Data Stewardship Career Paths: Recommendations of the EOSC Task Force Data Stewardship Curricula and Career Paths
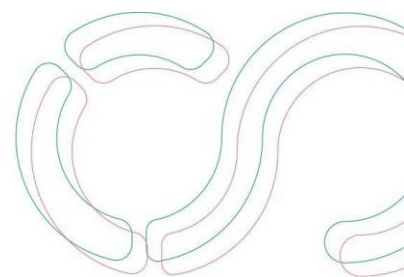
Authorship Community:

Tereza Kalová[1]
Francesca Frontini[2]
Laetitia Bracco[3]
Dunja Legat[4]
Joke Meeus[5]
Ilire Hasani-Mavriqi[6]

1. Vienna University Library, University of Vienna, Austria, ORCID: 0000-0002-1764-7228
2. CNR - Istituto di Linguistica Computazionale "A. Zampolli" & CLARIN ERIC, ORCID: 0000-0002-8126-6294
3. University of Lorraine, France, ORCID: 0000-0002-2939-9110
4. University of Maribor Library, Slovenia
5. Research Foundation Flanders, Belgium, ORCID: 0000-0001-5385-1747
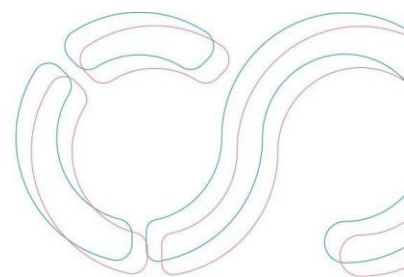6. Graz University of Technology, Austria, ORCID: 0000-0003-0758-0805

This report was open for community feedback in November 2023.

Table of Contents

Executive Summary

Data Stewardship Career Paths:
Recommendations of the EOSC
Task Force Data Stewardship
Curricula and Career Paths

Data Steward Career Paths

Reports, papers, initiatives, projects and surveys

Overview

EOSC Association Task Force
"Data stewardship, curricula and career paths"

Report

Recommendations

EOSC Association

- Keep current overview up-to-date
- Creation of an online "inventory" of Data Steward initiatives
- Ensure collaboration with international initiatives
- Ensure coordination among activities and studies carried out within EOSC Horizon Europe projects
- Promote and support the organisation of dedicated events.
- Assess the need for a professional network for Data Stewards

EOSC Projects

- Build on and extend the work of this task force

EOSC Partnership

- Establish a permanent Data Stewardship expert group

This document provides an overview of the topic of Data Stewardship Career Paths. Our review and summary of relevant reports and papers, ongoing initiatives, projects and surveys highlight the importance of ensuring more sustainable career paths for Data Stewards. The report further argues the need for further in-depth study and documentation of this topic. In particular, the analysis identifies relevant aspects that should be considered, ranging from employment conditions and salary to scientific recognition and roles. The report provides a list of recommendations and identifies activities that can be taken by the EOSC in the areas of Partnership, Association and Projects, summarised as follows:

1. The **EOSC Association** and related projects should ensure that the overview of the current situation is kept up-to-date as a reference point. A section on the EOSC Association public website should be dedicated to Data Stewardship initiatives including a dedicated bibliography. To ensure long-term sustainability, a governing body should own and maintain this "inventory" as a point of reference; future projects should be encouraged to use it as a reference and provide input.

2. The **EOSC Association** should ensure collaboration with international initiatives, in particular, the RDA IG Professionalizing Data Stewardship – TF Career Tracks and ensure coordination among current activities and studies carried out within the various EOSC Horizon Europe projects, and promote and support the organisation of dedicated events.

3. The **current and future EOSC projects** and initiatives in the field should build on the work of this task force, as well as on the results of the studies above, and extend them by
   a. Applying qualitative methods such as guided interviews or focus groups to investigate the aspects covered in section 6 ("Data Stewardship Careers - What Counts") in more depth
   b. Developing Data Steward *Personas* based on the proposed methodology detailed in Annex 1

4. The **EOSC Partnership** (EOSC Association, European Commission and Steering Board) should establish a permanent Data Stewardship expert group (including representatives from the various existing initiatives and this task force) with the following responsibilities:
   a. Develop and implement a monitoring framework that will allow the EOSC and other national and international institutions to support the career paths and development of personnel hired (at least in part) in Data Stewardship roles
   b. Advise the Association and the relevant (ongoing and future) projects and initiatives
   c. Issue recommendations on further activities of the Association regarding Data Stewardship

5. In consideration of the importance of the key role of Data Stewards in the development and implementation of the EOSC and to facilitate the exchange with and among Data Stewards, a further objective of the **EOSC Association** should be to assess the need for a professional network for Data Stewards, at least on the European level.

The use of an innovative methodology, the *Persona* workshops, is recommended alongside further surveys and in-depth interviews to explore the relevant aspects of career paths and

EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931
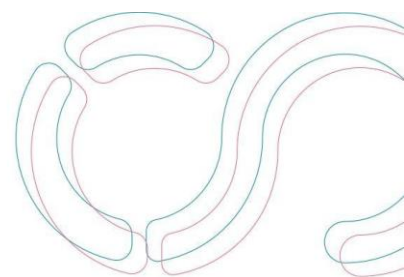
professional development of Data Stewards, including the roles and responsibilities of employers and Data Stewardship training programmes.

