lessons>
- Digital Preservation Coalition (DPC). Novice to Know-How: Online Digital Preservation Training: <https://www.dpconline.org/knowledge-base/training/n2kh-online-training>

---

<sup>133</sup> Several overviews were used:

Scholtens, S. et al., (2019). Towards FAIR data steward as profession for the lifesciences. <https://zenodo.org/record/3474789> (Section 7).

SSHOC (2020): <https://training-toolkit.sshopencloud.eu/>

UKB Working Group (2017): [https://dataverse.nl/dataverse/RDM\\_Courses\\_NL](https://dataverse.nl/dataverse/RDM_Courses_NL)

LCRDM (no date): <https://www.lcrdm.nl/trainingen>

- DTL Helix Academy Data Stewardship Course: <https://www.helixacademy.com/en/data-analysis-stewardship>
- ELIXIR-NL Technical Data Stewardship Course: <https://www.dtls.nl/courses/elixir-nl-training-technical-data-stewardship>
- EMBL-EBI: Bringing data to life: data management for the biomolecular sciences: <https://www.ebi.ac.uk/training/online/course/bringing-data-life-data-management-biomolecular-sciences>
- EUDAT services training: <https://github.com/EUDAT-Training>
- FOSTER. Assessing the FAIRness of data: <https://www.fosteropenscience.eu/node/2644>
- FOSTER. Managing and Sharing Research Data: <https://www.fosteropenscience.eu/node/2328>
- FOSTER. Data Protection and Ethics: <https://www.fosteropenscience.eu/node/2330>
- GO-opleidingen. Data-architectuur en datamanagement: <https://goopleidingen.nl/module/data-architectuur-en-datamanagement>
- Maastricht university. EU General Data Protection Regulation Essentials & EU General Data Protection Regulation: <https://www.maastrichtuniversity.nl/research/institutes/ecpc/professional-certification-education>
- PARTHENOS. Manage, Improve and Open up your Research and Data: <https://training.parthenos-project.eu/sample-page/manage-improve-and-open-up-your-research-and-data>
- RDMLA: <https://rdmla.github.io> (for librarians and information professionals)
- RDNL. Essentials 4 Data Support: <https://datasupport.researchdata.nl/en/> (introductory course for data supporters)
- RDNL/DCC. Delivering Research Data Management Services (MOOC): <https://www.futurelearn.com/courses/delivering-research-data-management-services>
- RUG. Privacy in Research (MOOC): [https://www.rug.nl/research/research-data-management/data\\_protection-gdpr/data\\_protection/training-privacy-in-research](https://www.rug.nl/research/research-data-management/data_protection-gdpr/data_protection/training-privacy-in-research)
- SERISS. Data Management (Youtube modules): <https://seriss.eu/training/training-overview/#tab-6a6f811cc6a30ac8f54> (social sciences)
- [several] Open Science MOOC: <https://opensciencemooc.eu> (contains modules on Open data and on Open research software / Open Source)
- SURF. Data management, Big data and other more or less ICT-related courses: <https://www.surf.nl/en/training-courses-for-research>
- TUD. Open Science: Sharing your research with the world (MOOC): <https://online-learning.tudelft.nl/courses/open-science-sharing-your-research-with-the-world>

- University of North Carolina at Chapel Hill & University of Edinburgh. Research Data Management and Sharing (MOOC): <https://www.coursera.org/learn/data-management>

#### Trainings only accessible for affiliated people

Listed here are a few examples of trainings which require institutional credentials. They typically have a small focus, like writing data management plans or dealing with personal data. Despite the limited access to the trainings, this list is useful as a starting point for coordinating a national training provisioning through reuse.

- University of Amsterdam: Data Management: <https://rsp.uva.nl/en/training/data-management/data-management.html> (RDM Introduction, workshop on Encryption)
- Vrije Universiteit: Training 'Writing a Data Management Plan' for PhD students: <https://www.uv.vu.nl/en/university-library-for-researchers/research-data-support/training-and-seminars/index.aspx>
- TU/e: A basic course on research data management: <https://www.tue.nl/universiteit/bibliotheek/ondersteuning-onderwijs-onderzoek/wetenschappelijk-publiceren/data-coach/trainingen/>
- Leiden University: Training data management: <https://www.library.universiteitleiden.nl/researchers/data-management/training-data-management> (Different workshops (writing a DMP, make your own data FAIR, How to publish your data, Project & Data Management for PhD's))
- Maastricht University: An introduction to research data management: <https://library.maastrichtuniversity.nl/information-skills/workshops-courses/an-introduction-to-research-data-management>
- University of Twente: Data Management Bootcamp: <https://www.utwente.nl/en/ctd/courses/1000227/data-management-bootcamp> for First-year PhD-candidates from UT
- Utrecht University: <https://www.uu.nl/en/research/research-data-management/training-workshops> (Different workshops (Quick start to Research Data Management, Writing a DMP, Handling personal data in research, Introduction to R & Data, R Cafe, Best Practices for Writing Reproducible Code))
- Wageningen University & Research: Research Data Management <https://www.wur.nl/en/Library/Students/Courses-and-demos/Courses-and-demos-display/Research-Data-Management.htm>

## Training materials

The references in this list are no full-blown trainings but point to materials that data stewards should know about. The materials could well be embedded in trainings.

- Australian Research Data Commons. FAIR self-assessment tool: <https://www.andsnectar-rds.org.au/fair-tool>
- Brinkman, L., Oberski, D., Aarts, H. Open and Reliable Science Teaching Formats. 25 formats for bachelor education in the Social Science faculty of University Utrecht; targets teaching staff. [www.tinyurl.com/OSTeachingformats](http://www.tinyurl.com/OSTeachingformats)
- CLARIN Legal Information Platform: <https://www.clarin.eu/content/legal-information-platform>
- DANS. Training material: <https://dans.knaw.nl/en/about/services/training-consultancy/training>
- Digital Curation Centre. Guides and tools: <https://www.dcc.ac.uk/guidance>
- ELIXIR. TeSS training registry: <https://tess.elixir-europe.org>
- EOSC Hub. Training materials: <https://www.eosc-hub.eu/training-material>
- European Commission. Guidance on Open Access and data management: [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination\\_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm)
- FAIR evaluation services: <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!>
- FAIRsFAIR: FAIR-Aware: assess your awareness of making data FAIR: <https://fairaware.dans.knaw.nl>
- Federatie Auteursrechtbelangen. Auteursrecht/Intellectual property rights (in Dutch): <https://auteursrecht.nl>
- FOSTER. The Open Science Training Handbook (targets trainers). Various authors: <https://book.fosteropenscience.eu/en>
- HANDS: Handbook for Adequate Natural Data Stewardship v2.0. NFU Data 4 Lifesciences: <https://data4lifesciences.nl/hands2/data-stewardship>
- M. Jetten et al. 23 Things for Data Stewards. <https://doi.org/10.5281/zenodo.3773663> (Note: this package contains also sets of Things for other stakeholders)
- LCRDM wiki and knowledge platform: [https://www.edugroepen.nl/sites/RDM\\_platform/SitePages/Home.aspx](https://www.edugroepen.nl/sites/RDM_platform/SitePages/Home.aspx)
- OpenAIRE. Guides, factsheets and training materials: <https://www.openaire.eu/>
- PARTHENOS: Policies & Guidelines: [https://www.parthenos-project.eu/portal/policies\\_guidelines](https://www.parthenos-project.eu/portal/policies_guidelines) (humanities and cultural heritage)
- UK Data Service. Data management training resources: <https://www.ukdataservice.ac.uk/manage-data/training.aspx>



## Annex 9: Annotating trainings

### The annotation approach

To see if and how trainings match with the required competences of data stewards, we did a pilot study, mapping some of the trainings to the data steward competences that were identified previously (see Section 4.3). Originally, we looked for vocabularies and ontologies to describe trainings. A systematic description of, for instance, topic, learning goals, and type of training, would allow for comparing them. However, such an ontology seems to be lacking, although the emerging terms 4FAIRskills terminology<sup>134</sup> is promising. The data life cycle<sup>135</sup> doesn't provide a vocabulary and the Foster taxonomy of Open Science<sup>136</sup> has another focus than training. Furthermore, the Learning Object Metadata standard<sup>137</sup> concerns content elements within training, not the higher aggregation level of full trainings.

Eventually we used the list of data stewardship competence areas (Annex 3) with the added soft skills area. We selected five trainings (see Annex 8) and a university course, which we mapped in an annotation pilot study to the competence areas. Below we describe the different trainings and we end this annex with lessons learned.

### Helis FAIR Data Stewardship Course (DTL)

- *Website:* <https://helisacademy.com/en/data-analysis-stewardship>
- *Target audience:* wet-lab scientists - from industry and academia - and graduate students in the life sciences who wish to improve their digital scholarship on data handling. Some basic experience with programming and scripting languages like Python, Perl, R, MATLAB, etc. are an advantage, but not needed. The course is also relevant for data stewards who have to support life science Researchers.
- *Training description:* the course will introduce the trainees to important concepts of data stewardship<sup>138</sup>. We start with a general introduction covering the data life cycle, the FAIR principles<sup>139</sup> and a definition of data stewardship and data stewards. We will pass the stages of the data life cycle in more detail in the training modules of this 3-day training.

---

<sup>134</sup> <https://terms4fairskills.github.io>

<sup>135</sup> <https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>

<sup>136</sup> <https://www.fosteropenscience.eu/taxonomy/term/7>

<sup>137</sup> [https://www.edustandaard.nl/standaard\\_afspraken/nl-lom/nl-lom-versie-1-01](https://www.edustandaard.nl/standaard_afspraken/nl-lom/nl-lom-versie-1-01)

<sup>138</sup> <https://www.dtls.nl/fair-data/research-data-management/research-data-management>

<sup>139</sup> Wilkinson, M.D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data volume 3, Article number: 160018. <https://doi.org/10.1038/sdata.2016.18>

- *Goals:* gaining knowledge on research data management life cycle, FAIR principles and FAIR data stewardship; practical experience in data management planning, cleaning data, semantic interoperability between data, archiving and publishing of data, and persistent identifiers and their use cases

Competence area	Matching with training objectives
Policy/Strategy	Doesn't match with training goals or topics
Compliance	Doesn't match with training goals or topics
Alignment with FAIR data principles	Training is specifically focused on this topic
Services	Doesn't match with training goals or topics
Infrastructure	Training focuses on FAIRifying your data (very hands-on)
Knowledge management	Training contributes to solving a FAIR knowledge gap
Network	Doesn't match with training goals or topics
Data archiving	Training also focuses on the data archiving aspect of FAIR
Soft skills, e.g., communication and consultancy	Doesn't match with training goals or topics

*Table Annex 9.1 Mapping Helis FAIR Data Stewardship training*

**Conclusion:**

- This training focuses on competence area 3 (alignment with FAIR data principles), strongly connected to competence areas 5 and 8: infrastructure (hands-on FAIRifying your data) and data archiving (as an explicit part of the data life cycle/FAIR data).
- The training specifically targets data steward and data stewardship skills, so a very good example for training data stewards. It helps that it is a 3-day training, so it offers in-depth knowledge.

