Practical course on FAIR Data Stewardship in Life Science

Professionalising Data Stewardship in the Netherlands

Mijke Jetten, Celia van Gelder
DTL, ELIXIR-NL, Health-RI
March 9, 2022

Outline

- Quick intro
- Learning goals
- The need for skills and capacity building in data stewardship
  - Capacity needed
  - Data stewardship roles and competences
- The Netherlands
  - Data stewardship roles and responsibilities
  - Data stewardship training
  - Data stewardship community building
- ELIXIR-CONVERGE
- Other players in the ecosystem
  - EOSC competences and skills
  - RDA interest groups
- Takeaways
- Discussion with the audience

Slides: https://doi.org/10.5281/zenodo.6341447

health RI enabling data driven health
elixir CONVERGE
elixir NETHERLANDS
DTL DUTCH TECHCENTRE FOR LIFE SCIENCES
national programme open science

Quick intro: who are we

- Celia van Gelder
  - Training Programme Manager (DTL, ELIXIR-NL, Health-RI)
  - WP lead for Training & Capacity Building (WP2) in ELIXIR-CONVERGE
  - Member Data stewardship Curriculum and Career Paths Task Force (EOSC Association)

- Mijke Jetten
  - Community Manager Data Stewardship (DTL)
  - Programme Manager FAIR Data (Health-RI)

Learning goals

After this presentation …

- you are able to position yourself in the data steward landscape
- you are aware of tools/resources with information on the required knowledge, skills and abilities
- you are aware of training resources that help you acquire this knowledge, skills and abilities
- you understand how community building may help you to become a better data steward
- and you might even be inspired to start/participate in similar communities yourself

Data-related roles

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[43]

OECD report Building digital workforce capacity and skills for data-intensive science (2020)

Table from our NPOS report (page 21)

nationaal programma open science

Our focus today: data stewardship

- **Data stewardship:** Responsible planning and executing of all actions on digital data before, during and after a research project, with the aim of optimizing the usability, reusability and reproducibility of the resulting data

- Data stewardship and data management skills are **essential in research**
  - The lack of consensus on the responsibilities, knowledge, and skills of data stewards hampers building adequate data steward capacity

- **Professionalisation of data stewardship is needed!**
  - (Inter)national alignment and coordination are needed to achieve coherent training/ education, accompanied with a consistent human resource (HR) policy

Role	Task	FTEs needed per 1000 researchers
Data Steward	Assisting researchers with effective management of research data	26
Trainer on data stewardship	Training researchers on data management skills	4

WORLD VIEW · 25 FEBRUARY 2020

Invest 5% of research funds in ensuring data are reusable

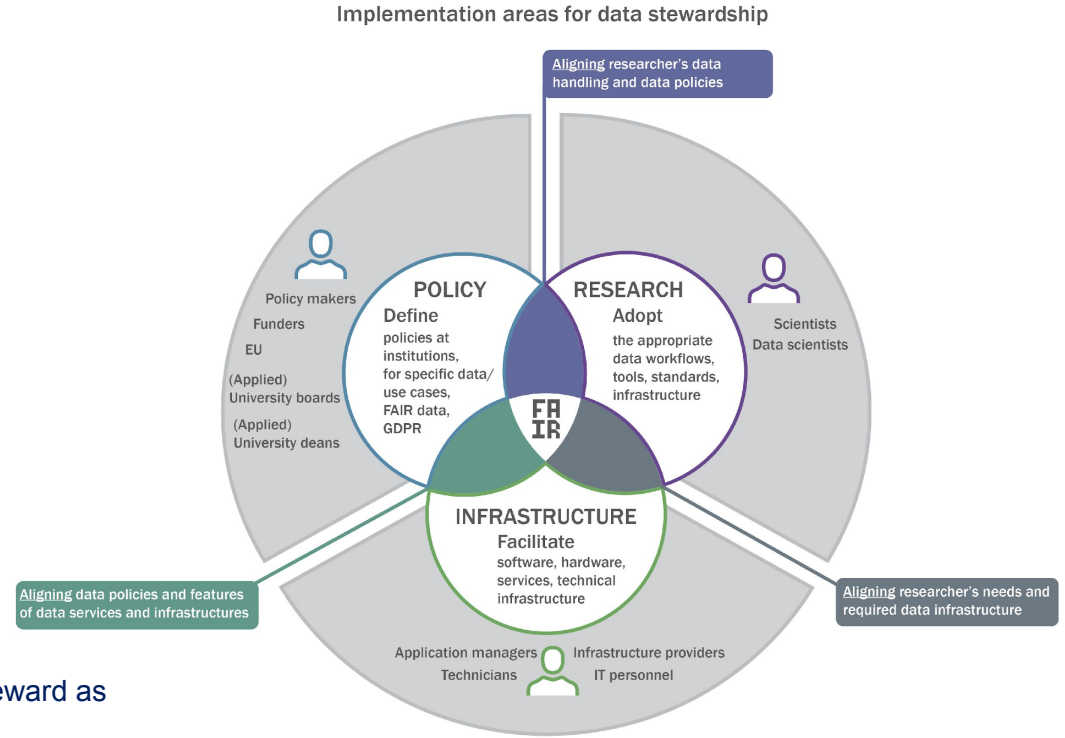
It is irresponsible to support research but not data stewardship, says Barend Mons.

Barend Mons

https://doi.org/10.1038/d41586-020-00505-7

Data Steward Roles in the Data Steward Landscape

- **Data Steward Policy**
  institute and policy focused
- **Data steward Research**
  project and research focused
- **Data Steward Infrastructure**
  data and e-infrastructure focussed



https://doi.org/10.5281/zenodo.4461378

- ZonMw/ELIXIR-NL funded project "Towards FAIR Data Steward as profession for the Life Sciences"
- All project output: https://zenodo.org/communities/nl-ds-pd-ls/
- Final report (2019): https://doi.org/10.5281/zenodo.3471707
- Matrices: https://doi.org/10.5281/zenodo.3239079

Radboudumc umcg Radboud University UMC Utrecht DTL DUTCH TECHCENTRE FOR LIFE SCIENCES elixir NETHERLANDS NFU NEDERLANDSE FEDERATIE VAN UNIVERSITAIR MEDISCHE CENTRA ZonMw

NPOS/ELIXIR Data Stewardship Competency Framework

8 Competence Areas

For 3 Data Steward roles and 8 competence areas:

- Responsibilities & tasks
- Knowledge, skills and abilities (KSAs)
- Learning objectives (including Bloom's levels)

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

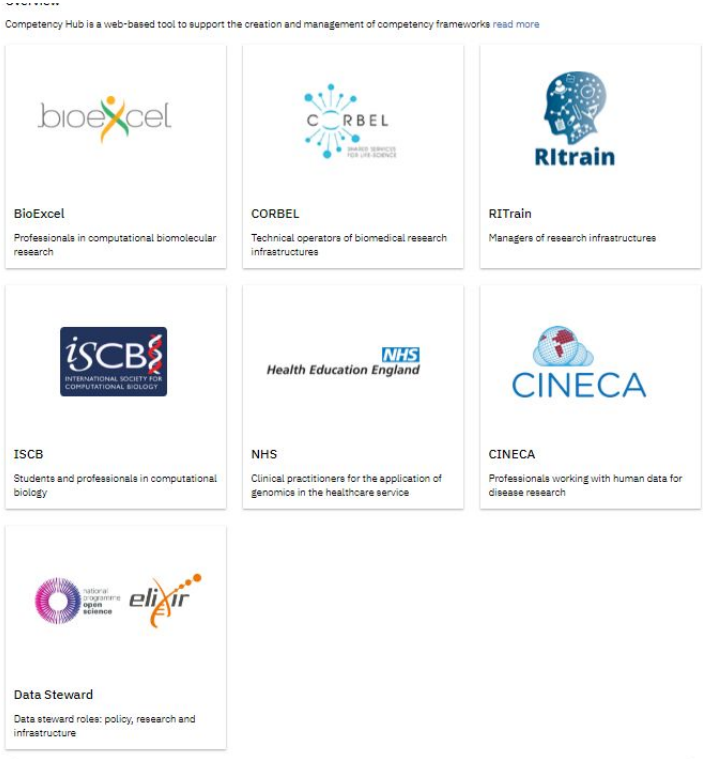
Competence area	This concerns
Policy/strategy	Development, implementation and monitoring of research data management policy and strategy for the research institute
Compliance	Compliance to the Netherlands Code of Conduct for Academic Practice, the Netherlands Code of Conduct for Research Integrity, the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

Competency Frameworks for Data Management/Stewardship

- NPOS/ELIXIR Data stewardship Competency Framework
  - responsibilities, activities, KSAs and Learning Objectives of three data steward roles matrices
  - since 2021also in Competency Hub https://competency.ebi.ac.uk/