Context

This work was carried out within the Data Stewardship Curricula and Career Paths EOSC Task Force framework, particularly its *Career Paths* work stream.
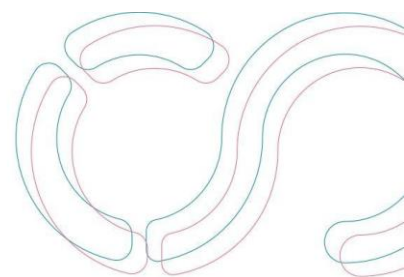
1. Introduction

The rationale for monitoring the career paths of Data Stewards stems from a simple consideration: Data Stewardship activities are essential for a healthy open science ecosystem, and a large number of Data Stewards will need to be trained and hired in the coming years in Europe (Mons, B. 2020). Alongside the appropriate capacity building, it is also important to make these careers interesting for the relevant communities, so that institutions can hire and retain Data Stewards.

Our understanding of "Data Stewards" is based on the two working definitions used by the task force[1] (see also the second deliverable of the task force: Basalti, C. et al. (2024) Recommendations for Data Stewardship Skills, Training and Curricula with Implementation Examples from European Countries and Universities). We will take a broad angle, thus considering professionals employed in Data Stewardship roles for a significant part of their time, independently from their official professional title.

The overall objective of the *Career Paths* work stream was to produce a report on how Data Steward roles map in an international, disciplinary and institutional context. The goal was to build on existing initiatives, projects and professional networks and issue recommendations for Data Steward career paths.

However, the initial discussions have led us to conclude that the current panorama on Data Stewardship is too scattered regarding training and careers. While many research institutions are hiring Data Stewards, there is not currently enough consensus on how these roles map onto existing professional positions within institutions. With most organisations focusing on training and embedding new Data Stewards in the institution, developing sustainable long-term career paths currently needs to be addressed systematically. Differences in approach among various countries and types of institutions (such as universities, libraries, scientific IT services, research performing organisations and research infrastructures) are also prevalent.

The relationship between Data Stewardship and research activities should be more precise. Researchers who fulfil the role of a Data Steward as part of their job may have different career path perspectives and issues concerning those hired solely as Data Stewards.

Crucially, Data Stewardship roles are too recent in many organisations and countries to allow for a diachronic investigation of the career paths of those hired in these positions.

Given these considerations, the possibility of producing a comprehensive report on the current panorama focusing on career paths was deemed premature. The Task Force believes that more research is needed to issue comprehensive recommendations. This document represents an intermediate step to present relevant initiatives and identifies key aspects to be considered. The core of our work consists in a review of the current Data Stewardship literature, in which we

---

[1] **"Data steward" definition 1**: A person [or a group of people] to ensure professional and careful treatment of data throughout all stages of a research process responsible for 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. *(Building on the Dutch Techcentre for Life Sciences and the definition listed in the glossary of OECD report Building digital workforce capacity and skills for data-intensive science (2020))*

"Data steward" definition 2: 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 to guarantee that data is appropriately treated at all stages of the research cycle (i.e., design, collection, processing, analysis, preservation, data sharing and reuse). *(From Jetten et al. 2021. Professionalising data stewardship in the Netherlands)*
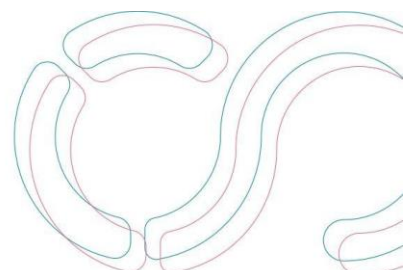
specifically highlight papers that address career paths-related issues. Also, the report references relevant ongoing projects and initiatives.

As for the long-term objectives, this document will provide recommendations for (a) gathering additional information on Data Steward's careers, (b) monitoring the development of this professional role at the European level and (c) facilitating the creation of a professional network at the European level.

Finally, our work is also based on the more general recommendations issued by the "The EOSC Skills & Training WG report" (Directorate-General for Research and Innovation (European Commission) et al. 2021), which indicates the precise definition of profiles and career paths for Open Science professionals is a primary driver for the development of Open Science. Making the position of Data Steward exciting and challenging is also indicated in the recommendations in the Dutch roadmap towards national implementation of FAIR Data Stewardship (Jetten et al. 2021).

2. Survey on Status of Data Stewardship

In exploring the current landscape, we considered existing documentation, ongoing projects, working groups and initiatives. While we cannot claim that the current review is exhaustive, it was based on the expertise and network of the members of the Task Force, with 28 members from 13 European countries.

We also monitored sister initiatives where our members are involved:

- RDA Interest Group - Professionalising Data Stewardship IG, specifically Career Paths Task Force
- OpenAIRE Task Force Data Steward

Finally, we have monitored closely the following EOSC projects which are engaging with similar issues:

- Skills4EOSC – This lists the definition of Data Stewards and open science-related profiles and curricula as one of its Key Exploitable results.
- EOSC Focus
- EOSC Future

3. Reports and Papers on Data Stewardship

Has the current literature on Data Stewardship engaged with the question of the career paths of Data Stewards? To answer this question, we conducted desk research, including any documents, articles or reports relating to Data Stewardship activities and roles that specifically touch upon the topic of career paths[2]. Here is a list of our findings, with the reference, the main contribution to the topic and main findings or recommendations.