- The training contains both theoretical lectures and hands-on practice sessions. Training is face-to-face currently, but there are plans to adjust it into an (additional) online training.

#### Essentials 4 Data Support (RDNL, the national alliance of DANS, SURF, and 4TU)

- *Website:* <https://datasupport.researchdata.nl/en>
- *Mission:* the Essentials 4 Data Support (E4DS) training aims to contribute to professionalisation of data supporters and coordination between them. Data supporters are people who support researchers in storing, managing, archiving and sharing their research data.
- *Target group:* the training focuses on anyone wanting to support researchers in storing, managing, archiving and sharing research data: a data supporter. Think, for instance, of (data) librarians, ICT staff and researchers with duties involving data management. Research Data Netherlands (RDNL) intentionally uses the term “Data supporter”, to avoid discussion about job titles.

The training Essentials 4 Data Support was designed with the following competences<sup>140</sup> in mind:

Competence	Description
Skilfully handles ICT	Efficiently uses available information technology.
Shows entrepreneurship	<ul style="list-style-type: none"> <li>• Aims to improve data services in response to changing needs in the field. Keeps an eye on trends which emerge in the profession, knows where knowledge is available (networks) and disseminates important information to key people in the organisation.</li> <li>• Regularly inquires into perceived needs in the field, e.g., by using questionnaires, interviews or focus groups.</li> <li>• Actively contributes to developments in the field by visiting or contributing to training sessions, conferences etc.</li> </ul>
Sees from the whole	<ul style="list-style-type: none"> <li>• Acknowledges that data are only part of the scientific lifecycle and is aware of the significance research data have for carrying out scientific research.</li> </ul>

<sup>140</sup> <https://datasupport.researchdata.nl/en/about-the-course/competencies>

	<ul style="list-style-type: none"> <li>• Sees data- and information services as part of larger whole in which decisions are made.</li> </ul>
Consulting skills	<ul style="list-style-type: none"> <li>• Can handle questions skilfully. Knows when to give advice and when to refer question about data management to a dedicated expert (e.g., questions about data formats, data documentation, storage, data citation (persistent identifiers), writing a data management plan (DMP), intellectual property and funder requirements).</li> <li>• Can empathise with customer perceptions.</li> <li>• Asks for feedback on one's consulting skills and adjust one's behaviour accordingly.</li> </ul>
Co-operative skills	<ul style="list-style-type: none"> <li>• Examines how collaboration with others (employees, researchers, institutions) may enhance service provision.</li> <li>• Acknowledges the necessity of a forum where data supporters can communicate and stand up together to make a fist when it comes to important themes like data policies, copyright and information-infrastructure.</li> <li>• Takes responsibility for one's contribution to these partnerships.</li> </ul>
Skilfully handles ICT	Efficiently uses available information technology.

*Table Annex 9.2 Competences in Essentials 4 Data Support*

When we try to map those competences and the competence areas used in this report, the following table results:

Competence area	Matching with Essentials 4 Data Support training competences
Policy/Strategy	Matches with sees from the whole
Compliance	Doesn't match with an explicit competence. Addressed in Chapter V - Data, legislation and policy <a href="https://datasupport.researchdata.nl/en/start-the-course/v-legislation-and-policy">https://datasupport.researchdata.nl/en/start-the-course/v-legislation-and-policy</a>

Alignment with FAIR data principles	Doesn't match with an explicit competence. Ubiquitously addressed in E4DS
Services	Matches with skilfully handles ICT and shows entrepreneurship
Infrastructure	Doesn't match with an explicit competence. Addressed in Chapter III - Research phase <a href="https://datasupport.researchdata.nl/en/start-the-course/iii-research-phase">https://datasupport.researchdata.nl/en/start-the-course/iii-research-phase</a> , e.g., storing data, Virtual Research Environments
Knowledge management	Matches with skilfully handles ICT and shows entrepreneurship
Network	Matches with shows entrepreneurship
Data archiving	Doesn't match with an explicit competence. Addressed in Chapter IV - Harvest phase <a href="https://datasupport.researchdata.nl/en/start-the-course/iv-harvest-phase/data-archives">https://datasupport.researchdata.nl/en/start-the-course/iv-harvest-phase/data-archives</a>
Soft skills, e.g., communication and consultancy	Matches with consulting skills and co-operative skills

*Table Annex 9.3 Mapping Essentials 4 Data Support*

#### Conclusion:

- This training addresses all competence areas, as it is a foundational training for data supporters. The face-to-face version with practical assignments runs six weeks with two contact days, and all content is also available for online self-paced learning. Many data stewards mentioned the training as truly essential (see the interviews in Chapter 3).
- Some competence areas match the competences defined by RDNL or training chapters, but this is accidental. Alignment with FAIR data principles (competence area 3) is intentionally addressed throughout Essentials, as it links to all other competences.
- RDNL pays much attention to strong soft skills for RDM support, e.g., with respect to collaboration and advice. In the training the skills are directly applied to RDM and the various RDM stakeholders. A competence area “professional soft skill” is missing in the original list, however.

## FOSTER. Assessing the FAIRness of Data

- *Website:* <https://www.fosteropenscience.eu/node/2644>
- *Target group:* the training is suitable for both researchers and research support staff.
- *Mission:* in this short training, you'll learn how to go about assessing how findable, accessible, interoperable, and reusable (FAIR) your research data are using freely available tools and resources.
- *Introduction to the training:* you have likely heard people using the term 'FAIR' data a lot recently and might wonder what exactly is meant by this term. FAIR data are those that are Findable, Accessible, Interoperable and Reusable. Sounds simple enough, but what do each of these terms mean in a practical sense and how can you tell if your own research data is FAIR? This short training will: introduce you to the key terms and explain what they mean in a practical sense; tell you how data management planning can help to make data FAIR from the very start of research projects; show you how you can use freely available tools to help assess the FAIRness of data.

Competence area	Matching with training objectives
Policy/Strategy	Doesn't match with training goals or topics
Compliance	Doesn't match with training goals or topics
Alignment with FAIR data principles	Training is specifically focused on this topic
Services	Doesn't match with training goals or topics
Infrastructure	Training focuses on tools for FAIR data (including assessment)
Knowledge management	Training contributes to solving a FAIR knowledge gap
Network	Doesn't match with training goals or topics
Data archiving	Training also focuses on the data archiving aspect of FAIR, i.e., tools for FAIR data archiving (including assessment)

Soft skills, e.g., communication and consultancy	Doesn't match with training goals or topics
--	---

Table Annex 9.4 Mapping FOSTER: Assessing the FAIRness of Data

#### Conclusion:

- This training focuses on competence area 3: alignment with FAIR data principles.
- The training has such a broad audience, that it doesn't specifically focus on competences for a data steward.

#### FOSTER. Data Protection and Ethics

- *Website:* <https://www.fosteropenscience.eu/node/2330>
- *Target group:* the training is suitable for both researchers and research support staff.
- *Mission:* this training helps you to get to grips with data protection and the ethics around responsible data sharing.
- *Introduction to the training:* This training covers data protection in particular and ethics more generally. It will help you understand the basic principles of data protection and introduces techniques for implementing data protection in your research processes. Upon completing this training, you will know what personal data are and how you can protect them; what to consider when developing consent forms; how to store your data securely; how to anonymise your data.

Competence area	Matching with training objectives
Policy/Strategy	Doesn't match with training goals or topics
Compliance	Training is specifically focused on the compliance to the GDPR on personal data protection during research. In the sub-topic "What are personal data?", legal requirements (GDPR) are discussed. Also, the ELIXIR Webinar on the GDPR in research is included. Informed consent, and the development of a proper ethical/legal form, is also an important topic in this training.
Alignment with FAIR data principles	Training is specifically focused on this topic

Services	Doesn't match with training goals or topics
Infrastructure	Training focuses on secure data storage and gives examples of archives
Knowledge management	This training contributes to solving a GDPR compliance knowledge gap in the organisation
Network	Doesn't match with training goals or topics
Data archiving	Training also focuses on the data archiving aspect personal data within the scope of the GDPR
Soft skills, e.g., communication and consultancy	Doesn't match with training goals or topics

Table Annex 9.5 Mapping FOSTER: Data Protection and Ethics

#### Conclusion:

- This training focuses on part of competence area 3: compliance to the GDPR. This touches also on areas 5 (Infrastructure) and 7 (Network), in the context of secure storage and archiving personal data, respectively.
- The training has such a broad audience, that it doesn't specifically focus on competences for a data steward. However, it does contribute to gaining knowledge on compliance to the GDPR in research, as one of the competences for data stewards.

#### FOSTER. Managing and Sharing Research Data

- *Website:* <https://www.fosteropenscience.eu/node/2328>
- *Target group:* the training is suitable for researchers.
- *Mission:* in this training, you'll focus on which data you can share and how you can go about doing this most effectively.
- *Introduction to the training:* data-driven research is becoming increasingly common in a wide range of academic disciplines, from Archaeology to Zoology, and spanning Arts and Science subject areas alike. To support good research, we need to ensure that researchers have access to good data. Upon completing this training, you will: understand which data you can make open and which need to be protected; know how



to go about writing a data management plan; understand the FAIR principles; be able to select which data to keep and find an appropriate repository for them; learn tips on how to get maximum impact from your research data.

Competence area	Matching with training objectives
Policy/Strategy	Doesn't match with training goals or topics
Compliance	Doesn't match with training goals or topics
Alignment with FAIR data principles	On a basic level, the FAIR principles are addressed in this training
Services	Doesn't match with training goals or topics
Infrastructure	DMPonline is mentioned as a tool for writing a data management plan. Also, different archives are appointed in the training
Knowledge management	This training contributes to solving a basic RDM knowledge gap in the organisation. The training can serve as a basis for developing RDM training within the institute
Network	Doesn't match with training goals or topics
Data archiving	On a basic level, data archiving is covered in the training. The information is tailored towards the researcher, but gives a basic overview of the topic
Soft skills, e.g., communication and consultancy	Doesn't match with training goals or topics

*Table Annex 9.6 Mapping FOSTER: Managing and Sharing Research Data*

#### Conclusion:

- This training is really tailored towards researchers at the start of their RDM knowledge. Therefore, a lot of topics are discussed at a basic level, which are mainly related to competence areas 5 (Infrastructure) and 6 (Knowledge management), where the training mentions a tool for writing DMPs and aims to fill an organisational knowledge gap, respectively.
- The training could be used for a beginning data steward. It could also serve as a basis for developing RDM training within the institute.