- FAIR4S framework EOSCPilot D7.5 https://eoscpilot.eu/content/d75-strategy-sustainable-development-skills-and-capabilities

- Frameworks discussed in FAIRsFAIR D7.2 Briefing on FAIR Competences and Synergies https://doi.org/10.5281/zenodo.4009006



Data stewardship roles and responsibilities

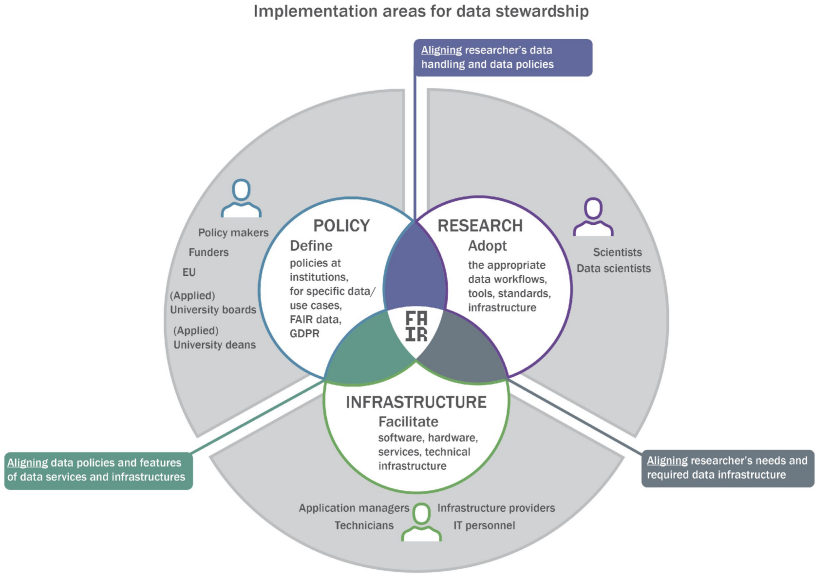


Figure 4.2 Basic job profile components of a data steward

national programme **open science**

ZonMw/ELIXIR-NL funded project "Towards FAIR Data Steward as profession for the Life Sciences"

- All project output: https://zenodo.org/communities/nl-ds-pd-ls
- Final report (2019): https://doi.org/10.5281/zenodo.3471707
- Matrices: https://doi.org/10.5281/zenodo.3239079
- Since 2021 also in the https://competency.ebi.ac.uk

NPOS project "Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship"

- Final report (2021): https://doi.org/10.5281/zenodo.4320504

Radboudumc umcg Radboud University UMC Utrecht DTL DUTCH TECHCENTRE FOR LIFE SCIENCES elixir NETHERLANDS NFU NEDERLANDSE FEDERATIE VAN UNIVERSITAIR MEDISCHE CENTRA ZonMw

What does a data steward do?

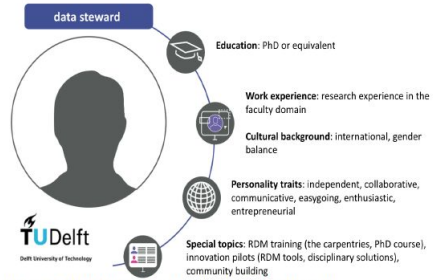


Figure 3.1 Delft University of Technology data steward

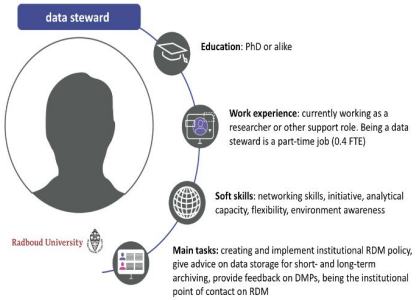


Figure Annex 2.3 Radboud University data steward

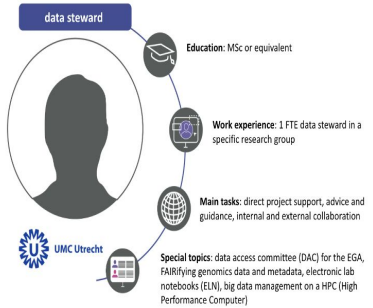


Figure Annex 2.5 UMC Utrecht data steward

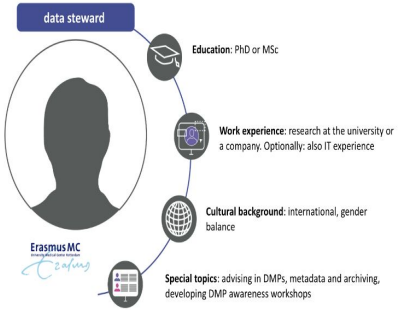


Figure Annex 2.7 Erasmus MC data steward

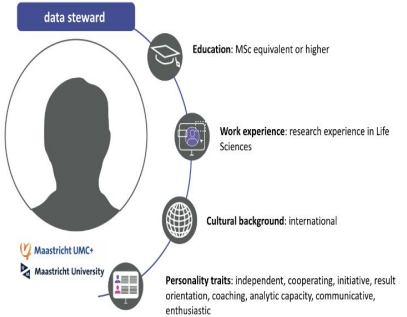


Figure Annex 2.9 Maastricht UMC data steward

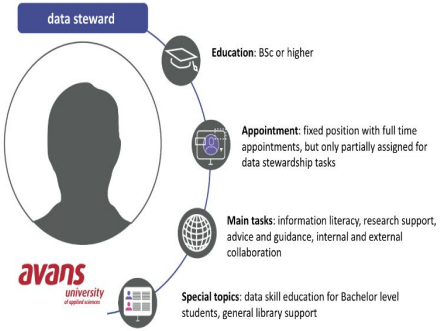


Figure Annex 2.11 Avans University of Applied Sciences data steward

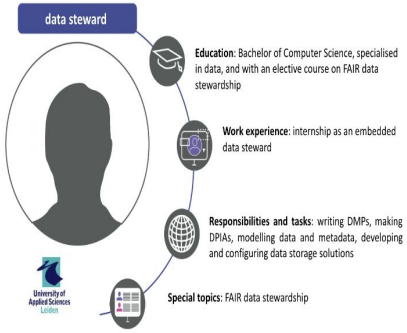


Figure Annex 2.13 University of Applied Sciences Leiden data steward

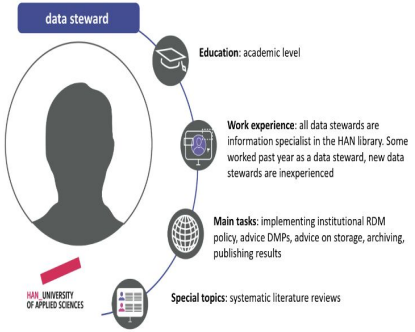


Figure Annex 2.15 HAN University of Applied Sciences data steward

Read these data stewards' full stories at https://doi.org/10.5281/zenodo.4320504

national programme open science

Data steward personas

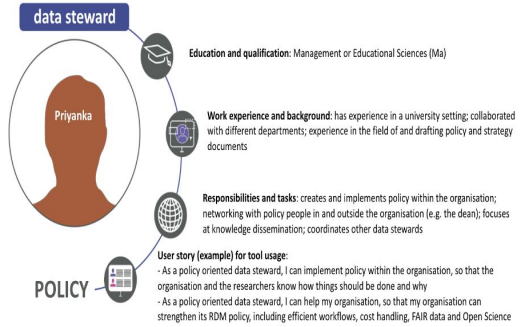


Figure Annex 12.1 Policy data steward persona (individual perspective)

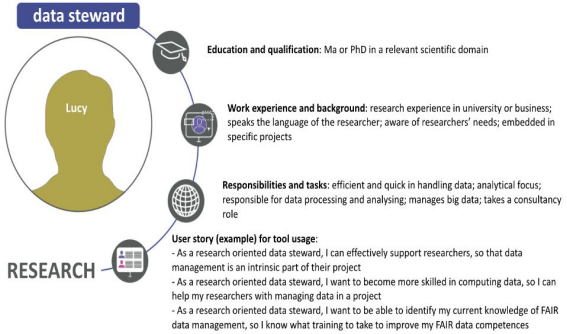


Figure Annex 12.2 Research data steward persona (individual perspective)

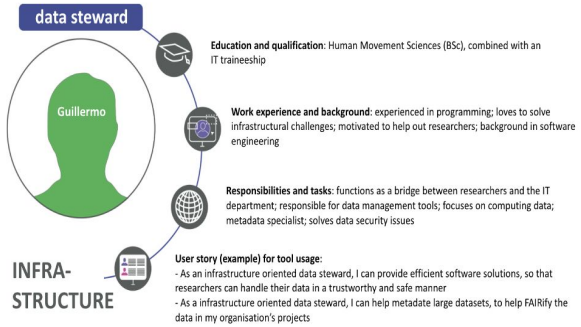


Figure Annex 12.3 Infrastructure data steward persona (individual perspective)