1. Dederke Julian (2023), in "Towards Decentralised Data Stewardship – A perspective from ETH Zurich", advocates for a paradigm shift in research, promoting decentralised data stewardship and empowering existing practitioners". The author highlights the importance of recognising and empowering individuals with research data management (RDM) or data stewardship responsibilities rather than imposing central structures or introducing new

---

[2] For this survey, we limited our search to publications available in English; however we are aware of a number of relevant publications in other languages; such as Seidelmayer et al. (2023) "Forschung unterstützen - Empfehlungen für Data Stewardship an akademischen Forschungseinrichtungen: Ergebnisse des Projektes DataStew" in German that provides an interesting take on data steward career paths, including a study of training opportunities in Germany, an analysis of job adverts, and expert interviews
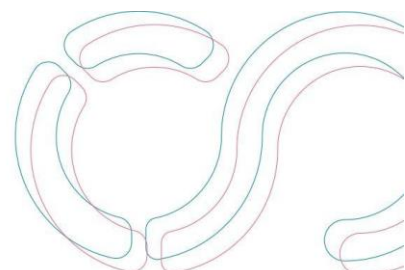
roles. By embracing this model, the author claims that research communities can tap into their expertise, foster a culture of effective data management, and pave the way for more responsible and impactful research in the future. Researchers can leverage their understanding of data stewardship to develop tailored approaches that best serve their needs. One of the key insights provided is that data stewardship should not be imposed from above but rather nurtured and empowered.

2. Patrick Helling, Felix Rau, Monika Linne, Jens Dierkes, Mijke Jetten, Anna Walek, & Magdalena Szuflita-Żurawska (2022), *The Importance of Demand and Environment for Defining and Establishing the Role of Data Stewards*
This workshop report presents different Data Stewardship programs in various institutions. The conclusion focuses on the fact that it is difficult to draw a line between Data Stewards, Data Librarians or Data Managers. But one element can be kept in mind for career paths: it is crucial to building professional communities of Data Stewards, regardless of their position in the institution, so that this role can be developed in the future.

3. Barr Moses (2022), *Data Stewards Have the Worst Seat at the Table*
This paper addresses the challenge of professional recognition of Data Stewards in a non-academic context, particularly in companies where data governance is central. Nevertheless, it is relevant for our career path reflection because data governance is very important for research. This article's important input must be underlined: "A modern data governance and stewardship approach must also go beyond *describing* the data to *understanding* its purpose". The Data Stewards must be recognised for their understanding of the data they manage; therefore, in an academic context, they need disciplinary training related to the scientific community in which they work.

4. Esther Plomp (2022), *A Data Steward journey*
In this feedback, the author explains how she became a Data Steward after following general training in research data management and basic programming skills. She explains, "A Data Steward can facilitate the exchange of data, identify gaps in services, provide insights in best practices and point researchers to existing tools *they can use*". This role is much closer to that of a Data Librarian than that of a Data Manager, which means it is still a sometimes vague concept.

5. Marta Teperek, Maria Cruz & Danny Kingsley (2022), *Time to re-think the divide between academic and support staff*
Data Stewards are generally considered members of the support, not the research staff. However, this distinction is not effective, for it introduces a status difference among researchers, which does not encourage collaboration. Indeed, support staff members often write papers, train students or manage projects – all of which could be considered the tasks of researchers. The careers relative to data management are not considered as interesting or valorising as research careers. Several ideas are suggested to tackle this issue: allow support staff to supervise scientists, to do research and unify remunerations.

6. Yuri Demchenko, Lennart Stoy (2021), *Research Data Management and Data Stewardship Competences in University Curriculum*
In this paper, an analysis was carried out on job advertisements related to Data Stewardship. The required skills are impressive: data management and governance, of course, but also data science engineering and domain knowledge. Data analysis is less required, for it is more often required of data scientists than Data Stewards. While this paper mostly concentrates on Data Stewardship curricula, it is important for our topic because it underlines the importance of scientific skills, which must be recognised in a career path.
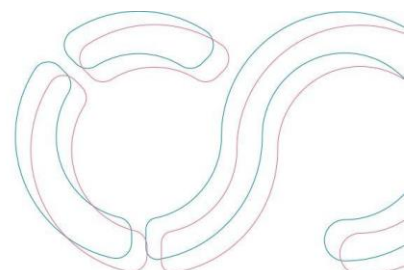
7. Reichmann et al. (2021), in "Between Administration and Research: Understanding data management practices in an institutional context", discusses the ambivalences and challenges associated with research data management (RDM). While there is a general agreement on the benefits of data sharing, the actual practices vary across different disciplines. The dual nature of RDM as both a research practice and an administrative task is explored. The uniqueness of data practices within each discipline should be acknowledged, and attempts to fit data management into predefined administrative categories may overlook these specificities. Researchers often perceive RDM as a separate task rather than integral to good research practice. Fluctuation within research teams, limited recognition and skills, and the perception that RDM takes away responsibilities from primary data producers directly impact career paths for data stewardship roles.

8. Daniel Bangert, Joy Davidson, Steve Diggs, Marjan Grootveld, Frans Huigens, Sanjin Muftić, Cristiana Pisoni, Hugh Shanahan, Lina Sitz, Ana Slavec, Sothearath Seang & Shanmugasundaram Venkataraman (2020), "The CODATA-RDA Data Steward school". This paper lists all the topics that should be included in a Data Stewardship curriculum. Students were recruited to test this pilot school for two weeks. The skills listed go far beyond the traditional skills required for a data librarian, for example, R, Git, machine learning… From the career paths perspective, it is important to remember that such skills require time and, therefore, recognition through a diploma or certification.

