DTL
DUTCH TECHCENTRE FOR LIFE SCIENCES

Welcome to ask questions during the presentation
Link to slides: https://doi.org/10.5281/zenodo.4773340

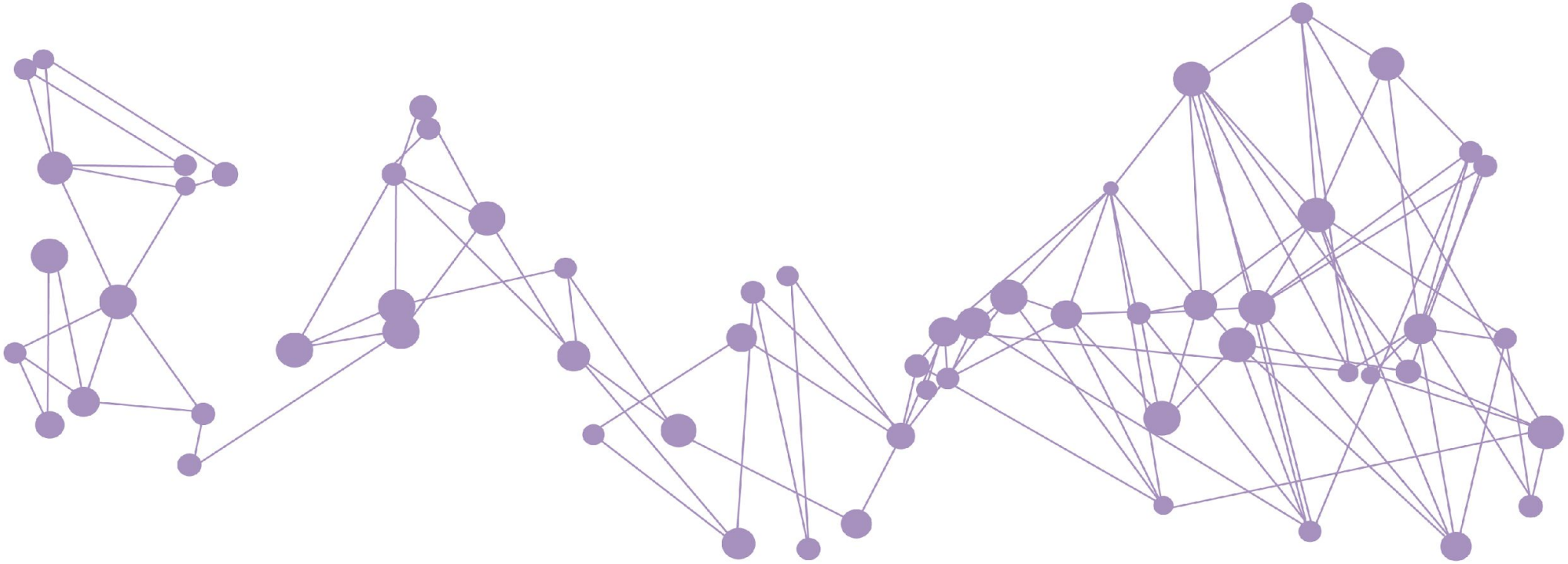
nationaal programma open science

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

Dutch roadmap towards national implementation of FAIR data stewardship

Mijke Jetten, DTL/Health-RI, mijke.jetten@dtls.nl

FAIR data Austria webinar, 20 May 2021



Need for skills and capacity building in FAIR data stewardship

- Data stewardship and data management skills are essential in research.
  Challenges:
  - There is no consensus on the responsibilities and tasks of data stewards, and formal profiles (including knowledge, skills and abilities (KSAs)) are lacking
  - No/too little and/or not findable tailored education & training
  - This hampers the adequate capacity building and complicates efficient data management which are necessary for open science implementation
  - (Inter)national alignment and coordination are needed to achieve coherent training and education, accompanied with a consistent human resource (HR) policy

- By **bundling expertise and manpower**, with support of the **research funders**, in the Netherlands, we are making steps to tackle FAIR data handling challenges
  - These joint capacity building efforts will help build a skilled workforce that is ready for professional data handling
  - Our approach: by the community, for the community

CC BY

nationaal programma open science

How much capacity is needed? Still a very long way to go!

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

That is 3 FTE per 100 researchers!

Number adapted from OECD (2020), "Building digital workforce capacity and skills for data-intensive science"

The EC High Level Expert Group EOSC estimated 5 FTE per 100 researchers

Landelijk Coördinatiepunt Research Data Management

Results of LCRDM "Do I PASS for FAIR" 2020 survey

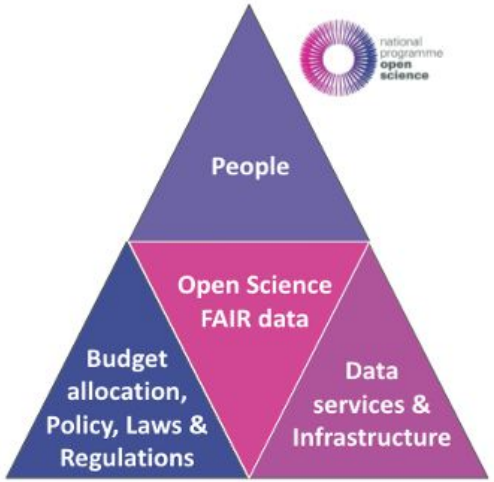
CC BY

Dutch roadmap towards professionalising data stewardship

"Invest 5% of research funds in ensuring data are reusable. It is irresponsible to support research but not data stewardship. (…) Many top universities are starting to see that the costs of not sharing data are significant and greater than the associated risks. Data stewardship offers excellent returns on investment." (Barend Mons, GO FAIR/CODATA)