#### UAS Leiden

The University of Applied Sciences Leiden (Hogeschool Leiden) teaches Computer Science at a Bachelor level. One of the specialisations of Computer Science is Business Data Management (BDaM), which is a domain-agnostic course. Also, there is an elective course 'Introduction to FAIR Data stewardship'. BDaM and the elective course were compared to the learning objectives for a data steward of type "C/Infrastructure"<sup>141</sup>, which lead to the following observations:

- The University of Applied Sciences Leiden aims to teach skills and competences that qualify students for a variety of jobs, not specifically for becoming a data steward within a research institute.
- The Data Steward C learning objectives are too fine-grained for a Computer Science bachelor's degree curriculum or course.
- However, it would be possible to match the Data Steward C learning objectives and the course learning goals.
- A student who has completed the BDaM curriculum is able to work as a data steward C – Infrastructure. If that student has also followed the elective 'Introduction to FAIR Data Stewardship' course, he or she is even more prepared.

#### Lessons learned from annotating trainings with competence areas

From the annotation pilot study, we draw the following lessons:

- The process was straightforward: select a training and compare its - preferably explicit - learning goals to the competence areas. The goal of this exercise was to find an

---

<sup>141</sup> See Section 2.3 and Scholtens, S., Jetten, M., Böhmer, J., Staiger, Ch., Slouwerhof, I., Van der Geest, M. & Van Gelder, C.W.G. (2019, October 3). Final report: Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. <http://doi.org/10.5281/zenodo.3474789>

appropriate training to bridge a knowledge or skill gap, identified in terms of one or more competence areas.

- The mapping process, also known as annotating, was quite feasible. Per course it took between one and four hours. Clearly, trainings with more focus or a smaller remit - e.g., “Data protection and ethics” - match fewer competence areas than trainings with a larger remit, e.g., “Essentials 4 Data Support”.
- One needs some knowledge about the training that is to be annotated. The title is probably insufficient, and the training’s learning goals may be orthogonal to the competence areas. (For instance, while “Alignment with FAIR data principles” is a separate competence area, this topic is addressed in several chapters of “Essentials 4 Data Support”, without even being mentioned in the training’s learning goals.) Therefore, we recommend that training producers annotate their own training(s).
- The competences areas are few and high-level, whereas the learning objectives are many and fine-grained. In theory, a level in between might be ideal in terms of training annotation, but in practice this would require more time from the annotator, as well as more knowledge about both the training and the competences.
- For the time being, we find the competence areas useful for annotating training. They should be evaluated on the basis of more mappings, which might lead to rewriting and/or finer-grained items.

## Annex 10: Certification of data stewardship

For education, three universities/UMCs responded, and four universities of applied sciences. For training we received input of seven not for profit organisations, and three for profit organisations. We realise that our inventory is not extensive, but we feel that the results of our stock taking indicate the trend of certification in both education and training. Table 10.1 below shows an overview of certification methods and mechanisms for education and training. Table 10.2 shows an overview of certification methods and mechanisms related to data stewardship training and education. The summary of the outcomes of the inventory is given below.

### Education

- The quality of teachers at universities and university medical centres is assessed by the VSNU through Basiskwaliteit Onderwijs (BKO)<sup>142,143</sup>. Teachers at universities of applied sciences are certified through Basiskwalificatie Didactische Bekwaamheid (BDB)<sup>144</sup>.
- The Dutch-Flemish Accreditation Organisation (NVAO) oversees the quality of higher education in the Netherlands and Flanders, by accrediting the organisation itself and/or education. Certification is currently implicit in Bachelor (BSc) and Master (MSc) degrees. The students are rewarded with ECTS (European Credit Transfer System), and a leaving certificate after completion.
- Only one university and two universities of applied sciences provide data stewardship training. Thus far only one UAS (Hogeschool Leiden) offers certification (Hogeschool Leiden is also planning to offer the course as 'contractonderwijs').

### Training

- The majority of the non-profit organisations offer a certificate of attendance for both general courses and specific data stewardship courses, whereas the courses themselves are not certified, nor the trainer or organisation.
- One example of certification related to trainers relates to courses of the Netherlands eScience Center based on the Carpentries (Data, Software, Library Carpentry), which are taught by certified Carpentries trainers.
- The for-profit organisations that offer data stewardship courses do not have certification in place for their courses, trainers, trainees and their particular organisation. However, certificates of attendance are handed out to the trainees in case of the FAIR data stewardship courses of Phortos Consultants and the Hyve.

---

<sup>142</sup> <https://vsnu.nl/bko.html>

<sup>143</sup> [https://www.vsnu.nl/en\\_GB/utq](https://www.vsnu.nl/en_GB/utq)

<sup>144</sup> <https://www.hsleiden.nl/nascholingen/didactische-bekwaamheid/index.html>

Organisation	Certification for courses	Certification for trainees	Certification for trainers	Certification for organisation
<b>Educational Programs</b>				
<b>University &amp; UMC</b>				
Bachelor (BSc)	implicit in BSc degree	grades (official certificates), diploma	BKO	NVAO
Master (MSc)	implicit in MSc degree	grades, diploma	BKO	NVAO
Postgraduate				
Summerschools	certificates, ects	certificates, ects		
<b>University of Applied sciences</b>				
Bachelor post HBO	implicit in BSc degree	grades (official certificates), diploma	BDB/education on subject	NVAO
HBO master (funded)		certificates	BDB	NVAO
HBO master (not funded)		diploma	Msc?	NVAO
Minor at universities	implicit in BSc degree	official certificates	Msc? BKO	
<b>Organisation</b>				
<b>Training Providers</b>				
<b>Non-commercial</b>				
DTL/ELXIR/BioSB/Helis Academy	no	certificate of attendance with indication of workload (expressed in ECTS)	no	no
DANS	no	no	no	no
RDNL	no	certificate of active participation, not certification	no	no
The Netherlands eScience Center	data carpentry workshops	carpentries approved content	certificate of attendance on request	carpentries certified
SURF	yes (essentially for tools)	no	only for participation (usually upon request)	no
Hogeschool Leiden	yes (planned)			
<b>Commercial</b>				
Phorfos	no	certificate of attendance	no	companies that follow a
The Hye	no	no	no	OHDSI training receive
Micelio	no	no	no	certificate of attendance

Table Annex 10.1 Certification methods and mechanisms for education and training

Organisation	Data stewardship training in your organisation	Certification for data stewardship courses	Certification for data stewardship trainees	Certification for data stewardship trainers	Certification for organisation
<b>Educational Programs</b>					
<b>University &amp; UMC</b>					
TU Delft	yes	no	no	no	no
Erasmus MC / Erasmus University	no	no	no	no	no
UM / MUMC+	no	no	no	no	no
<b>University of Applied sciences</b>					
Hogeschool Leiden	yes	yes			
Hogeschool Avans	no	no	only school certificate to the students who finish the bioinformatics minor successfully	no	no
Fontys Hogeschool	yes, but extern and/or online	no	no	no	no
HAN	no	no	no	no	no
<b>Training Providers</b>					
<b>Non-commercial</b>					
DTL/ELIXIR/BioSB/Helix Academy	yes	no	certificate of attendance with indication of workload (expressed in ECTS)	no	no
DANS	yes	no	no	no	no
RONL	yes	no	certificate of active participation	no	no
The Netherlands eScience Center	data carpentry workshops	carpentries approved content	certificate of attendance	carpentries certified instructors	
SURF	yes (essentially for tools)	no	only for participation	no	only for PRACE courses
Hogeschool Leiden	yes (planned)				
<b>Commercial</b>					
Phortos	yes	no	certificate of attendance	no	no
The Hye	yes	no	no	no	no
Micelio	yes	no	no	no	no

Table Annex 10.2 Certification methods and mechanisms for data stewardship education and training

## Annex 11: Existing tool examples (Competency Hub, CILIP PKSB)


### Competency Hub

#### BioExcel


2.0 live

The BioExcel competency framework consists of a series of competencies that professionals in the field of computational biomolecular research might need. A competency is an observable ability of any professional, integrating multiple components such as knowledge, skills and behaviours. A competency framework lists and defines all the competencies that characterise specific user groups. The competencies that an individual might need to fulfil a particular role are listed in reference profiles, which can be used to guide career choices.


Career profiles   Competencies   Training resources

 Discover and explore

Explore career profiles within this competency framework

 Create your own profile

Create your own profile and assess yourself against the competencies


 Compare profiles

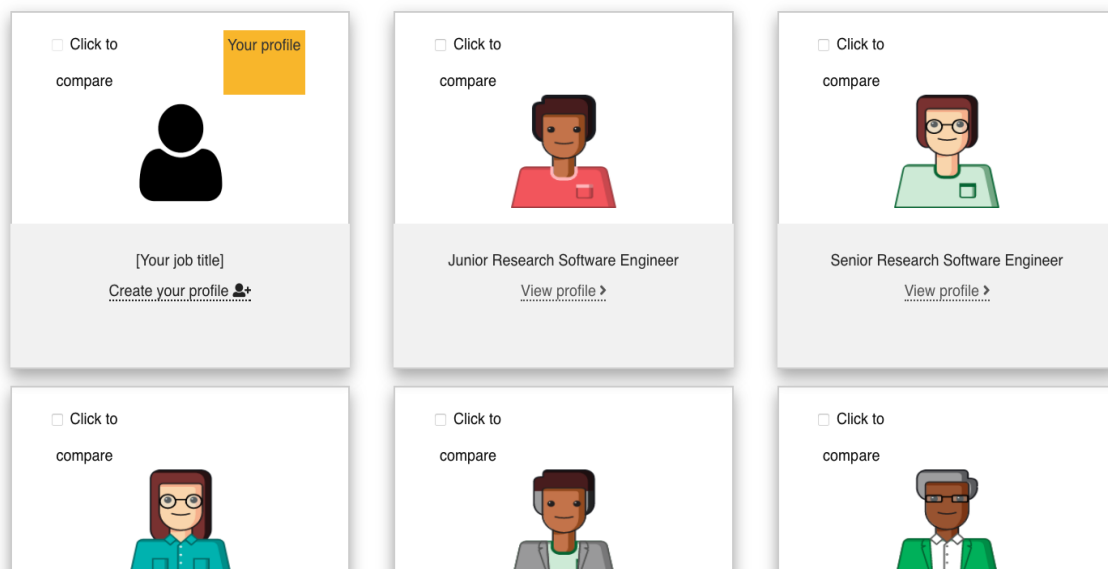
Compare your profile with other reference profiles to help you make career choices based on your competencies and interests

 Identify training opportunities

Training opportunities will help boost your career

Career profiles

Compare selected profiles 



The screenshot displays a grid of six career profile cards. Each card includes a 'Click to compare' button, a profile picture, a job title, and a 'View profile' link. The first card is highlighted as 'Your profile'.