9. Shalini Kurapati (2019), *Becoming a data steward* post describes how the author became a Data Steward. From the career point of view, it is interesting to notice that her experience in this role was a way for her to develop a career outside academia after a PhD while using her research skills. Another important consideration relates to organisational strategies: it is not uncommon to assign Data Stewardship tasks, such as metadata quality control, to central library services. Data librarians, however, cannot, according to the author, replace Data Stewards who are embedded in the faculties and have the necessary disciplinary knowledge. It is a very important element for Data Steward's career paths, which may differ for the two types of profiles.
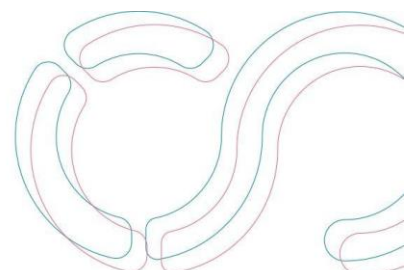
In addition to the reviewed papers, a dataset of 32 Data Steward related job vacancies was published by Thom Hartog and Yuri Demchenko (2022). These vacancies were collected from the *indeed.com* website using the keyword "data steward" and location "Amsterdam". It is interesting to note that many of these offers are from the private sector, and only a few require a Master's degree or a PhD as a qualification from applicants.

4. Ongoing Initiatives and Survey on Data Stewardship

Besides the literature review, the Task Force monitored various communication channels for initiatives to study Data Stewardship Career Paths. The most important studies are two surveys conducted in 2022. They are:

1) A survey conducted by the OpenAIRE Data Steward Task Force (WG RDM)
2) RDA Data Steward Career Tracks survey - Survey dataset published in November 2023, Report by Newbold et al. (2023) currently under review

The results of the OpenAIRE survey – based on a sample of 18 respondents – highlight the need to address the current lack of recognition for individuals in Data Stewardship roles, possibly correlating with the lack of official job descriptions at most of the respondents' research institutions. The results from the survey conducted by the RDA provide very useful insights on job titles, organisational context, contract types, educational background, alignment between educational background and Data Stewardship practice, and perceived career progression of respondents. The

EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

two open-ended questions in the survey focus on professional future and, through that, provide some insights on job satisfaction. 241 respondents were retained for analysis and their data is included in the published dataset. This survey sheds a new light on the field, adding the much-needed global perspective. Any future work should build upon this important report.

However, based on the points raised in our literature review, we notice that authors concentrate on some but not all aspects that the Task Force identifies as critical. In particular, job title, employment history and educational profiles are targeted, whereas job satisfaction and alignment of competencies and roles seem more difficult to tackle in a close-ended questionnaire. Both initiatives seem to foresee the possibility of follow-up interviews with Data Stewards that might address the missing points. Other qualitative approaches such as guided interviews or focus groups would benefit the common understanding of Data Stewardship career paths. A proposal for one such approach is presented in the appendix 1 of this report.

5. Main Conclusions and Recommendations for Further Action

As stated in the introduction, Data Stewardship is relatively new, and jobs take on very different shapes and forms within a very diverse academic landscape. It becomes clear from experiences on the work floor that there is a need to create more sustainable career paths for Data Stewards. Unsurprisingly, there is still little documentation and evidence on this matter, and most of the reports are testimonials or opinion pieces that reflect on indirect aspects and conditions of career paths as well as individual experiences. Nevertheless, taking the literature review as a starting point, it is clear that there are various challenges and expectations when considering career paths for Data Stewards. While new research might emerge in the future, we can already draw some interesting conclusions:

- Several papers have been published, as well as studies on the skills that Data Stewards must master, but there are very few (empirical) sources on career paths.
- A certain amount of vagueness still exists among the different roles and job titles: Data Librarian, Data Manager, Data Steward etc. This might present a challenge in both filling Data Stewardship positions and allowing for appropriate career progression.
- Based on the literature (2 and 9 in particular), Data Stewards rarely have equal footing with researchers, which can lead to frustration, as they often share the same academic background and experience. This may also hinder collaboration and prevent organisations from realising the potential of Data Stewards. In particular the thesis put forward in Teperek et al. (2022) highlights the need to rethink the divide between academic and support staff and recognize the collective value each brings to the research ecosystem.
- Data Stewards often have a strong disciplinary focus, which is perceived as crucial to understanding and processing the data they deal with and the domains they support.
- Data Stewardship roles often require technical skills, including basic computer programming and data science understanding. This implies specialised training for Data Stewards that is typically not part of the curricula in most disciplines[3].