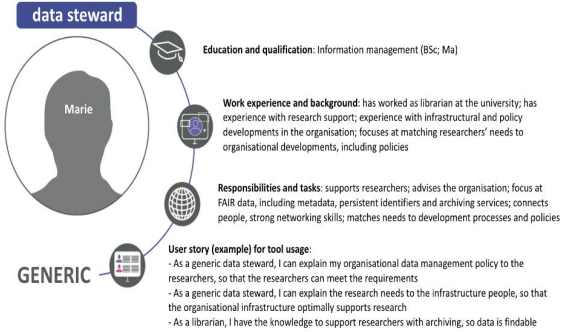


Figure Annex 12.4 Generic data steward persona (organisational perspective)

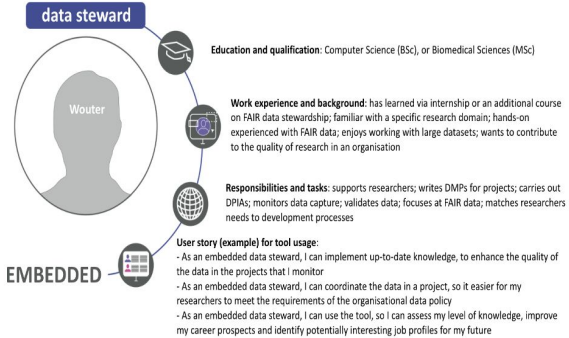


Figure Annex 12.5 Embedded data steward persona (organisational perspective)

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

national programme open science

Data steward competences

Network: Obtaining and maintaining a network of aligned expertise areas and relevant organisations by the institute

› [Policy Oriented Data Steward] Responsible for obtaining and maintaining a network of aligned expertise areas and relevant departments and organisations inside and outside the institute with regard to research data management

⌄ [Research Oriented Data Steward] Responsible for liaison and alignment of research data management within the department or project group and with relevant stakeholders outside the department or project group

Activities and tasks

More details

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

Knowledge, skills and abilities (KSAs)

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

Learning objectives (LOs)

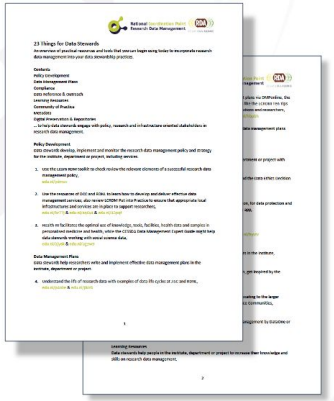
- Explain where to find department or project stakeholders and researchers (remembering)
- Liaise with department or project stakeholders and researchers (applying)
- Establish an active network in which regular consultations are taking place (evaluating)

https://competency.ebi.ac.uk

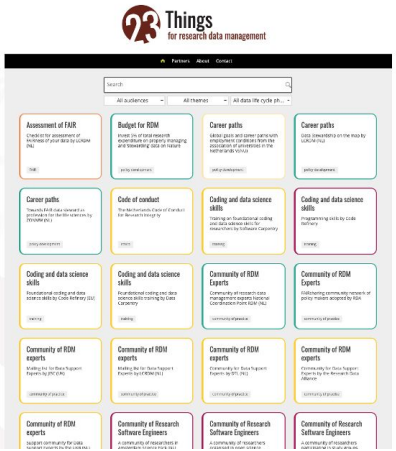
Data steward tool

RDA Output

National Coordination Point Research Data Management



Digital sheets for training purposes
https://doi.org/10.5281/zenodo.3773663



Toolkit (beta version)
https://23things.sites.uu.nl

Filter by
Audience (7 options)
Theme (17 options)
Data life cycle phase (6 options)

Audiences
- Researcher
- Student
- Librarian
- Data steward
- IT staff
- Research Software Engineer
- Policy maker

Work in progress

https://www.lcrdm.nl/en/23things

Developing and delivering FAIR data stewardship training for researchers and data stewards

- Helis Academy course “FAIR Data stewardship for life sciences”
  - https://www.aanmelder.nl/fair-data-stewardship-2021
- ELIXIR-NL Technical Data Stewardship Course (2019)
- Data stewardship training for Early Career Researchers from Marie Curie ITNs (2020, 2021)
- Various data stewardship webinars and metadata workshops for data stewards and project leads of ZonMw granted Covid-19 projects (2020)
  - https://www.health-ri.nl/workshops-delivering-fair-metadata-covid-19-data-portal
- All our modules have DOIs and are available for re-use

health RI enabling data driven health

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

elixir NETHERLANDS

Helis course FAIR data stewardship

https://www.aanmelder.nl/fair-data-stewardship-2021/part_program
https://tess.elixir-europe.org/events/helis-course-fair-data-stewardship

3rd edition of *FAIR data stewardship (for the life sciences)*

Did you ever wonder how to improve your data handling in your daily research practice? Want to improve the FAIRness of your data?

Did you recently start a job as data steward? Are you thinking of starting a career as a data steward in the life sciences in academia or industry?

Join this introductory course to FAIR data stewardship in the life sciences!

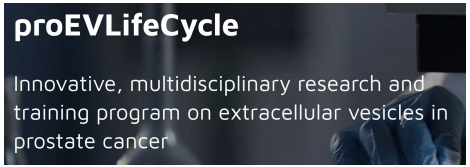
Course topics

We will pass the stages of the data life cycle in more detail in the training modules of the course. The following topics will be discussed at an introductory level:

- Data stewardship competency framework & the FAIR data stewardship landscape
- (Reviewing) data management plans
- Informed consent procedures
- Data discovery and capturing data
- Preregistration
- Data security and privacy
- Infrastructure for storing and sharing data
- Tools for processing and analysing data
- Organising, versioning and documenting data
- Data and software carpentry
- Archiving data
- Data rights
- FAIR data
- Metadata & ontologies

Related training materials:

- Helis Academy course FAIR data stewardship 2021, Day 1, Introduction Landscape Wrap up
- Helis Academy course FAIR data stewardship 2021, Day 1, Institute requirements
- Helis Academy course FAIR data stewardship 2021, Day 1, Funder requirements
- Helis Academy course FAIR data stewardship 2021, Day 1, Publisher requirements
- Helis Academy course FAIR data stewardship 2021, Day 1, (Reviewing) Data Management Plans
- Helis Academy course FAIR data stewardship 2021, Day 2, Introduction Wrap up
- Helis Academy course FAIR data stewardship 2021, Day 2, Informed consent procedures
- Helis Academy course FAIR data stewardship 2021, Day 2, Finding and capturing data part 1
- Helis Academy course FAIR data stewardship 2021, Day 2, Finding and capturing data part 2
- Helis Academy course FAIR data stewardship 2021, Day 2, Data security and privacy
- Helis Academy course FAIR data stewardship 2021, Day 2, Infrastructure for storing and sharing data
- Helis Academy course FAIR data stewardship 2021, Day 2, Tools for processing and analysing data
- Helis Academy course FAIR data stewardship 2021, Day 3, Introduction Wrap up
- Helis Academy course FAIR data stewardship 2021, Day 3, Electronic Lab Notebooks (ELN)
- Helis Academy course FAIR data stewardship 2021, Day 3, Software carpentry/Versioning
- Helis Academy course FAIR data stewardship 2021, Day 4, Introduction Wrap up
- Helis Academy course FAIR data stewardship 2021, Day 4, Software carpentry
- Helis Academy course FAIR data stewardship 2021, Day 4, Data carpentry
- Helis Academy course FAIR data stewardship 2021, Day 5, Introduction Wrap up
- Helis Academy course FAIR data stewardship 2021, Day 5, Archiving data
- Helis Academy course FAIR data stewardship 2021, Day 5, Data rights
- Helis Academy course FAIR data stewardship 2021, Day 5, Preregistration
- Helis Academy course FAIR data stewardship 2021, Day 6, Introduction Wrap up
- Helis Academy course FAIR data stewardship 2021, Day 6, FAIR Data
- Helis Academy course FAIR data stewardship 2021, Day 6, Metadata
- Helis Academy course FAIR data stewardship 2021, Day 6, Metadata FAIR Data Point



https://fairaware.dans.knaw.nl/
https://doi.org/10.5281/zenodo.5704716

Welcome and introduction

Your needs

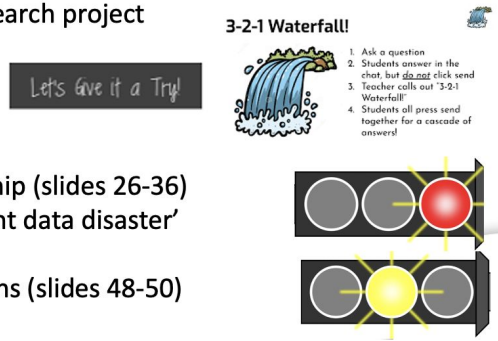
Via the assignments, you have provided the content of today's session. We invite you to participate actively, ask questions and use examples so we can tailor the session to your needs. This session is a success if you feel that it is time well spent!