Profile	Job Title	Action
Your profile	[Your job title]	Create your profile 
	Junior Research Software Engineer	View profile 
	Senior Research Software Engineer	View profile 
		
		
		

Figure Annex 11.1 Competency Hub example: career profiles, competences and training resources<sup>145</sup>

<sup>145</sup> <https://competency.ebi.ac.uk/framework/bioexcel/2.0>

# Competency Hub

Home / BioExcel 2.0

Admin login 

## Kyo - Senior Research Software Engineer



36 years

[Create your profile !\[\]\(17acf1afa8cdf0b67c53d4865a5ed469\_img.jpg\)](#)

### Qualification and background

Kyo wrote an initial version of the cellular **image analysis application** from scratch during a postdoc position in another group. They decided to create it because they enjoyed **writing software** and because they saw an opportunity to advance their own research. Following some initial novel results, the emergence of new imaging techniques meant that a fundamental rewrite of the code was needed soon after Kyo joined the current research group. Whilst the resulting software was innovative and pushed the state of the art in what could be done, the time Kyo had to spend developing the software meant they weren't able to do as much research and as a result published fewer papers. In recognition of the critical importance of Kyo's software to the success of the group and indeed the field, senior colleagues have been finding ways to fund Kyo from their research budgets to spend most of the time developing the code. This has become easier in recent years as funding bodies have come to recognise the value of software that enables cutting edge computational research, including in the life sciences, and Kyo is considering applying for a Research Software Engineer Fellowship to directly fund them in this career path.

### Activities of current role

Kyo is 36 and is the primary developer of a popular **software package** that performs **quantitative image analysis** of cellular structures and their 3D spatial relationships in complex biological systems based on results from a variety of types of microscopic imaging techniques. Kyo is part of a research group mainly performing experimental work and the software Kyo has developed over the course of the past 5 years has enabled researchers in the group and elsewhere to perform cutting edge analyses that have led to the publication of a number of high impact papers.

Kyo's main activities are:

- **Staying up to date** with evolving imaging and cutting edge experimental techniques and research questions by talking to local colleagues and to researchers at relevant conferences and workshops
- **Identifying new functionality** to add to their software package, e.g. being able to handle new types of imaging data or perform more sophisticated analyses
- **Architecting** how this new functionality can best be implemented within the existing framework of the application and given the hardware it runs on in order to ensure good performance and usability
- **Testing new code** to make sure it gives scientifically meaningful results
- **Fixing bugs** reported by users of the package
- **Writing documentation**
- **Training** colleagues how to use new functionality in the software

BioExcel 2.0 / Competencies

0 Not applicable   1 Awareness   2 Working knowledge   3 Specialist knowledge

### Scientific Competencies

Apply expertise in formal & natural sciences appropriate to the discipline, and follow best practice in experimental design

1

User-driven service provision and support

2

Search for, assess and compile appropriate literature and data sets to address specific research questions


2

*Figure Annex 11.2 Competency Hub example: career profile and matching competences<sup>146</sup>*

<sup>146</sup> <https://competency.ebi.ac.uk/framework/bioexcel/2.0>










## Competency Hub


Admin login 

### Overview

Competency Hub is a web-based tool to support the creation and management of competency frameworks [read more](#) 

 <p>Professionals in computational biomolecular research</p>	 <p>Technical operators of biomedical research infrastructures</p>	 <p>Managers of research infrastructures</p>
 <p>Students and professionals in computational biology</p>	 <p>Clinical practitioners for the application of genomics in the healthcare service</p>	 <p>Professionals working with human data for disease research</p>
 <p>Data steward roles: policy, research and infrastructure</p>		


If you have any questions, comments or suggestions, please contact us: [competency \[at\] ebi.ac.uk](mailto:competency@ebi.ac.uk)

[Developer documentation](#) 

*Figure Annex 11.3 Competency Hub: result of a pilot to include the data steward competency framework<sup>147</sup>*

<sup>147</sup> <https://competency.ebi.ac.uk>

## Competency Hub

Admin login 

### Data Steward

1.0 

The Data Stewardship Competency Framework distinguishes three data steward roles (policy, research and infrastructure) and eight competence areas (policy/strategy, compliance, alignment with FAIR data principles, services, infrastructure, knowledge management, network, data archiving and transferable skills). For each of these three data stewardship role competencies, KSAs (knowledge, skills and abilities) and learning outcomes (including Bloom's level) are given.

Competencies Training resources

Filter competencies

#### Policy and Strategy: Development, implementation and monitoring of the research data management policy and strategy of the institute

- › [Policy Oriented Data Steward] Responsible for advice on and development, implementation and monitoring of a research data management policy and strategy for the 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) data management plans
- › [Research Oriented Data Steward] Responsible for the development and implementation of a data management plan 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 research data management policy and supports FAIR data and Open Science
- › [Infrastructure Oriented Data Steward] Responsible for identifying the requirements of an adequate data-infrastructure and tool landscape that fits with the institute's research data management policy and supports FAIR data and Open Science

#### Compliance: Compliance with relevant codes of conduct, legislation and field specific standards

- › [Policy Oriented Data Steward] Responsible for compliance of the research data management policy with codes of conduct, as well as continuous alignment with legal and ethical standards
- › [Research Oriented Data Steward] Responsible for advice on compliance of the project or data collection with the data management plan and the institute's research data management policy, relevant codes of conduct, legislation and field specific standards
- › [Infrastructure Oriented Data Steward] Responsible for compliance of the data-infrastructure and tool landscape to the institute's research data management policy, relevant codes of conduct and legislation

#### FAIR data: Alignment to the FAIR data principles and the principles of Open Science

- › [Policy Oriented Data Steward] Responsible for alignment of the research data management policy to the FAIR data principles and the principles of Open Science
- › [Research Oriented Data Steward] Responsible for alignment of the data management plan to the FAIR data principles and the principles of Open Science, and for facilitating and supporting FAIR data
- › [Infrastructure Oriented Data Steward] 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

#### Services: Availability of adequate support on research data management, in staff or services

- › [Policy Oriented Data Steward] Responsible for the availability of adequate support on research data management, in staff or services, for the researchers and research

*Figure Annex 11.4 Competency Hub data steward competences: result of a pilot to include the data steward competency framework<sup>148</sup>*

<sup>148</sup> <https://competency.ebi.ac.uk/framework/datasteward/1.0>

## My Progress

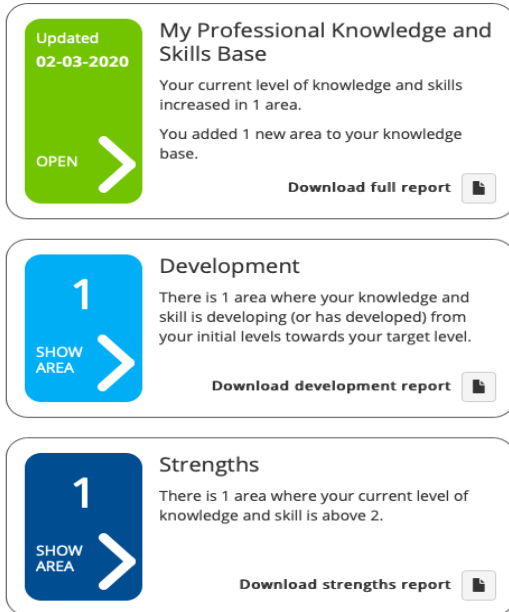


Figure Annex 11.5 CILIP PKSB example: skills areas graded according to the self-assessment ratings (scale 0 to 4)<sup>149</sup>

Knowledge / Skills Area	RATING			Comments
	Initial	Current	Target	
<b>4.1</b> Understanding research Appreciating the nature and value of research, including practitioner research, and research carried out on behalf of clients	0	2	3	Please enter any comments here
<b>4.2</b> Evaluative research Evaluations and comparisons of systems and services, including experiments, observations and surveys	1	1	1	Please enter any comments here
<b>4.3</b> Desk research Undertaking research, including literature analysis and methods such as content analysis and historical research	2	4	3	Please enter any comments here

Figure Annex 11.6 CILIP PKSB example: skills areas graded according to the self-assessment ratings, using colour-coding to indicate areas where development is needed (orange) and where rating meets or exceeds the target required (green)

<sup>149</sup> <https://www.cilip.org.uk/page/PKSB>

Key to self assessment ratings ✕

Rating	Descriptor	Description
0	None	I have no understanding of this subject or practical experience of this skill
1	Basic	I have a basic awareness and understanding of this area <b>AND/OR</b> I am able to use this to carry out simple tasks under instruction or supervision
2	Good	I have a good understanding of the concepts, principles and theories <b>AND/OR</b> I can use this to carry out tasks with a measure of complexity with limited supervision or on my own
3	Comprehensive	I have a comprehensive understanding of the subject area <b>AND/OR</b> I am able to apply this in my day to day work and have experience in doing so. I can support others in developing their knowledge and skills
4	Advanced	I have a specialised understanding of this area and can use this to solve problems in specialist areas <b>AND/OR</b> I am highly skilled in this area and am able to use my skills to deliver change. I am recognised as an expert in this area

Figure Annex 11.7 CILIP PKSB example: rating scale for self-assessment



Figure Annex 11.8 CILIP PKSB basic components overview, which also functions as the tool's navigation wheel

## Annex 12: Data steward personas and pathways

### Personas

Based on the ZonMw data stewardship and the LCRDM data stewardship projects (see Section 2.2), we distinguish the following five different data steward personas. The first three personas reflect the individual role a data steward could have in organisation, i.e., policy oriented, research oriented and infrastructure oriented. The last two data stewards reflect an organisational perspective, namely a generic data steward, for instance hired in the organisation's central library or research support service, and an embedded data steward, working for a specific faculty, discipline, department or project.

The user stories we drafted for each of the personas can help identify the wishes and needs for a data steward skills tool (see pathways below). For formulating the user stories, we used the format "As a <persona (who)> I can <capability (what)> so that <receive benefits (why)>"<sup>150</sup>.