6. Data Stewardship Careers - What Counts

These findings must be considered when establishing the scope of a Data Steward's activities, initial and lifelong training, role within the organisation and recognition. In particular, the review identifies clear training needs and specific skills for better recognition of Data Stewardship in the academic

---

[3] The last two items on the list are largely dependent on the model of Data Stewardship implemented (i.e. organisational context in which the data steward is active). The RDA IG PSD Current Models of Data Stewardship report (Ayres et al. 2022) also sheds some light on this.
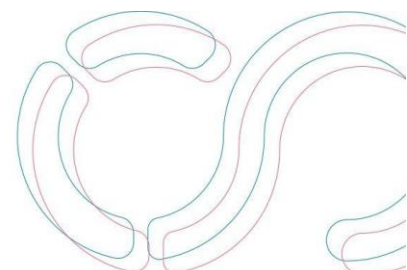
field and discipline-specific career paths. For this reason, studies on Data Stewardship career paths should take into account the following **aspects**:

- (Highest achieved) qualification (including whether they have a PhD)
- Employment History
- Official job title
- Role in the institution, function description (embedded/disciplinary vs generic)
- Satisfaction with salary
- Satisfaction with prospects (such as in terms of job security)
- Type of contract and funding
- Possibility to run (third-party funded) projects
- Leadership possibilities
- Authorship of publications
- Previous research activities (titles such as PhD had a role in the hiring process)
- Research role vs a support role
- Active in research (mixed role)
- Overall job satisfaction
- Evaluation of Data Stewards and Data Stewardship programmes
- Satisfaction with professional networking opportunities
- Experience with institutional ecosystem(s) – working relationships with the library, IT department, research infrastructures, legal department, tech transfer etc.
- Satisfaction with training and learning opportunities, both at the Data Stewardship level and disciplinary level

However, these preliminary findings can only be considered as a starting point, and an overarching, in-depth analysis of the specific needs and expectations from the perspective of Data Stewards is still missing. Better coordination and a broader set of questions might be necessary to understand the current situation regarding career paths. Also, qualitative studies, such as interviews or focus groups, could help explore the questions listed above. In particular, this task force believes that the methodology of *Persona* workshops (see the proposal in Annex 1) is an appropriate approach to analyse these aspects in depth.

Before designing sustainable career paths, it is first necessary to investigate the current needs and expectations of Data Stewards. It is key to better understand the general aspects of Data Stewards' careers: How would they prefer to shape their career? What are their essential requirements to do so? What are their expectations, and how can they be linked to the expectations of their institution/current job profile? These questions should be addressed in the context of the needs and requirements of the target groups as well as best practices in data stewardship.

7. Recommendations

In light of these considerations, the Task Force issues the following recommendations. Given our objectives and scope, the following list highlights activities that can be taken by the EOSC (partnership, association and projects). However, this focus is not meant to downplay the key role that institutions, such as ministries, or evaluation and funding agencies as well as research-performing organisations, such as universities and research infrastructures, play in ensuring the recognition of Data Stewards' careers and professional development at the regional, national and international level.
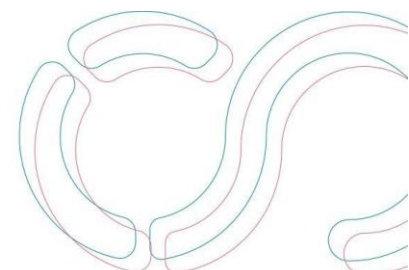
1. The **EOSC Association** and related projects should ensure that this overview is kept up-to-date as a reference point. A public website, as well as a dedicated Zotero library, could be created for this purpose. To ensure long-term sustainability, a governing body should own and maintain this "inventory" as a point of reference; future projects should be encouraged to use it as a reference and provide input.

2. The **EOSC Association** should ensure collaboration with international initiatives, in particular, the RDA IG Professionalizing Data Stewardship – TF Career Tracks and ensure coordination among current activities and studies carried out within the various EOSC projects, and promote and support the organisation of dedicated events.

3. The **current and future EOSC Horizon Europe projects** and initiatives in the field should build on the work of this task force, as well as on the results of the studies above, and extend them by
   a. Applying qualitative methods such as guided interviews or focus groups to investigate the aspects covered in section 6 ("Data Stewardship Careers - What Counts") in more depth
   b. Developing Data Steward *Personas* based on the proposed methodology detailed in Annex 1

4. The **EOSC Partnership** (EOSC Association, EC and Steering Board) should establish and support a Data Stewardship expert group (i.e., Opportunity Area 5: Skills, Training, Rewards, Recognition, Upskilling[4] including representatives from the various existing initiatives and this task force) with the following responsibilities:
   a. Develop and implement a monitoring framework that will allow the EOSC and other national and international institutions to support the career paths and development of personnel hired (at least in part) in Data Stewardship roles
   b. Advising the Association and the relevant (ongoing and future) projects
   c. Issuing recommendations on further activities of the Association regarding Data Stewardship

5. In consideration of the importance of the key role of Data Stewards in the development and implementation of the EOSC and to facilitate the exchange with and among Data Stewards, a further objective of the **EOSC Association** should be to assess the need for a professional network for Data Stewards, at least on the European level.

While the training of Data Stewards is an urgent and important task for the EOSC, as indicated by the conclusions of this Task Force, we urge the EOSC Association not to neglect the well-being and career paths of this new workforce.

A recent publication "What is data stewardship? Towards a comprehensive understanding" (Wendelborn et al. 2023) highlights the important role of Data Stewards and their "responsibilities towards the scientific community and the public to effectively manage data for reusability and reproducibility" but also towards the "legitimate interests of other affected stakeholders, for instance, data subjects or vulnerable groups". In a context where both research and the private sector are increasingly "data hungry" (as shown by the creation of the Data Spaces, also motivated by the need to fuel the fast advancements of Artificial Intelligence in many domains), the

---

[4] https://eosc.eu/oa5-skills-training-rewards-recognition-upscaling/
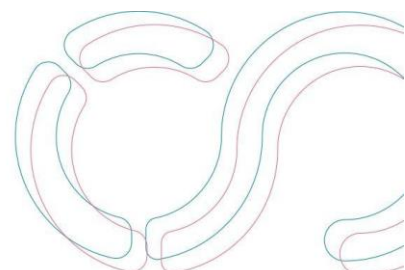
professional satisfaction and status of the individuals that are in charge of sharing and depositing data is crucial. This will ensure their ability to safeguard, respect and consider all stakeholders' rights and legitimate interests.