Objectives

- By the end of this session, you will be able to recognise the basics of FAIR data stewardship in the various stages of the data life cycle
- And you will be able to evaluate what actions need to be taken to solve FAIR data stewardship issues in the proEVlifecycle research project

Content

1. FAIR data stewardship generics (slides 4-9)
2. How to be FAIR aware (slides 10-25)
3. Recap: group efforts on FAIR data stewardship (slides 26-36)
4. Recap: *ten practices*: 'prepare well to prevent data disaster' (slides 37-47)
5. Next steps: FAIR awareness and group actions (slides 48-50)



Findable data (cf. FAIR Aware)

1. Your dataset should be assigned a globally unique persistent and resolvable identifier

   - ✓ So your dataset can be located unambiguously by humans or machines
   - ✓ Identifiers (such as a DOI) are usually assigned by data repositories
   - ✓ Persistent identifiers (PID) remain stable and direct users to the object consistently over time
   - ✓ Not all data will need a PID; in general, those that underpin published findings or have longer term value are worth assigning a PID
   - ✓ PIDs ideally point to an online page that contains metadata for context and the link to access the actual data, or details about how to request access

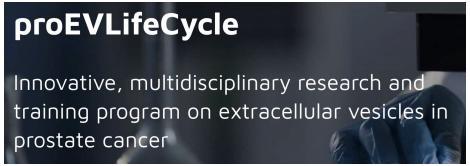
FAIR Aware

2. You will need to provide details (discovery metadata) to make the data findable, understandable and reusable

   - ✓ Metadata: descriptive information about the data object (e.g. creator, title, publisher, creation date, publication date, summary, keywords, PIDs, license)
   - ✓ Data content: an accurate reflection of the data deposited (e.g. resource type, format, size, variables, methods)
   - ✓ Other research outputs: include links to other research output to increase reuse (e.g. prior version, other datasets, publications, data source, data creators/collectors, funders, host institution)
   - ✓ A discipline-specific repository will likely be using common metadata standards (see for instance Re3data or FAIRsharing)
   - ✓ It pays to spend time on providing a good description rather than just the minimum information required, so your data is clearly understood

You have indicated this as a difficult topic

FAIR Aware



https://doi.org/10.5281/zenodo.5704716

4. Recap: ten practices: 'prepare well to prevent data disaster'

Reread last session's article: Briney KA, Coates H, Goben A (2020) Foundational Practices of Research Data Management. Research Ideas and Outcomes 6: e56508

Reread last session's slide deck for example best practices from your own team, and details of the recommended actions

Check the Data Management Plan (D6.3) and EV database report (D4.3)

What steps have you taken in the past year to improve FAIR data stewardship?

Practice 1. Keep sufficient documentation

Easy steps to take

- Create documentation standards/templates to ensure recording of the same information (group effort)
- Take the 'outsider' perspective: can others understand my data activities? (group effort)
- Document so that research is reproducible (document more and more)
- Don't wait: document your data right away

How does this help you?

- Producing documentation in the course of your research ensures that data can be
  - ✓ properly interpreted as relevant context is available
  - ✓ verifiable and reproducible
  - ✓ reusable (by you or by others)
- It helps to explain
  - ✓ the content of the dataset, at the data level (codebook)
  - ✓ the context of the dataset, and how the research was done (methodology section)
  - ✓ the structure of the dataset (readme.txt file with the structure of the dataset)
- Optimising the usability, reusability and reproducibility of the resulting data

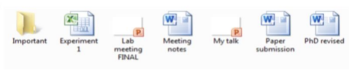
Practice 2. Organise files and name them consistently

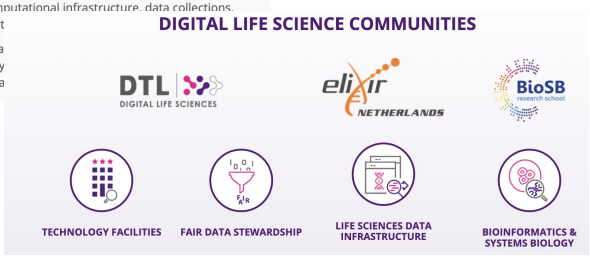
Easy steps to take

- Record the structure you choose, and create standards/templates to ensure organising in the same way (group effort)
- Separate raw, analysed, processed data
- Separate ongoing from closed work (milestone versions)
- If you use abbreviations, make sure to explain them
- Keeping folder and file names as short as possible
- Apply the same to physical data (samples)
- Don't wait: organise your data right away

How does this help you?

- Optimising the usability, reusability and reproducibility of the resulting data
- In 3 years time, would you know what these are?

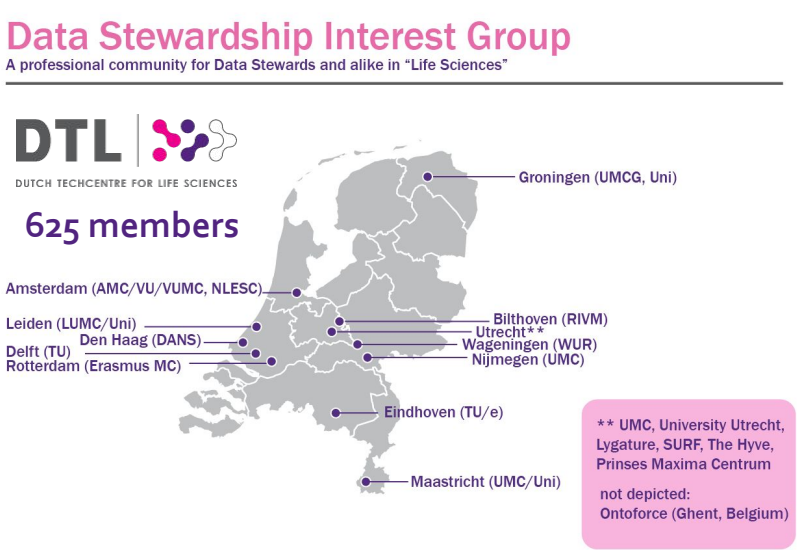




← Communities

Data Stewardship Community

The Health-RI Data Stewardship Community (DSC) will establish a community hub for health data stewards to facilitate collaboration.



- Facilitating communities is an essential element of professionalising data stewardship and capacity building, to exchange experiences and good practices jointly tackle data challenges
- For many years, DTL facilitates the Data Stewards Interest Group (DSIG), with regular meetings and a vibrant (slack channel) community for data stewards and like-minded in the Netherlands **and beyond** to share experiences and foster the (Dutch) national implementation of data stewardship
- In Health-RI, the Data Stewardship Community (DSC) unites healthcare data stewards in national collaborations, so each institute does not have to reinvent the wheel locally

Implementation Plan Investments Digital Research Infrastructure

With the extra structural investments and the kickstarter funding, deployed as an integral coherent plan, NWO wants to stimulate the following aspects:

- Data sharing according to the principles of FAIR (Findable, Accessible, Interoperable, Reusable) and open science.
- Federated digital infrastructure both locally and at the (inter)national level.
- Making research data easier to analyse by using more powerful computer capacity, larger storage capacity and new technological developments.
- Making more use of software already developed because this is better known and more accessible.

Subprogrammes

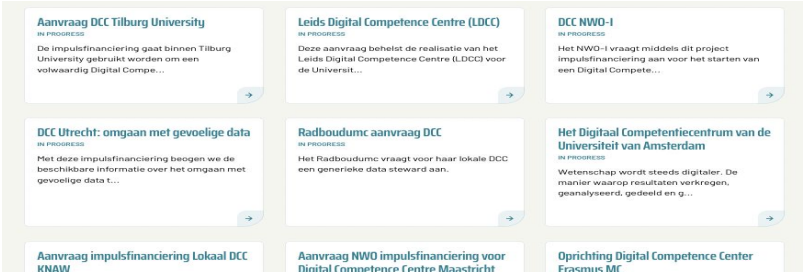
There are four funding lines within this programme: Local DCCs, thematic DCCs, Investments in eScience, Computing facilities (supercomputer and computing time). The PC-LRI subcommittee ICT is in charge of the programme.

- Local DCCs
- Thematic DCCs
- Investments in eScience
- Computer facilities
- ICT subcommittee of the PC-LRI

https://www.nwo.nl/en/researchprogrammes/implementation-plan-investments-digital-rese

Local DCCs

This call was a one-off stimulus for the setting up or further development of local Digital Competence Centers. Research institutions could use this funding to appoint data stewards and data managers for an existing DCC or the central setting up of a new DCC within the institution. With this call, NWO also wanted to ensure that the institution would safeguard the DCC concerned from both a policy and financial perspective.