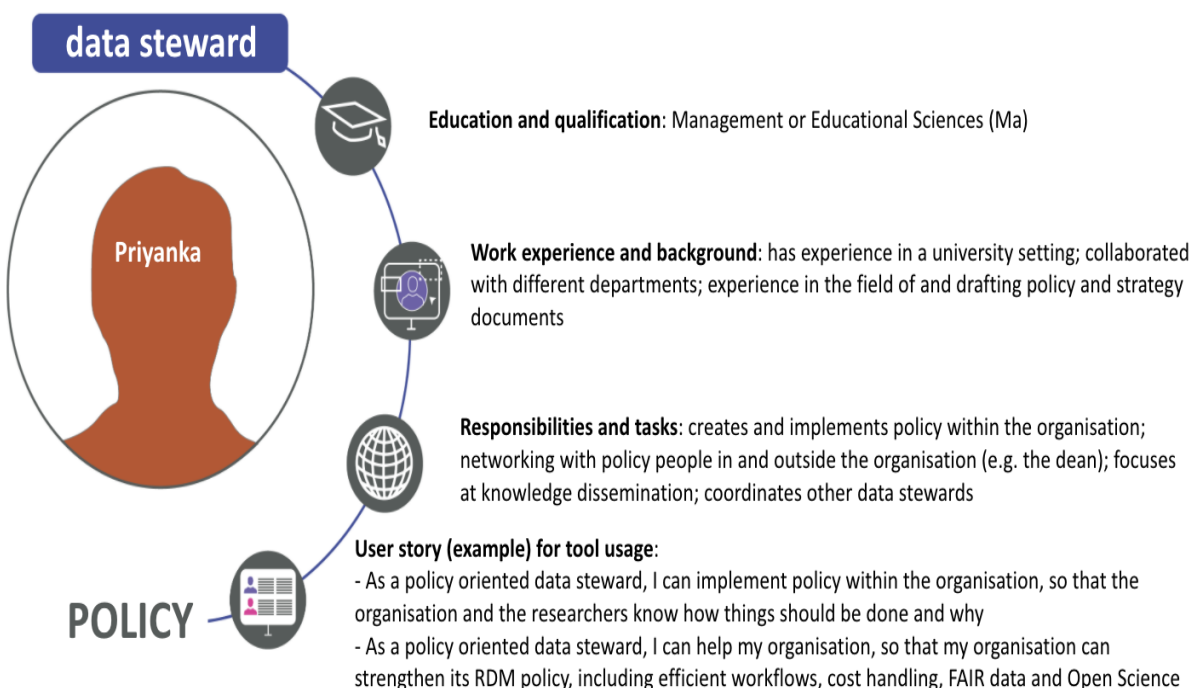


Figure Annex 12.1 Policy data steward persona (individual perspective)

<sup>150</sup> Ralph, Paul (2015). "The Sensemaking-coevolution-implementation theory of software design". Science of Computer Programming. 101: 21–41. <https://doi.org/10.1016/j.scico.2014.11.007>

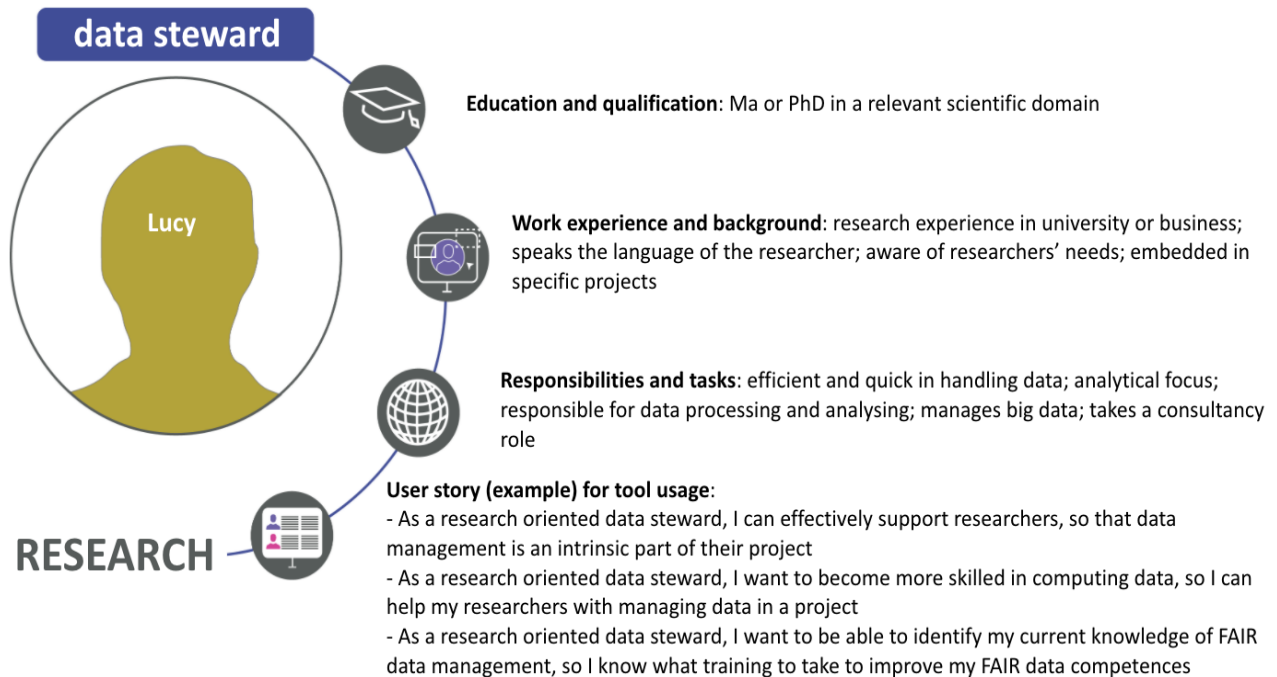


Figure Annex 12.2 Research data steward persona (individual perspective)

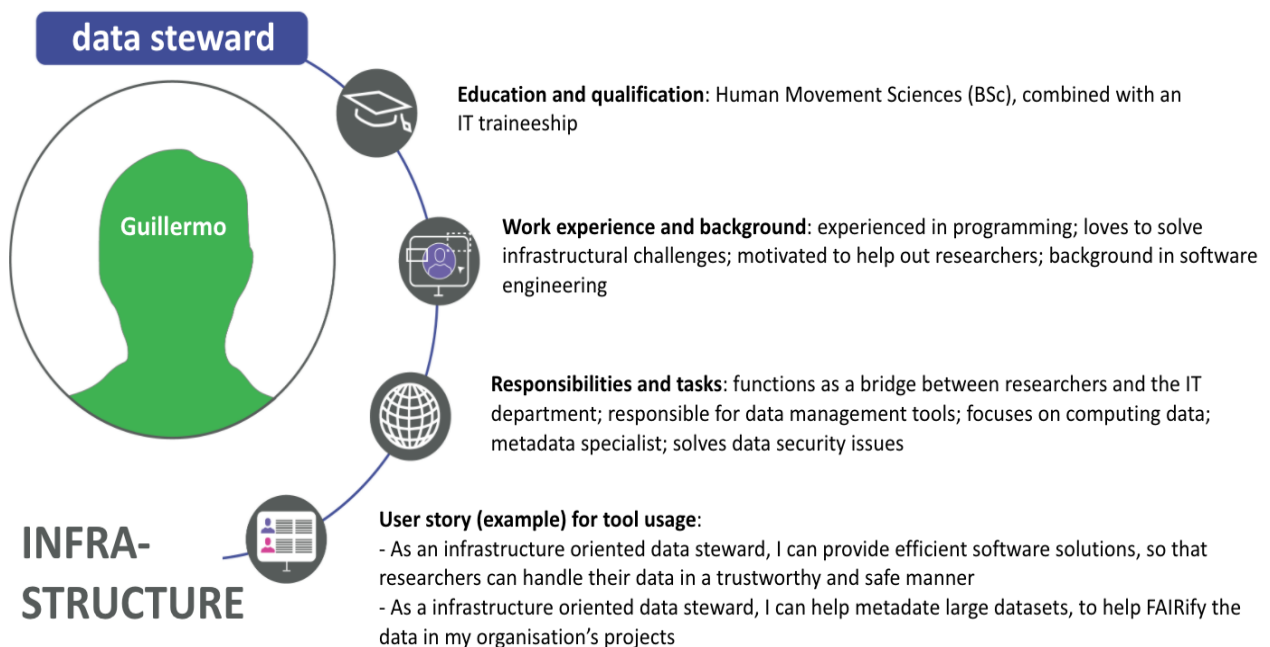


Figure Annex 12.3 Infrastructure data steward persona (individual perspective)



Figure Annex 12.4 Generic data steward persona (organisational perspective)

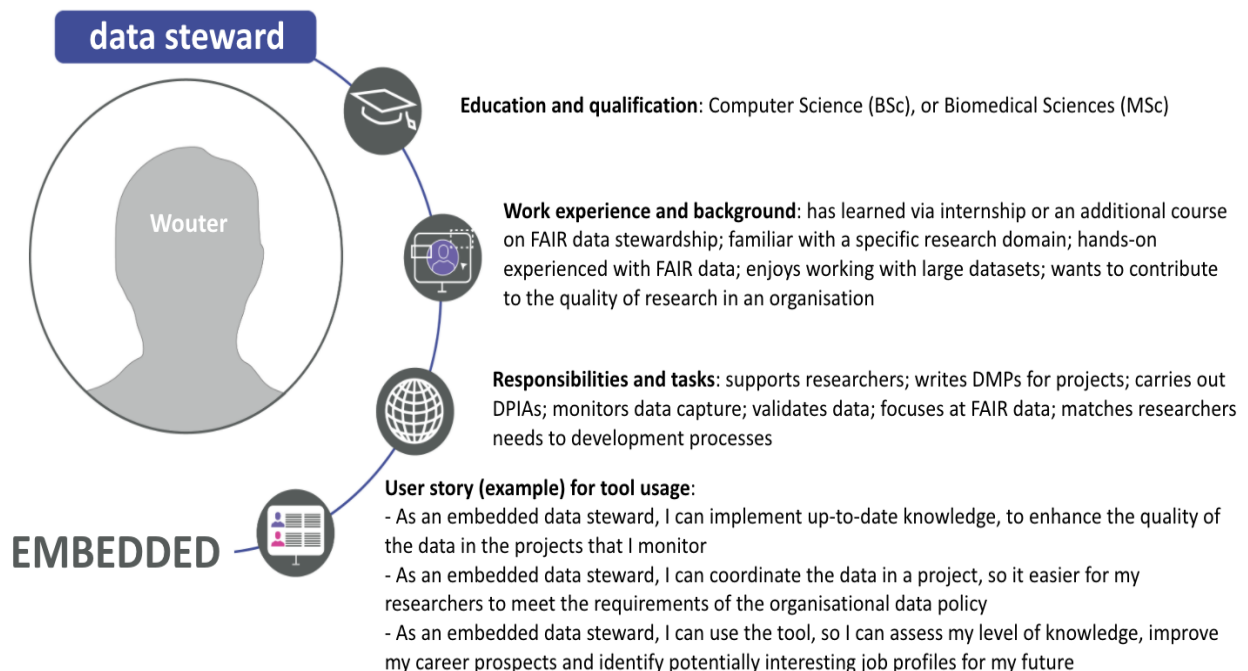


Figure Annex 12.5 Embedded data steward persona (organisational perspective)

## Tool pathways

Following the above-mentioned user stories, there are several situations where a data steward may use a skills tool.