8. Annex 1 - Concept for Persona Workshops on Data Stewards Career Paths

Data stewardship is increasingly critical in today's research landscape, ensuring the effective management, preservation, and accessibility of data across various disciplines. To address the diverse career paths and skill sets required for Data Stewards, Persona Workshops could be the next step to facilitate interactive exploration and discussion on this profession's evolving demands. Persona workshops aim to bring together professionals, educators, and stakeholders to collaboratively identify key competencies, develop Persona profiles, and create tailored strategies for nurturing and supporting Data Stewards throughout their careers.

8.1 Persona Method

Persona is a well-known segmentation technique in IT development and design, as well as marketing and communication practices. A Persona represents patterns of representative users' behaviour, goals and motives, compiled in a fictional description of a single individual (Cooper, 2004). These Personas are presented as fictitious profiles that describe them in realistic detail, typically specified with names, photos, likes and dislikes, habits, backgrounds, expectations and needs (Blomquist & Arvola, 2002; LeRouge et al. 2013). Thus, Personas are not real, single, or average users. Instead, they visualise an explicit user category into a descriptive sketch that can be accessed and comprehended by both expert and non-expert groups.
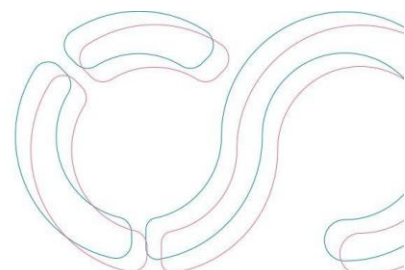
Personas are not a scientific method and are often based on assumptions and experiences of the professional expert group trying to conceptualise their (user) target groups. Nevertheless, they have proved useful as working hypotheses that help structure user-centred service development approaches. Following the belief that Personas need to be based on sound user/target group research, Pruitt and Adlin (2006) present the persona lifecycle that starts with family planning, which is followed by conception and gestation, birth and maturation, adulthood, and ends with lifetime achievement and retirement. In other words, personas should prove their functionality as strong working hypotheses that keep being tested with reality until they lose their function because the understanding of the target groups has reached the next phase.

8.2 Proposal: Workshop on Types of Data Stewards' Career Paths

We suggest applying the technique of Personas to gather a common understanding of the various types and roles in Data Stewardship, focusing on their careers. This approach would complement work previously done and described in the present report. Institutions and expert groups seeking to gather information on this topic can organise one or a series of such events. Quality of results will be ensured by 1) selecting experts from various national settings and institutional contexts and 2) securing preparation by gathering information and research on the target group, in this case, Data Stewardship.

Workshop participants: Participants are selected based on their expertise and knowledge of Data Stewardship. Expertise can be broad (having an overview of different types of stewardship in a specific context) or in-depth (having expertise on specific Data Stewardship roles and/or specific institutional contexts).

Preparatory information gathering: Building on their expertise, participants can be further informed

EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

by additional documentation concerning the target group. It is important that this information is broad and does not yet contain suggestions of specific segmentation so as not to influence the creation of the actual Personas that will take place during the workshop; examples are interview materials, results of surveys or descriptive lists of the target groups. The organisers can prepare participants based on the information from the OpenAIRE and RDA surveys and use cases from various countries. The following could be used to structure these examples.

8.2.1 Background

- Briefly describe the institution: University, technical university, applied university, research centre, research infrastructure, other
- Briefly describe the person filling out the use case

8.2.2 Description of Data Stewards

- How many "Data Stewards"[5] are there? Can you provide rough numbers on gender distribution, age range, and estimated mean age?
- What is their scientific background? Years of experience in research, PhD, scientific discipline etc.
- How were they trained as Data Stewards? Did they have relevant professional experience before this job? Was this done internally (via structured onboarding) or externally (informal/formal training)? What areas were covered by the training?

8.2.3 Description of Job Profile

- In which department/structure do the Data Stewards work? Library, Research coordination, management, research group, IT department, legal office, and others
- Do they have a permanent or fixed-term position? How are their positions funded (through internal or external funds)? Is the funding general or project-based? Is there a formal role in the functional classification system of the organisation/institute? One example could be the university job classification system used in the Netherlands.
- Are they called "Data Stewards"? If not, provide a term or a literal translation. Refer to these multilingual terminologies if applicable:
    - https://skosmos.loterre.fr/TSO/en/page/-GN16Q9CK-8