Thematic DCCs

In the spring of 2022, the thematic DCC (TDCC) network organisations will start. NWO will subsequently publish a call for proposals each year for the realisation of these networks.

The local DCCs provide generic support to researchers within a single institution. However, many researchers collaborate across institutions on specific research themes, and they could benefit from joint support for their research theme. For example, this could concern how you make agreements to better disclose and share data, or about the development of software that makes data analysis within that research theme easier.

To support this, the implementation plan provides for the formation of TDDCs. These network organisations will be established at the start of 2022 with funding from NWO. There will be three TDCC networks that will bring together researchers in three 'domains' (just like the domains in the National Roadmap for Large-scale Research Facilities):

- Life Sciences and Health (LSH)
- Natural and Engineering Sciences (NES)
- Social Sciences and Humanities (SSH)

Dutch roadmap towards professionalising data stewardship

Creating FAIR data implies

Making well informed choices about

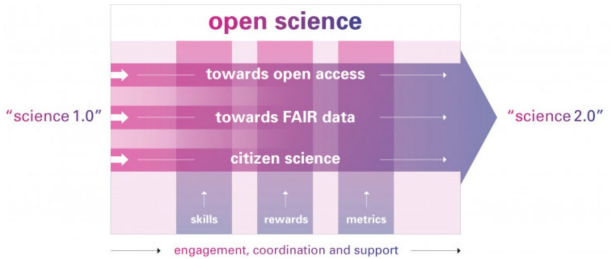
- … the number of data stewards
- … where in the organisation
- … and with what competences
- … including training

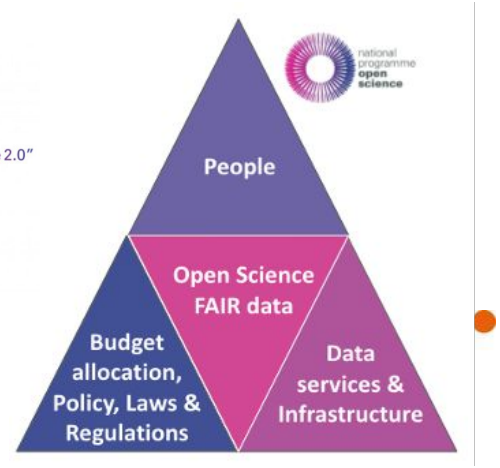
Realising data steward capacity implies

- Changes in research-performing organisations, including HR management
- National funds for science and institutional budgets
- Coordinated action of research institutes, policy makers and research-funding organisations for the required changes

Dutch National Programme Open Science (NPOS)

- Three key areas: 100% Open Access publishing, optimal reuse (FAIR) of research data, and corresponding evaluation and valuation systems
- The data stewardship report (NPOS F) links to the second key area, together with a report on the Dutch data infrastructure and services landscape (NPOS E)
- NPOS F: for a quick overview, we advise to read the preambule, executive summary and Chapter 7



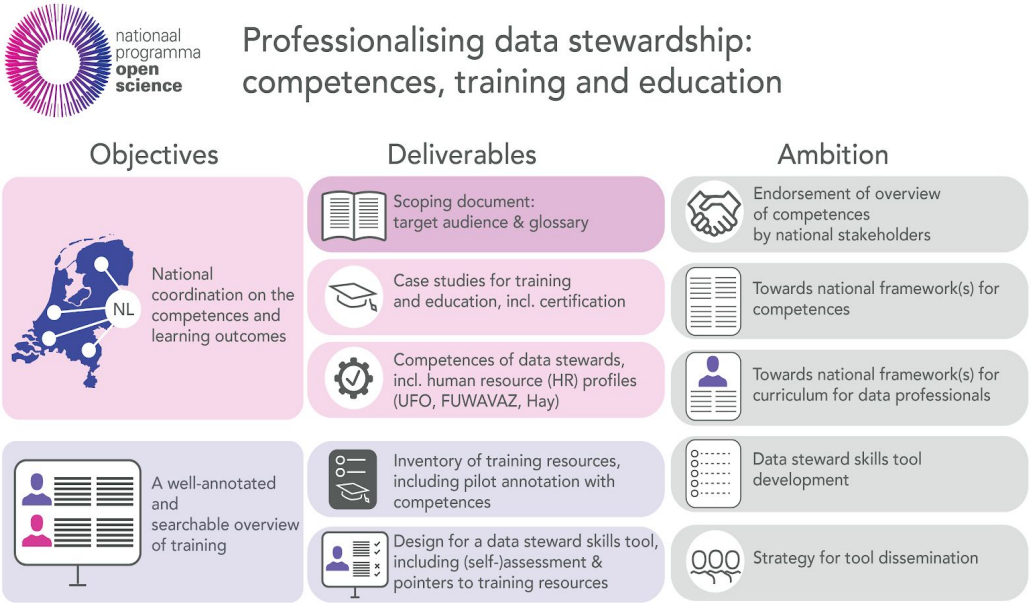


Dutch roadmap towards national implementation of FAIR data stewardship

- A collaborative effort of over 30 representatives of universities, university medical centres (UMCs), universities of applied sciences (UASs), service providers, and representatives of the major Dutch umbrella organisations
- Recommendations will be taken further by the Dutch NPOS 2021-2030 FAIR Data Programme

Professionalising data stewardship in the Netherlands. Competences, training and education. *Dutch roadmap towards national implementation of FAIR data stewardship*

Mijke Jetten, Marjan Grootveld, Annemie Mordant, Mascha Jansen, Margreet Bloemers, Margriet Miedema, & Celia W.G. van Gelder. (2021) https://doi.org/10.5281/zenodo.4320504



Recommendations - Job profiles

- Define data stewardship competences and formalise the job profiles via job classification systems
- Recognise and reward data stewards, secure their position and include development and remuneration

- 2022: The basic job components have been transformed into a **formal UFO (university) data steward profile**, which was formalized in August 2021. In the meantime, the UMC & UAS data steward profiles presented in the report, are already informally adopted, with the expectation to be included in one of the next updates of the system.

Figure 4.2 Basic job profile components of a data steward

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

Recommendations - Training

- Define, develop and deliver tailored training programmes to match the required competences
- Create an continuous culture for data steward education, incl. soft skills, networking and peer exchange
- Formalise existing training efforts incl. certification as a joint effort of local and umbrella organisations

- 2022: The competency work is/will be taken along into **various Dutch training efforts**, e.g., Helis (DTL), DCC Implementation Network (LCRDM), RDNL (E4DS training)

Figure 3.2 Delft University of Technology case study reference card

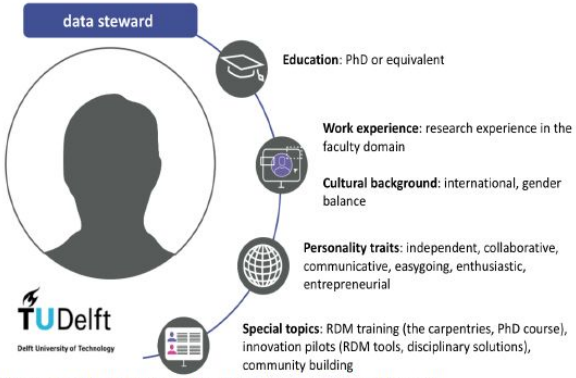


Figure 3.1 Delft University of Technology data steward

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

National Coordination Point Research Data Management

The National Coordination Point Research Data Management (LCRDM) is a national **network** of experts in the field of research data management (RDM).

The LCRDM forms the link between **policy and solution**. Close consultation between educational and research institutions is crucial for this. Within LCRDM, experts work together to put RDM subjects on the agenda that are too big for one institute to tackle and need a national plan of action.

LCRDM brings together research support services, policy makers, ICT specialists, managers of diverse research institutes and research funding organizations. The LCRDM coordinates and facilitates the collaboration between the various RDM stakeholders.

https://www.lcrdm.nl/en

Pool of Experts

The LCDRM also organises meetings on RDM themes, such as November 3, 2020 in Maastricht on adoption and implementation, training and competences of data professionals, and RDM in times of Corona.

RDM maillijst

Are you organizing a symposium, do you have a vacancy, do you want to discuss something with colleagues: use this list. Everyone involved in RDM is welcome. To

Do you have RDM questions that could benefit from a national approach? You will find our working method on the Task groups page.

RDM in the Netherlands

Glossary

TG Software Archiving

Recommendations on encouraging of software archiving.

Pitch and team

TG FAIR enabling

Principles on the basis of which an organisation could be assessed on the degree of 'FAIR' enabling.

Pitch and team

Digital consent

A legal and ethical framework for obtaining digitized informed consent.

Pitch and team

Privacy risks

Implement and maintain a DPIA per research scenario, so that a scenario can be selected for new research proposals.