- A data steward could use the tool as a general information resource, for example at the start of a new job.
- A data steward may want to find out which knowledge gaps he or she has.
- A data steward already knows that he or she has knowledge gaps and wants to solve them through training or other useful resources.
- A data steward may want to know his or her level (beginner, intermediate, advanced) in different competence areas. The outcome could be used during the annual talk with a manager or when the data steward aims to be promoted to a different salary scale.
- A data steward may also have spare time to follow training or wants to know what might be the best investment of his or her time to get the most benefit.

Consequently, the tool should at least match the following criteria:

- A data steward should be able to view the competences that correspond to the different data steward personas.
- A data steward should be able to self-assess these competences.
- A data steward should be able to view the learning resources corresponding to these competences, ideally including suggestions for a learning path.
- A data steward should be able to export the results of the self-assessment, to allow sharing in a team or organisation.

Since there are different starting points for a steward, a data steward skills tool should at least include the following features, to facilitate the potential, different learning paths of a data steward:

- A clear menu with the following sub elements:
  - *List of personas*: each persona in the list links to competences that are highly recommended for this persona.
  - *List of competence areas*: each competence in the list will link to relevant training resources that match the competence. Each competence will also link to self-assessment.
  - *List of training resources*: each training resource will link to relevant competences. Additionally, an agenda of upcoming training events could be part of the tool.
  - *Self-assessment*: this menu section offers the option to assess your knowledge, skills and abilities on each of the competence areas.



- Optionally: a *search box* to land on a specific page containing the search term
- To assist the data steward in finding what he or she wants to get out of the tool in the first place, based on the situations sketched above, the menu could also have a question-like approach. The data steward is asked what he or she would like to do, followed by some options to choose from:
  - I would like to do a self-assessment (self-assessment menu).
  - I would like to search the list of competence areas (competence areas menu).
  - I would like to see an overview of competences linked to my own profile (personas menu).
  - I would like to see a list of training resources (training resources menu).

These situations, needs, and criteria ask for a clear sequencing of the elements that will be presented in a data steward skills tool. For this sequencing it is important to keep the real context of the data steward in mind, to make connections between different concepts and to connect to the needs and interests of the data steward (Posner & Strike, 1976<sup>151</sup>; Morrison, Ross & Kemp, 2013<sup>152</sup>), resulting in different learning paths or journeys, as visualised below.

The tool should also be flexible to the needs and preferences of the data stewards and provide the user with so-called 'learner control' (Lawless & Brown, 1997)<sup>153</sup>. The learning pathway is not strict, but the tool can guide the data steward, for example by already showing training resources related to a specific competence area, when the data steward reads information about that competence area.

All data stewards start the tool the same way. However, depending on the initial purpose of the data steward to use the tool, the actions that follow may differ. They match at least the following perspectives, which should be reflected in the tool's menu options personas, competences, training resources and self-assessment. Visualised below are these four different paths to proceed in the tool, after the data steward chooses either of the four main menu options.

---

<sup>151</sup> Posner, G.J. & Strike, K.A. (1976). A categorization scheme for principles of sequencing content. *Review of Educational Research*, 46, pp. 665-690.

<sup>152</sup> Morrison, G. R., Ross, S. M., Kalman, H. K., & Kemp, J. E. (2013). *Designing Effective Instruction*. John Wiley & Sons Inc, 7th Edition.

<sup>153</sup> Lawless, K.A., Brown, S.W. (1997) *Multimedia learning environments: Issues of learner control and navigation*. *Instructional Science*, 25 (2), pp. 117-131.

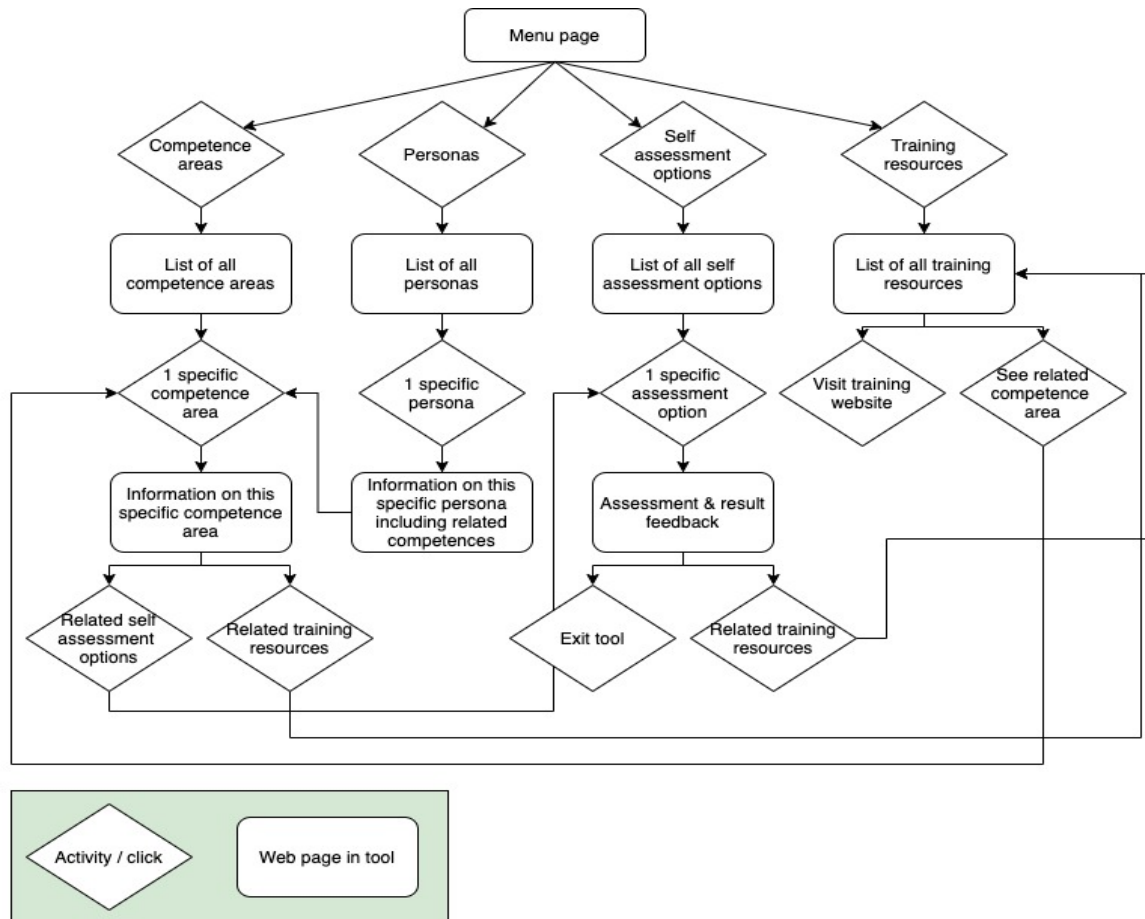


Figure Annex 12.6 Example of four different learner journeys

### Path 1. Competence areas

Marie, our generic data steward, wants to learn something about the competence area policy and strategy. Marie enters the tool and checks the competence areas menu to see if her training needs match a competence area. By doing that, she discovers that policy (and data archiving, another area of interest, based on her user story) is a main competence area.



#### User story (example) for tool usage:

- As a generic data steward, I can explain my organisational data management policy to the researchers, so that the researchers can meet the requirements
- As a generic data steward, I can explain the research needs to the infrastructure people, so that the organisational infrastructure optimally supports research
- As a librarian, I have the knowledge to support researchers with archiving, so data is findable

Figure Annex 12.7 User story generic data steward

1. Marie enters the tool.
2. Marie looks for a competence area connected to policy and policy requirements.
3. Marie clicks to learn about the different competence areas and levels (beginner, intermediate, advanced).
4. Marie clicks for self-assessment and runs the self-assessment test.
5. Marie identifies her own current level. Marie could be advanced and leaves the tool. Or Marie is not advanced yet and the tool suggests which skills might be developed.
6. Marie could leave the tool or clicks to see a list of suggested training resources.
7. Marie browses the training options and sees more information about the resources.
8. Marie leaves the tool.

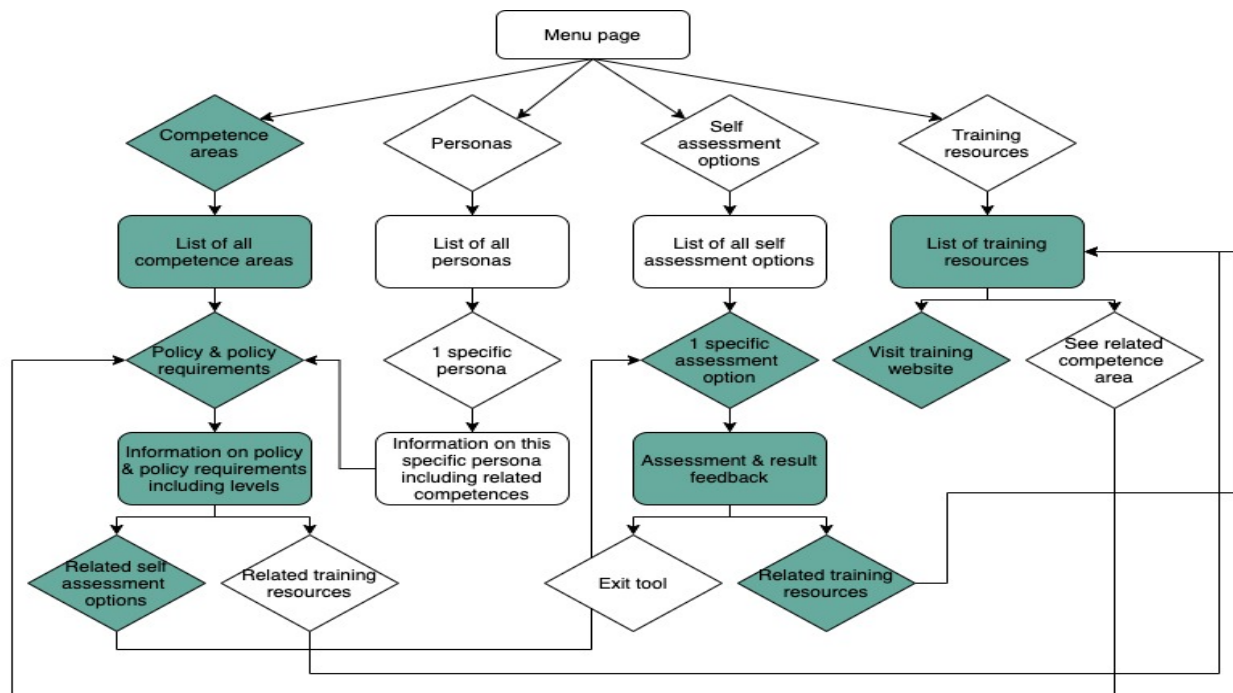


Figure Annex 12.8 Path 1, competence areas, example learner journey of a generic data steward

### Path 2. Personas

Wouter, our embedded data steward, wants to have an overview of the competences that are linked to his persona. This way, the tool helps him to assess his level of knowledge of the competence areas that fit his persona and could improve his career prospects (even discover potentially interesting future job profiles that better match his competences).