---

[5] **"Data steward" definition 1**: A person [or a group of people] to ensure professional and careful treatment of data throughout all stages of a research process responsible for planning and executing of all actions on digital data before, during and after a research project, to optimise the usability, reusability and reproducibility of the resulting data. *(Building on the Dutch Techcentre for Life Sciences and the definition listed in the glossary of OECD report Building digital workforce capacity and skills for data-intensive science (2020))*
"Data steward" definition 2: 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).
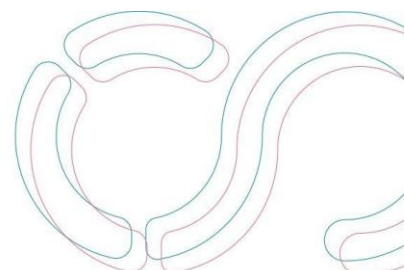*(from Jetten et al. 2021. Professionalising data stewardship in the Netherlands).*
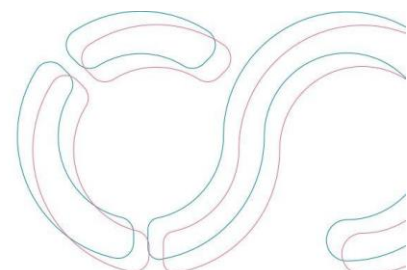
- ○ https://vocabs.sshopencloud.eu/vocabularies/sshocterm/en/page/data_steward_1 18
- How would you describe their main job functions/profile: Such as advising and support, training, needs assessment, requirements engineering, or other?
- Are all their tasks RDM-related, or do they have additional responsibilities or roles within the organisation (e.g. as research engineers, librarians, trainers)? Is their position full or part-time (such as 20% Data Steward and 80% research)? Are the RDM-related tasks considered in their evaluation/yearly review?
- How can you best describe their type of position? Do they have a research position? Do they have an administrative position such as research support? A management position? An IT/tech position? A mix of the above (include ratio)?
- What is their position in the institutional hierarchy? Do they actively shape their role and activities within the institution, or are their actions more service-related (such as responding to questions from researchers)? Do they have a position where they can suggest and/or initiate improvements to institutional processes? Do they have a mandate within their institution to propose structural changes or optimise procedures? Can they speak with the leadership? Who do they report to?
- How do they promote the existence of Data Stewards at the institution (such as via a website or information sessions)
- Where on the generic to domain-specific Data Stewardship scale is their role situated?
- Is there a provision for career paths at their institution? What kind of support for career development can they take advantage of?
- What are their aspirations? Are they aiming to stay in this role or want to grow within or outside the institution?
- What steps (if any) are they taking regarding career development?

The organisers should provide a 1-2 page summary for each use case. Alternatively, they can make a short video covering the main points. The use cases can constitute valuable material in their own right and should be considered for publication with the informed consent of the interested parties.

8.2.4 Workshop Methodology

The workshop methodology should follow the general principles of *Persona* workshops. First, participants hold discussions (in subgroups) about tendencies they perceive based on the preparatory documents provided by the TF (the developed use cases but also this deliverable and the results of the surveys). This will serve as a warm-up and a way to get to know fellow participants. The participants will start to create common insights (on post-its). Afterwards, participants will work in subgroups to elaborate on some rudimentary personas by clustering characteristics of typical Data Steward profiles and roles. Following this activity, these *Personas* will be presented and discussed with the whole group. Deep democracy techniques will be used to select the most relevant personas (5 to 7). Lastly, these personas will be enriched with more details in subgroups. In this exercise, it is important to provide the various personas with specific details like names, gender, age, and career path-related aspects like goals, intentions, aspirations, frustrations and needs. Working groups can rotate so that different people work on the individual personas. Working

EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

time is estimated at one day with 15 to 20 people. The workshop can occur online (see the following example), but we expect a physical meeting to be more productive. Another option is to organise the first part as an online workshop with a face-to-face follow-up.

8.3 Possible Outcome

The various *Personas,* as well as the reflections from the participants, could be detailed in a dedicated publication. Working hypotheses could be made about future actions to optimise Data Steward's training and career paths. These working hypotheses could create a basis for subsequent research, such as interviews addressing specific aspects identified in the *Personas*.
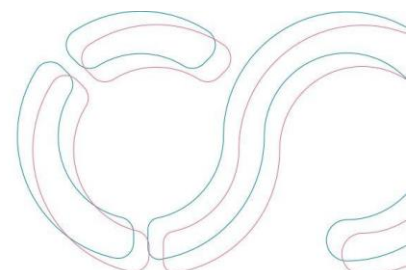
9. References

1. Adlin, Tamara, John Pruitt, Kim Goodwin, Colin Hynes, Karen McGrane, Aviva Rosenstein, and Michael J Muller. 2006. 'Putting Personas to Work'. In CHI'06 Extended Abstracts on Human Factors in Computing Systems, 13–16.

2. Ayres, B., Lehtsalu, L., Parton, G., Száldobágyi, Á., Warren, E., Whyte, A., & Zimmer, N. (2022). RDA Professionalising Data Stewardship - Current Models of Data Stewardship: Survey Report (1.0). Zenodo. https://doi.org/10.15497/RDA00075

3. Bangert, Daniel, Joy Davidson, Steve Diggs, Marjan Grootveld, Frans Huigens, Sanjin Muftić, Cristiana Pisoni, et al. 2020. "The CODATA-RDA Data Steward School". Dublin, Ireland, January 15th. https://doi.org/10.5281/zenodo.3609205.

4. Basalti, C., Fazekas-Paragh, J., Forni, M., van Gelder, C., Hasani-Mavriqi, I., Janik, J., Kalová, T., Kuchma, I., Lindroos, H., Lütcke, H., Pinnick, J., Raga, N., Thorpe, D., & Wildgaard, L. (2024). Recommendations for Data Stewardship Skills, Training and Curricula with Implementation Examples from European Countries and Universities. Zenodo. https://doi.org/10.5281/zenodo.10573892

5. Blomquist, Åsa, and Mattias Arvola. 2002. "Personas in action: Ethnography in an interaction design team". In Proceedings of the second Nordic conference on Human-computer interaction, 197–200.