Data collaboration

Inventory of trans-institutional research projects (use cases), identification of bottlenecks and incompatible policy and description of implications and solutions.

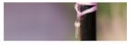
Pitch and team

"23 Things" adoption

A wider adoption of the existing RDA (Research Data Alliance Europe) guide "23 Things - Libraries for Research Data".

Pitch and team

Anonymisation
√2019



Pseudonymisation
√2019

Data Curation
√2019

Task groups

The task groups are composed by the advisory group for each submitted issue. The team members are drawn from the pool of experts by means of a pitch, and possibly supplemented with experts. The task groups are ideally supervised by a process supervisor or consultant. The motivation of the task group members is: 'I also have to do something with my subject in my own working environment'.

Read more about pitching new task groups.

VRE/Workspaces
√2019

Data Stewardship
√2019

"23 Things" Grant
√2019

Working groups 2015 - 2017

ELIXIR-CONVERGE WP2: Training and Capacity Building in Data Management

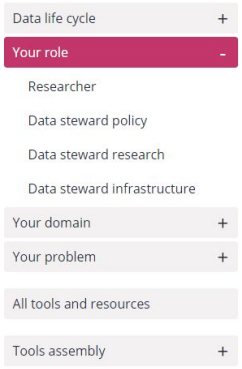
- For data stewards, researchers, and trainers (TtT)
- In alignment with Node priorities and/or national Strategies

71
Training events
from Feb 2020-Dec 2021

16
Countries
organised training in Data Management

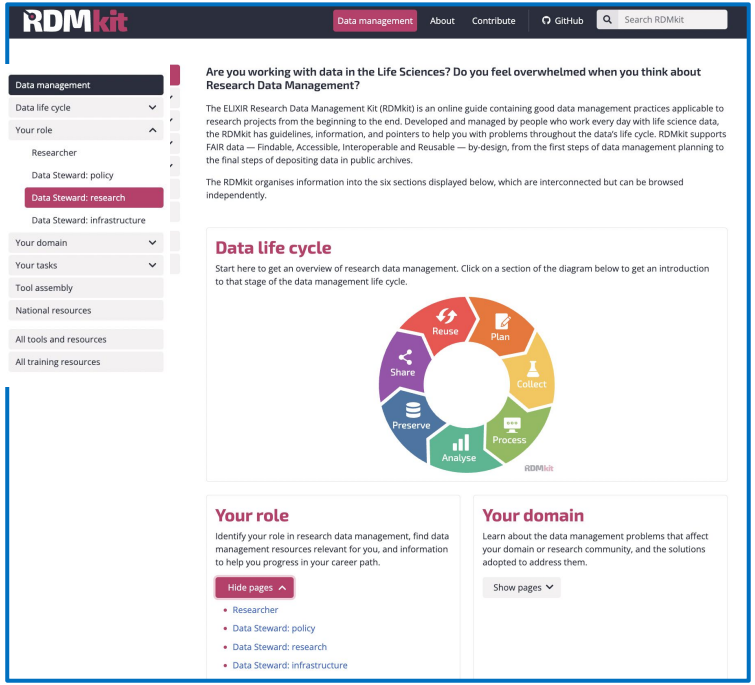
2105
Participants

- ELIXIR-CONVERGE WP3:
  - RDMkit https://rdmkit.elixir-europe.org/
  - Data Stewardship Roles implemented in the Your Role pages in the RDMkit



https://elixir-europe.org/about-us/how-funded/eu-projects/converge

Data steward tool



Your role

Data Steward: research

- Description
- Focus
- Learning path
- Related pages
- More information
- Relevant tools and resources

Description

As a research data steward, I support and work in close collaboration with the main data producers and users in academia: the researchers, ranging from undergraduate students to full professors. I advise researchers, make sure data is handled in a manner compliant with the institute's policy and may also perform hands-on work in a project.

My work focuses on implementing the institute's data guidelines and translating them into domain and project specific procedures, for example by managing a database or reviewing data management plans. My responsibilities and tasks focus on translating the researcher needs on data into infrastructural and service requirements.

Focus

- Develop and implement data management plans for projects and data collections and align Data Managements Plans (DMP) with the FAIR (Findable, Accessible, Interoperable, Reusable) data principles and the principles of Open Science
- Advise projects and data collections on compliance with codes of conduct, regulations and field specific legal and ethical standards
- Provide adequate research data management (RDM) support to researchers. This involves, for example, supporting researchers in improving the reproducibility of their computational analyses or directing researchers to appropriate data management and archival solutions
- Monitor a project's needs regarding data-infrastructure and tools for RDM
- Determine the adequate level of knowledge and skills of researchers on RDM
- Identify the requirements of adequate support and data infrastructure for FAIR and long-term archiving of data of a project

Learning path

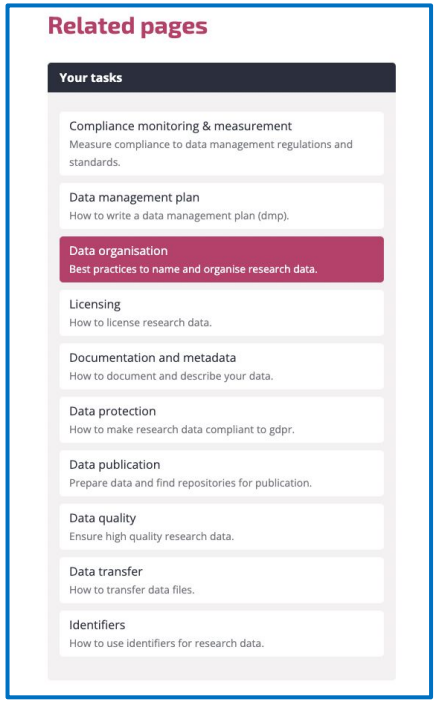
Institutes across Europe have started hiring professional data stewards. A research oriented data steward is expected to be competent in the following areas:

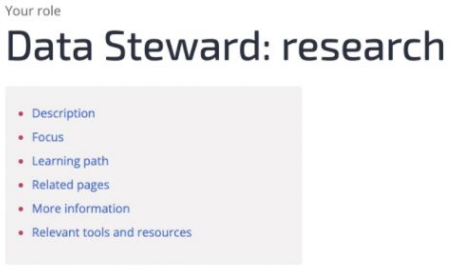
- Create awareness and communicate about RDM and the FAIR data principles and translate RDM policies into guidelines for researchers
- Transform discipline specific research data into FAIR data with help of available services and tools
- Advise and assist researchers on short and long term actions for RDM
- Assess RDM knowledge and skills, identify gaps among researchers and take action when needed
- Understand the purpose and use of a DMP in a project and have the skills to utilise the available tools and templates to produce a DMP
- Assist researchers in developing a DMP, review DMPs, and support researchers in putting DMPs into action
- Liaise with the surrounding environment (department, project, national stakeholders and international network) and continuously follow the field to gain knowledge of relevant facilities, tools and emerging standards available for RDM

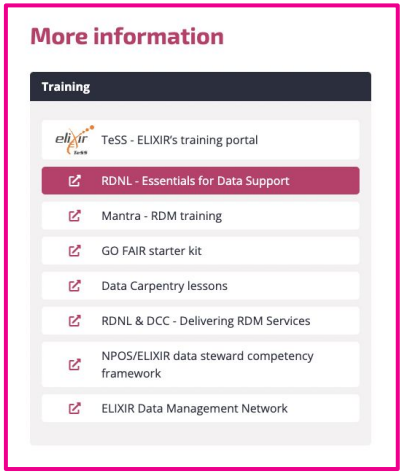
https://rdmkit.elixir-europe.org

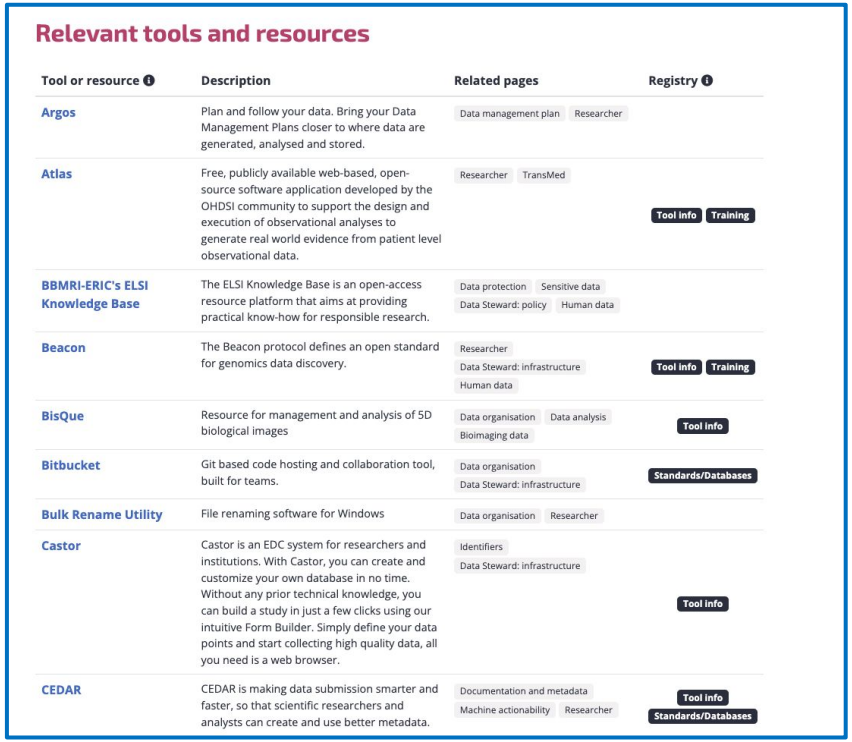
RDMkit

Data steward tool









https://rdmkit.elixir-europe.org

RDMkit

European Open Science Cloud - EOSC

- The EOSC (https://eosc.eu/) will offer researchers a virtual environment with **open and seamless services for storage, management, analysis and re-use of research data**, across borders and scientific disciplines by federating existing data infrastructure