**EMBEDDED**



**User story (example) for tool usage:**

- As an embedded data steward, I can implement up-to-date knowledge, to enhance the quality of the data in the projects that I monitor
- As an embedded data steward, I can coordinate the data in a project, so it easier for my researchers to meet the requirements of the organisational data policy
- As an embedded data steward, I can use the tool, so I can assess my level of knowledge, improve my career prospects and identify potentially interesting job profiles for my future

Figure Annex 12.9 User story generic data steward

1. Wouter enters the tool.
2. Wouter clicks on I would like to see an overview of competences linked to my own profile or persona.
3. Wouter selects the profile or persona embedded data steward.
4. Wouter sees an overview of competences emphasising the embedded data steward.
5. Wouter clicks on a competence to read more about that competence.
6. Wouter sees an overview of training resources that have been mapped to the competence.
7. Wouter clicks on related training resources to find more information.
8. Wouter leaves the tool.

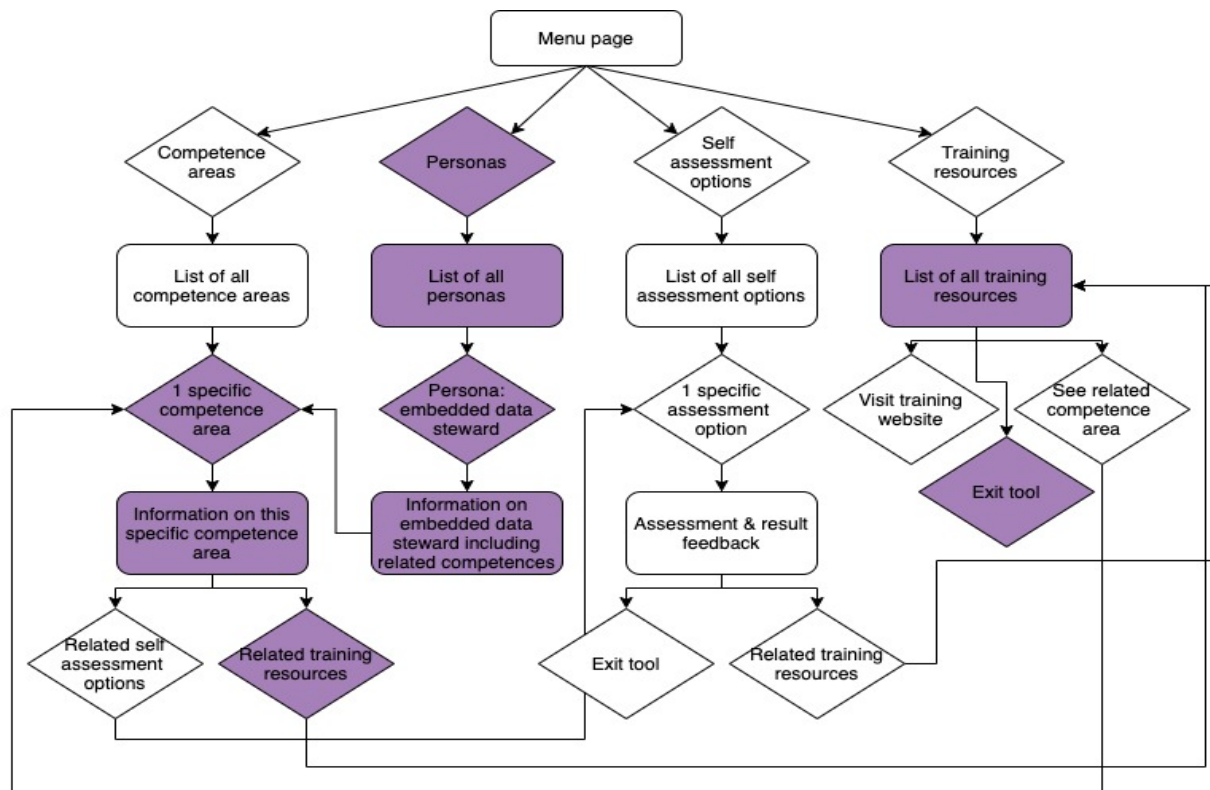



Figure Annex 12.10 Path 2, personas, example learner journey of an embedded data steward

### Path 3. Self-assessment options

Guillermo, our infrastructure data steward wants to assess his skills to identify knowledge gaps.

**INFRA-  
STRUCTURE**



**User story (example) for tool usage:**

- As an infrastructure oriented data steward, I can provide efficient software solutions, so that researchers can handle their data in a trustworthy and safe manner
- As a infrastructure oriented data steward, I can help metadata large datasets, to help FAIRify the data in my organisation's projects

Figure Annex 12.11 Path 3, user story infrastructure data steward

1. Guillermo enters the tool.
2. Guillermo clicks the menu option self-assessment.
3. Guillermo chooses one of the self-assessment tests.
4. Guillermo identifies his own current level, the tool suggests which skills might be developed and presents relevant training resources.
5. Guillermo sees an overview of training resources that matches his level (intermediate), and helps him to increase his knowledge, skills and abilities.
6. Guillermo clicks on related training resources to find more information.
7. Guillermo runs the self-assessment test and gets his new level (advanced).
8. Guillermo leaves the tool.

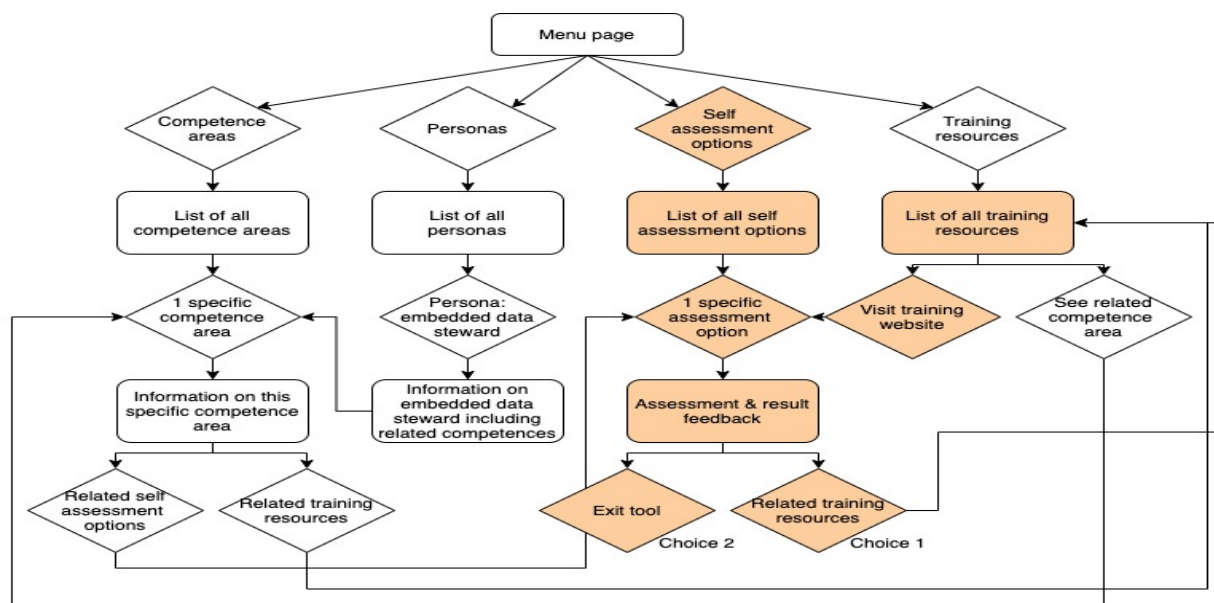


Figure Annex 12.12 Path 3, self-assessment options, example learner journey of a infrastructure data steward

#### Path 4. Training resources

Lucy, our research data steward, recently started her job and wants to have a general overview of training resources. This way, Lucy can become more skilled in computing data and FAIR data management, so that she can help her researchers with managing data in a project.

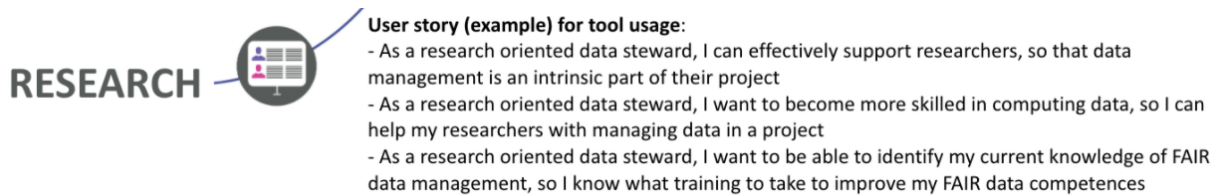


Figure Annex 12.13 User story research data steward

1. Lucy enters the tool.
2. Lucy clicks on I would like to see a list of training resources.
3. Lucy sees an overview of training resources.
4. Lucy browses the training options and sees more information about the resources.
5. Lucy clicks for self-assessment and runs the self-assessment test.
6. Lucy identifies her own current level, which is for some competence areas beginner. The tool suggests which skills she can develop.
7. Lucy clicks to see a list of suggested training resources.
8. Lucy leaves the tool.