6. Cooper, Alan. 2004. The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity. Second Edition. Indianapolis, IN: Sams.

7. Dederke, Julian. 2023. Towards Decentralised Data Stewardship – A perspective from ETH Zurich. https://doi.org/10.3929/ethz-b-000605802 https://ethz.ch/staffnet/en/news-and-events/internal-news/archive/2023/01/start-des-data-stewardship-netzwerks-an-der-eth-zuerich.html

8. Demchenko, Yuri, and Lennart Stoy. 2021. "Research Data Management and Data Stewardship Competences in University Curriculum". Vienna, Austria (Virtual), April 21st. https://doi.org/10.5281/zenodo.4633752.

9. Directorate-General for Research and Innovation (European Commission), EOSC Executive Board, Natalia Manola, Emma Lazzeri, Michelle Barker, Iryna Kuchma, Vinciane Gaillard, and Lennart Stoy. 2021. Digital Skills for FAIR and Open Science: Report from the EOSC Executive Board Skills and Training Working Group. LU: Publications Office of the European Union. https://data.europa.eu/doi/10.2777/59065.

10. Jetten, Mijke, Marjan Grootveld, Annemie Mordant, Mascha Jansen, Margreet Bloemers, Margriet Miedema, and 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.4623713.

11. Gurdal, Gultekin and Yavuz, Begüm. 2023. OpenAIRE Survey: Who is Managing Research Data? 11th May 2023. https://www.openaire.eu/openaire-survey-who-is-managing-research-data
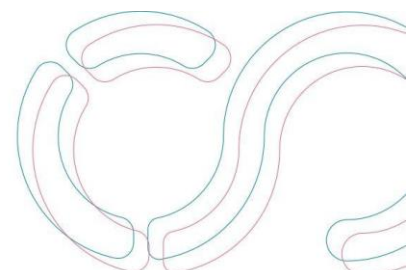
12. Helling, Patrick, Felix Rau, Monika Linne, Jens Dierkes, Mijke Jetten, Anna Walek, and Magdalena Szuflita-Żurawska. 2022. "The Importance of Demand and Environment for Defining and Establishing the Role of Data Stewards". Zenodo. https://doi.org/10.5281/zenodo.6511185.

13. Hartog, Thom and Demchenko, Yuri. Data Steward Professional: Reference dataset of Data Steward-related job vacancies for competencies assessment (Version 1) [Data set]. 2022. Zenodo. https://doi.org/10.5281/zenodo.6008122

14. Kurapati, Shalini. 2019. "Becoming a Data Steward". Impact of Social Sciences (blog). 30th April 2019.
https://blogs.lse.ac.uk/impactofsocialsciences/2019/04/30/becoming-a-data-steward/.

15. LeRouge, Cynthia, Jiao Ma, Sweta Sneha, and Kristin Tolle. 2013. "User Profiles and Personas in the Design and Development of Consumer Health Technologies". International Journal of Medical Informatics 82 (11): e251–68.

16. Mons, Barend. 2020. "Invest 5% of research funds in ensuring data are reusable". Nature. 20th September 2022. https://doi.org/10.1038/d41586-020-00505-7.

17. Moses, Barr. 2022. "Data Stewards Have The Worst Seat At The Table". Medium. 20th May 2022. https://towardsdatascience.com/data-stewards-have-the-worst-seat-at-the-table-e115ec65b639.

18. Lehtsalu, Liise, Valerie McCutcheon, Elizabeth Newbold, Yan Wang, and Biru Zhou. 2023. "Research Data Alliance Interest Group Professionalising Data Stewardship Career Tracks Survey Dataset". Zenodo. https://doi.org/10.5281/zenodo.10117910.

19. OECD. 2020. "Building Digital Workforce Capacity and Skills for Data-Intensive Science." Paris: OECD. https://doi.org/10.1787/e08aa3bb-en.

20. Plomp, Esther. 2022. "A Data Steward Journey". Open Working (blog). 22nd April 2022. https://openworking.wordpress.com/2022/04/22/a-data-steward-journey/.

21. Reichmann, S., Klebel, T., Hasani-Mavriqi, I., Ross-Hellauer, T. 2021. Between administration and research: Understanding data management practices in an institutional context. Journal of the Association for Information Science and Technology, 72(11), 1415– 1431. https://doi.org/10.1002/asi.24492

22. Seidlmayer, Eva, Fabian Hoffmann, Jens Dierkes, Birte Lindstädt, Ralf Depping, and Konrad Ulrich Förstner. 2023. "Forschung unterstützen - Empfehlungen für Data Stewardship an akademischen Forschungsinstitutionen : Ergebnisse des Projektes DataStew". https://repository.publisso.de/resource/frl:6441397.

23. Teperek, Marta, Maria Cruz, and Danny Kingsley. 2022. "Time to Re-Think the Divide between Academic and Support Staff". Nature, https://doi.org/10.1038/d41586-022-01081-8.

24. Wendelborn, Christian, et al. "What Is Data Stewardship? Towards a Comprehensive Understanding". Journal of Biomedical Informatics, vol. 140, April 2023, p. 104337. ScienceDirect, https://doi.org/10.1016/j.jbi.2023.104337.