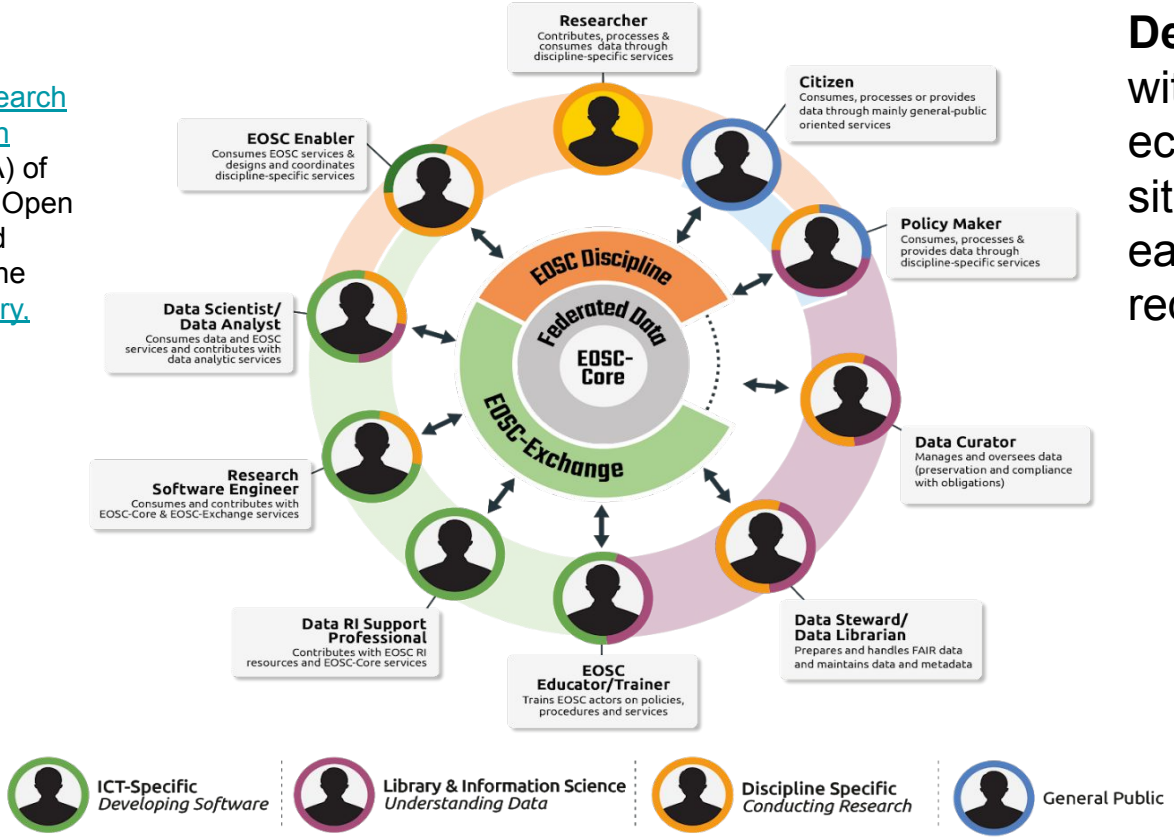
- An EOSC **network of skilled professionals** is essential to bring a culture change for sharing research outcomes, and to empower individuals and institutions to develop and maintain EOSC competences, capabilities and skills.

EOSCSkills&Training
Executive Board Working Group

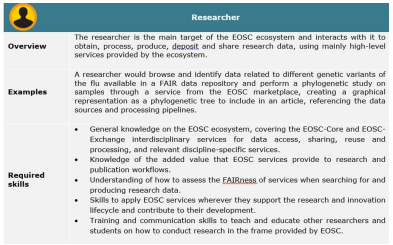
Report "Digital skills for FAIR and open science" from the Working Group Skills & Training (2021)

Framework of Actors in the EOSC Ecosystem

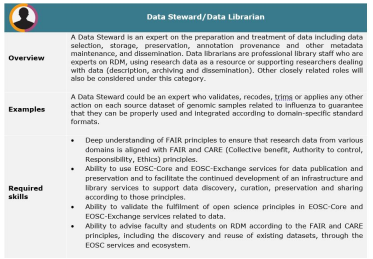
Also in the Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC) and the EOSC Glossary, Dec 2020



Description of 10 roles within the EOSC ecosystem, one situational example for each role and a list of required skills



Researcher	
Overview	The researcher is the main target of the EOSC ecosystem and interacts with it to obtain, process, produce, deposit and share research data, using mainly high-level services provided by the ecosystem.
Examples	A researcher would browse and identify data related to different genetic variants of the flu available in a FAIR data repository and perform a phylogenetic study on samples through a service from the EOSC marketplace, creating a graphical representation as a phylogenetic tree to include in an article, referencing the data sources and processing pipelines.
Required skills	• General knowledge on the EOSC ecosystem, covering the EOSC-Core and EOSC-Exchange interdisciplinary services for data access, sharing, reuse and processing, and relevant discipline-specific services. • Knowledge of the added value that EOSC services provide to research and publication workflows. • Understanding of how to assess the FAIRness of services when searching for and producing research data. • Skills to apply EOSC services wherever they support the research and innovation lifecycle and contribute to their development. • Training and communication skills to teach and educate other researchers and students on how to conduct research in the frame provided by EOSC.



Data Steward/Data Librarian	
Overview	A Data Steward is an expert on the preparation and treatment of data including data selection, storage, preservation, annotation provenance and other metadata maintenance, and dissemination. Data librarians are professional library staff who are experts on RDM, using research data as a resource or supporting researchers dealing with data (description, archiving and dissemination). Other closely related roles will also be considered under this category.
Examples	A Data Steward could be an expert who validates, recodes, trims or applies any other action on each source dataset of genomic samples related to influenza to guarantee that they can be properly used and integrated according to domain-specific standard formats.
Required skills	• Deep understanding of FAIR principles to ensure that research data from various domains is aligned with FAIR and CARE (Collective benefit, Authority to control, Responsibility, Ethics) principles. • Ability to use EOSC-Core and EOSC-Exchange services for data publication and preservation and to facilitate the continued development of an infrastructure and library services to support data discovery, curation, preservation and sharing according to those principles. • Ability to validate the fulfilment of open science principles in EOSC-Core and EOSC-Exchange services related to data. • Ability to advise faculty and students on RDM according to the FAIR and CARE principles, including the discovery and reuse of existing datasets, through the EOSC services and ecosystem.