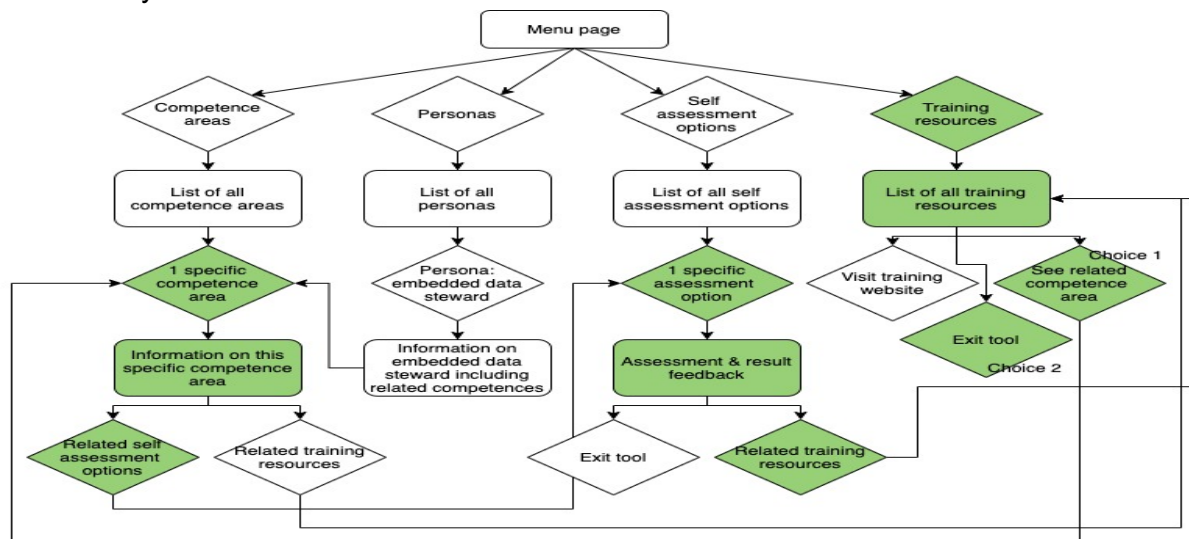


Figure Annex 12.14 Path 4, training resources, example learner journey of a research data steward

## Annex 13: Other features of a data steward skills tool

### Tool audience

Having an online competence and training tool would help the Dutch community of data stewards and employers by providing a clear single point of reference for all information related to data stewardship competences and training:

- Making visible an agreed, and up to date, set of competences against which institutions could plan their staff development strategy.
- Communicating changes to roles, competences and training opportunities, as the landscape rapidly evolves.
- Enabling individuals and an individual's employer to assess data stewardship skills and identify training opportunities, in order to plan and carry out career development.
- Enabling the Dutch community to identify gaps in current training opportunities that could be addressed by existing groups, companies or institutions.

A data steward skills tool could be useful for a wide range of stakeholders, such as individual data stewards, employers and training providers. With clear mandates for updating and maintaining the contents, it could evolve as the data steward landscape also evolves and remain a valuable tool for recruitment and career development of data stewards across the Netherlands, and beyond.

Individual data stewards could use the tool to:

- Use as a general information resource regarding competences and to find available training.
- Identify development areas and relevant training.
- Assess skills and knowledge, determine the level (beginner, intermediate, advanced) in different competence areas.
- Keep track of all opportunities and examples of best practice.
- Have clarity in potential career progression either within their employing institution, or in a different role elsewhere.

In addition to data stewards, their (future) employers could also benefit from the tool:

- To demonstrate that they have qualified data stewards, contributing to good data stewardship.
- To identify development areas and relevant training opportunities, in order to plan career development.
- To describe competences for new job profiles and assess potential candidates.

- To plan staff and development strategy with up-to-date competences.
- To communicate and stay up to date to changes in roles, competences and training opportunities, as the data stewardship landscape rapidly evolves.

Similarly, training providers could also use the tool:

- To showcase training opportunities in the tool.
- To identify gaps in current training provisioning.
- To identify what competences any future training opportunities might need to address.

### Content of the tool

This project focuses on the design and strategy of a data steward skills tool, i.e., the look and feel of the tool, how such a tool could facilitate the training of a data steward, and the strategy (umbrella) stakeholders and organisations could take to implement such a tool into the ambitions for training and education of data stewards.

The actual content of a data steward skills tool is only part of this project insofar as it concerns the strategy for the content, and not the content itself (although, if building on existing content, the overview of learning resources of Chapter 5 can be used as a basis for creating the content in a follow up process of building the actual tool by these (umbrella) stakeholders and organisations).

Topic wise, we build upon the earlier mentioned competence areas and the topics that derive from these areas (Chapter 4). Hence our earlier reference to competence areas in the proposed menu and learning path (Annex 12).

### Self-assessment

For all potential learners it is useful to assess where they 'stand' with regard to knowledge, skills and abilities. Given the diversity in backgrounds of data stewards, this is even more crucial. A self-assessment option in a skills tool could help learners to find the training that suits their needs and background. Moreover, self-assessment can be seen as a tool to stimulate intrinsic motivation and more meaningful learning (Macmillan & Hearn, 2008)<sup>154</sup>.

---

<sup>154</sup> McMillan, J.H. & Hearn, J (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. Educational Horizons, v87 n1, pp 40-49.



The outcome of the tool is, at least, that:

- The data steward has an idea of his or her level regarding one or more competences.
- The data steward identifies knowledge gaps.
- The data steward knows about the available training resources to work on these knowledge gaps.

Where should self-assessment be situated in the tool, to match self-assessment to the data steward's learner journey? If a data steward browses the skills tool, it should be really easy to start the self-assessment:

- The tool should be designed so that it facilitates the start of the learning path, via a call-to-action button, which starts the self-assessment when a data steward enters the tool.
- When browsing the tool, self-assessment should also be accessible via each of the competence areas. For those competence areas, a data steward should be able to self-assess his or her knowledge in combination with browsing resources that match the identified knowledge gaps on a competence area. Learners should also be referred to specific training resources, based on the result of their self-assessment.
- Self-assessment options should also match the end of a learning path, or at least the end of a specific competence-related learning path. The assessment is expected to confirm that the data steward has bridged an earlier knowledge gap.

The self-assessment functionality benefits from well thought-out visualisation. One possibility of visualising an assessment is the spider diagram with axes for the various competence areas. This offers a bird's eye view to what data stewards are learning and identify gaps in their knowledge. Alternatively, or in combination with a spider diagram, different levels could be distinguished, such as beginner, intermediate and advanced. Examples of how self-visualisation could be integrated in a data steward skills tool are given below.


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

Figure Annex 13.1 Example of a self-assessment question

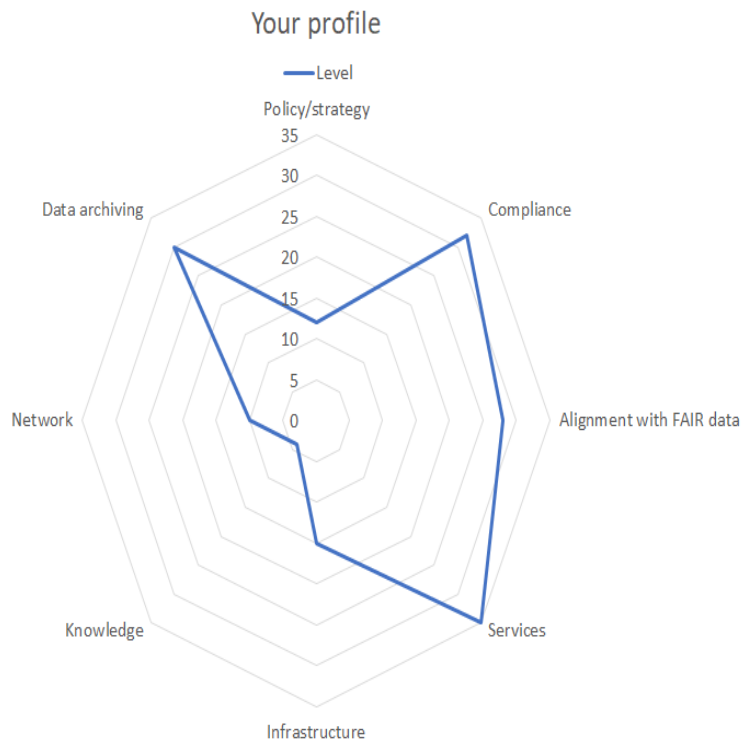
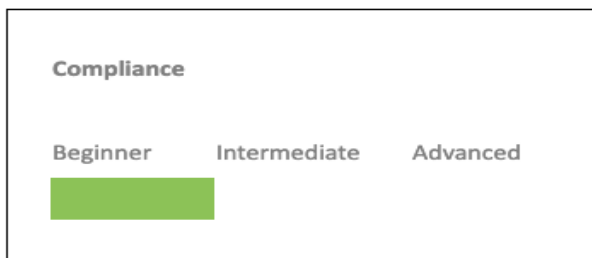
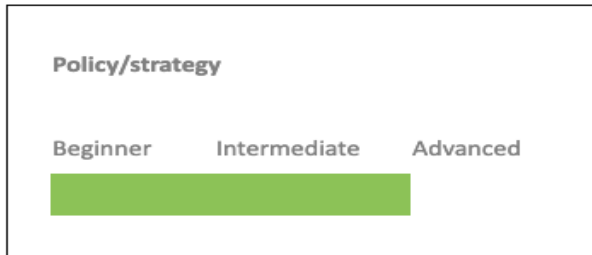


Figure Annex 13.2 Example of a self-assessment result page I



*Figure Annex 13.3 Example of a self-assessment result page II*

## Annex 14: Composition of project team

The project was initiated by the National Platform Open Science (NPOS) and coordinated on behalf of NPOS by Margreet Bloemers (ZonMw) and Margriet Miedema (LCRDM). Merlijn Bles (ZonMw) and Frederike Schmitz (DTL) were responsible for communication.

The project started in September 2019 and the final report was delivered in February 2021.

The following umbrella organisations participated on an in-kind basis in the project: VSNU (Darco Jansen), VH (Jort Diekerhof), NFU (Petra Drankier), PNN (Tess van Doorn), SURF (Annette Langedijk, Claudia Behnke) and ZonMw (Margreet Bloemers). Elevate Health (Nienke Verdonk) was asked to advise on the design and didactics of the digital data steward skills tool.

The daily project lead was done by Mijke Jetten and Celia van Gelder (DTL). To form the daily project team, the Dutch community of data professionals, including the LCRDM pool of experts<sup>155</sup>, was engaged.

The following organisations participated on an in-kind basis in the project team: Erasmus MC (Sabrina Gunput), Radboudumc (Mirjam Brullemans-Spansier, Peter-Bram 't Hoen), UMCG (Salome Scholtens), UMC Utrecht (Jasmin Böhmer), Eindhoven University of Technology (Iza Witkowska), Leiden University (Kristina Hettne, Joanne Yeomans), Maastricht University (Erik Jansen, Annemie Mordant), Radboud University (Inge Slouwerhof), TU Delft (Marta Teperek, Yan Wang), VU (Brett Olivier, Lena Karvovskaya), Avans University of Applied Sciences (Cora Bijlsma, Miaomiao Zhou), Fontys University of Applied Sciences (Masha Boosten-Ovtchinnikova), HAN University of Applied Sciences (Ingrid Busser), Saxion University of Applied Sciences (Renate Mattiszik), University of Applied Sciences Leiden (Mischa Barthel), DANS (Marjan Grootveld), GO FAIR Foundation (Mascha Jansen), Netherlands eScience Center (Mateusz Kuzak, Carlos Martinez), RDNL (Ellen Verbakel) and SURF (Carlos Teijeiro Barjas).

For more information about the project, you may contact Mijke Jetten ([mijke.jetten@dtls.nl](mailto:mijke.jetten@dtls.nl)).

---

<sup>155</sup> <https://www.lcrdm.nl/en/pool-of-experts-rdm-mailing-list>