Main recommendations of the EOSC Training & Skills Working Group

1. Utilise the **Framework of Actors in the EOSC Ecosystem** in the development and mainstreaming of FAIR and open science skills and training
2. Coordinate and align **relevant skills curricula and training frameworks**
3. Encourage and support the **competence centres approach** for FAIR and open science training
4. Facilitate increased **integration of FAIR and open science courses** with university qualifications.
5. Build a **learning and training catalogue** to maximise interoperability.
6. Include learning and training resources in the EOSC Interoperability Framework (EIF)
7. Develop an **EOSC Skills and Training Leadership Programme**

-

Recommendations of the EOSC Training & Skills WG

The report gives specific recommendations for the different stakeholders:

- Policy makers & funders
- Universities & research organisations
- Competence centres
- EOSC Association
- EOSC projects

Policymakers & funders	Universities & research organisations	Competence centres	EOSC Association	EOSC projects
Understand the importance of addressing gaps in provision of digital skills for FAIR and open science				
Identify actors/roles that need to be supported in the EOSC ecosystem	Consider diversity of employment and training needs of staff	Identify knowledge to be taught and skills to be learned	Understand range of roles that initiatives must address	Ensure projects consider appropriate roles for both their staff and users
Consider how to support competence centres and encourage collaboration	Consider how to establish competence centres	Learn about governance, business models and alignment approaches of other competence centres	Consider how to encourage and support competence centres approach	Ensure projects consider FAIR and open science training
Recognise the need for investment in hubs for training resources	Know how to provide and access learning and training resources	Learn how to maximise interoperability, FAIRness and usage of learning and training resources	Facilitate interoperability and discovery of learning and training resources	Identify and engage key stakeholders, and build learning and training catalogues according to their goals and actions
Identify actions to advance national policies on FAIR and open science	Consider how to encourage and/or implement national approaches	Understand the importance of aligning with national programmes and policies	Identify ways to increase international alignment	Not relevant

EOSC Association – 2022 Advisory Groups & Task Forces

- Implementation of EOSC
  - Rules of Participation compliance monitoring
  - PID policy and implementation
  - Researcher engagement and adoption
- Technical challenges on EOSC
  - Technical interoperability of data and services
  - Infrastructure for quality research software
  - AAI Architecture
- Metadata and data quality
  - Semantic interoperability
  - FAIR metrics and data quality
- **Research careers and curricula**
  - **Data stewardship curricula and career paths**
  - **Research careers, recognition and credit**
  - **Upskilling countries to engage in EOSC**
- Sustaining EOSC
  - Defining funding models for EOSC
  - Long-term data preservation

https://www.eosc.eu/advisory-groups

Building the social and technical bridges to enable open sharing and re-use of data

RDA EU RDA US CONTACT US LOGIN REGIST

ABOUT RDA GET INVOLVED GROUPS RECOMMENDATIONS & OUTPUTS RDA FOR DISCIPLINES PLENARIES & EVI

Home » Working and Interest Groups » Interest Group » Education and Training on handling of research data IG

IG

Education and Training on handling of research data IG

Taxonomy:

Posts | Wiki | Events | Repository | Outputs | Charter | Plenaries | Members

Group Status: IG Established

IG

ELIXIR Bridging Force IG

Taxonomy: Natural Sciences, Natural Sciences/Biological sciences, Medical and Health Sciences

Posts | Create Wiki index | Events | Repository | Outputs | Charter | Plenaries | Members

Group Status: IG Established

uirements for rese
no create and use

cation plans. Also,

IG

Status:
Chair (
Case St
IG

Life scien
"omics" te
Simultane
data of di
marine lif
European
society.
The ELIXI
data anal
relevant i
The ELIXI
organisat
Int

WG

RDA-COVID19-Omics

Taxonomy: Medical and Health Sciences

Posts | Wiki | Events | Repository | Outputs | Case Statements | Plenaries | Members

Group Status: WGs Getting started (~0-6 months after RDA endorsement)

Status: Recognised & Endorsed
Chair (s): Natalie Meyers, Rob Hooft, Juan Bicarregui

This is the subgroup of the RDA-COVID-19 working group focusing on **Omics**.

- **Useful Information for RDA COVID-19 Omics Subgroup**

Omics subgroup priorities:

1. A set of guideline documents, highlighting the primary data [and software/c
   resources in Omics research, addressing different data types and cross-cutti
2. Resource [data and software/code] List(s) in Omics.
3. A Decision Tree tool to facilitate navigation to specific Omics Resources.

RESEARCH DATA ALLIANCE

O&A Members 63

Active Organisational & Affiliate members

MEMBERSHIP Members: 10956

Becoming a member of RDA is simple and open to both individuals and organizations

Register now

R
Di
Gr
ou

ABOUT RDA GET INVOLVED GROUPS RECOMMENDATIONS & OUTPUTS RDA FOR DISCIPLINES PLENARIES & EVI

Home » Working and Interest Groups » Interest Group » Professionalising Data Stewardship IG

IG

Professionalising Data Stewardship IG

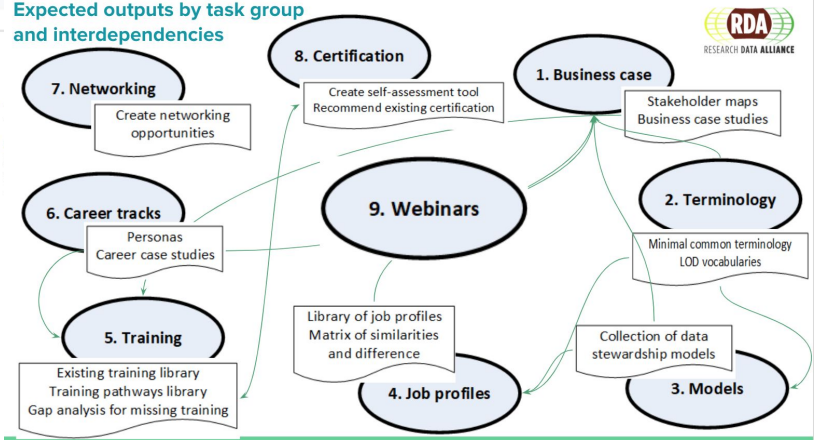
Taxonomy:

Posts | Create Wiki index | Events | Repository | Outputs | Charter | Plenaries | Members

Group Status: Not yet endorsed

Expected outputs by task group and interdependencies

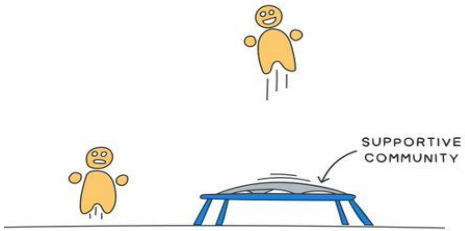
RDA RESEARCH DATA ALLIANCE



Takeaways

- Creating FAIR data implies making well informed choices about
  - … the number of data stewards
  - ... where in the organisation
  - … and with what competences
  - … including training

- Data stewardship is a new profession and things are not set in stone yet
  - There is a lot going on in Europe (and beyond) where you can build on, both for yourself as well as for your institute
  - Discuss your current/future roles, responsibilities and tasks in your local teams and organisations
  - Become part of the data stewards community!



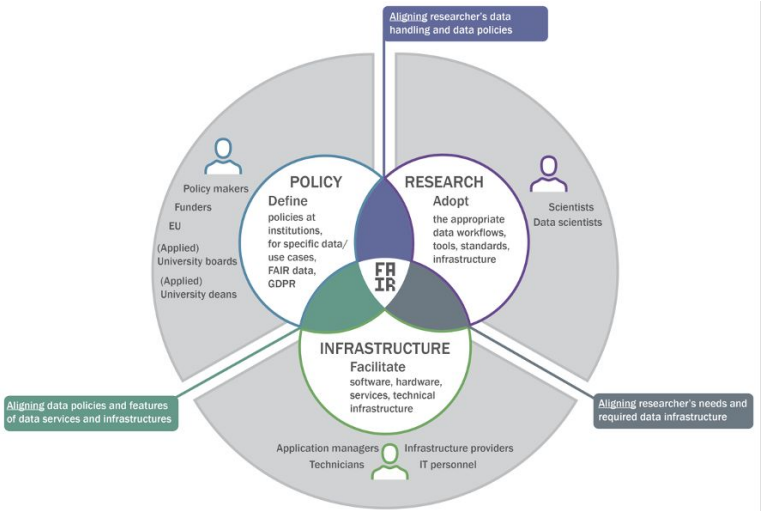
@OzolinsJanis

- Recommended next steps
  - Competencies and skills: what roles do you have in the team, what additional roles do you need?
  - Capacity: Identify capacity (fte) needed and work towards installing that capacity
  - Build a data stewardship community across the institutes
  - Look at the reports, tools and training resources in this presentation

health RI enabling data driven health
elixir CONVERGE
elixir NETHERLANDS
DTL DUTCH TECHCENTRE FOR LIFE SCIENCES
national programme open science

Discussion with the audience

- Do you recognise the sketched developments in professionalising data stewardship (challenges, needs, solutions) from your (local) organisation's perspective?
- In what way do you contribute yourself to professionalising data stewardship?
- Where are you and your colleagues in the diagram?



health RI enabling data driven health
elixir CONVERGE
elixir NETHERLANDS
DTL DUTCH TECHCENTRE FOR LIFE SCIENCES
national programme open science

Acknowledgements

- The ELIXIR-NL, DTL and Health-RI team
- ELIXIR-CONVERGE members
- NPOS-F team
- DTL Data Stewards Interest Group (DSIG) and the Health-RI Data Stewardship Community (DSC)
- RDA Professionalising Data Stewardship Interest Group

Thank you for listening!

Interested to learn more about DTL, Health-RI, ELIXIR-NL and NPOS activities? Contact us via mijke.jetten@dtls.nl and celia.van.gelder@dtls.nl