Dutch National Programme Open Science (NPOS)

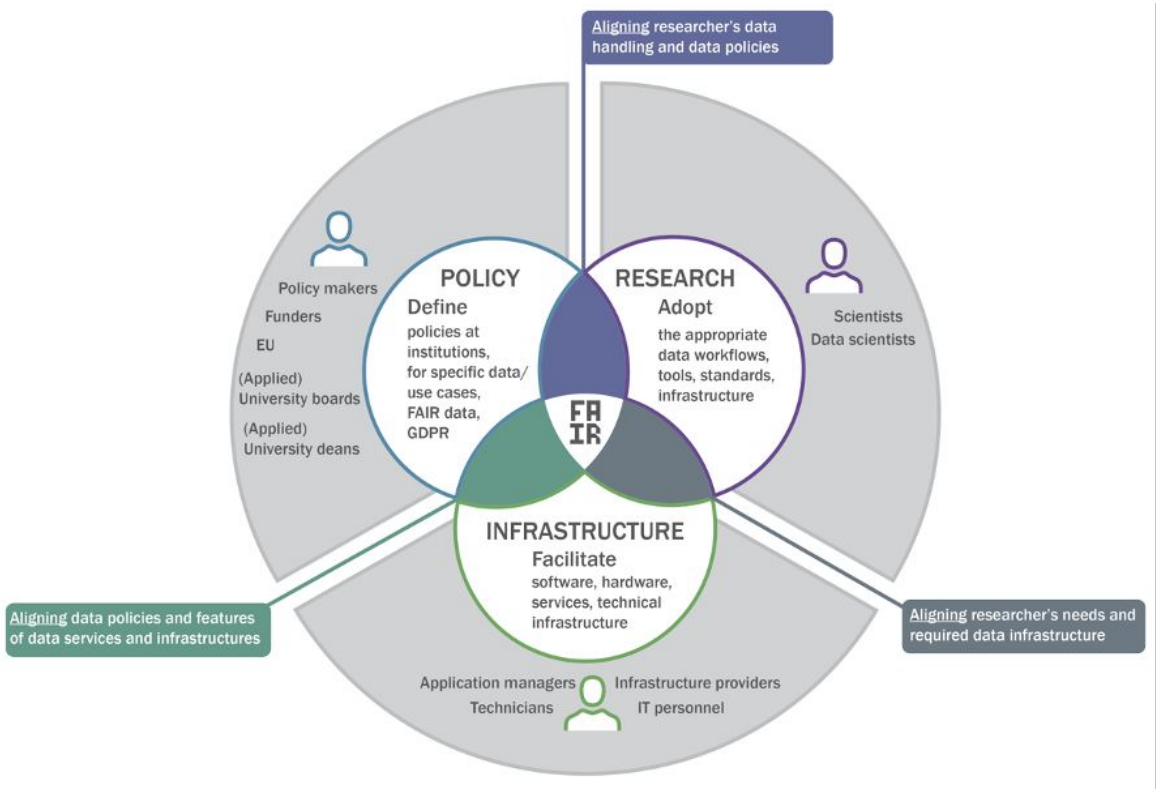
- Three key areas: 100% Open Access publishing, optimal reuse (FAIR) of research data, and corresponding evaluation and valuation systems
- The data stewardship report (NPOS F) links to the second key area, together with a report on the Dutch data infrastructure and services landscape (NPOS E)
- Three dimensions of the route to Open Science: 1. policy, regulations, and allocation of budgets, 2. research infrastructures and services, 3. cultural change in the research and the research support community
- The data stewardship report (NPOS F) consists of over 30 representatives of universities, university medical centres (UMCs), universities of applied sciences (UASs), service providers & representatives of Dutch umbrella organisations



BY
DTL
DUTCH TECHCENTRE FOR LIFE SCIENCES
nationaal programma open science

Previous reports

The NPOS report builds upon the widely supported outcomes and recommendations of two previous projects (2019), i.e. the ZonMw data stewardship project and the LCRDM data stewardship project



LCRDM data stewardship task areas
https://doi.org/10.5281/zenodo.2669150
https://doi.org/10.5281/zenodo.3066366

NPOS/ELIXIR data stewardship roles in the data stewardship landscape
http://doi.org/10.5281/zenodo.3474789
https://doi.org/10.5281/zenodo.3243909

nationaal programma open science

NPOS F report content

- **Chapter 1**: variety of international developments and introduces NPOS and the NPOS-F project

- **Chapter 2**: definition of data steward(ship), training & education, results of 2 earlier projects

- The next four chapters address the four angles from which we approached the project:
  - **Chapter 3**: analysis of the data stewardship training and education landscape, based on case studies
  - **Chapter 4**: components of a job profile for data stewards (and research software engineers)
  - **Chapter 5**: existing data-related trainings, pilot annotation of courses based on competences and at certification of education and training
  - **Chapter 6**: design of a data steward skills tool with navigation pathways, as a single point of reference for up-to-date information

- Whereas **Chapters 3 through 6** contain recommendations, the **final Chapter 7** addresses higher-level recommendations to specific stakeholders

For a quick overview, we advise to read the preambule, executive summary and Chapter 7

CC BY

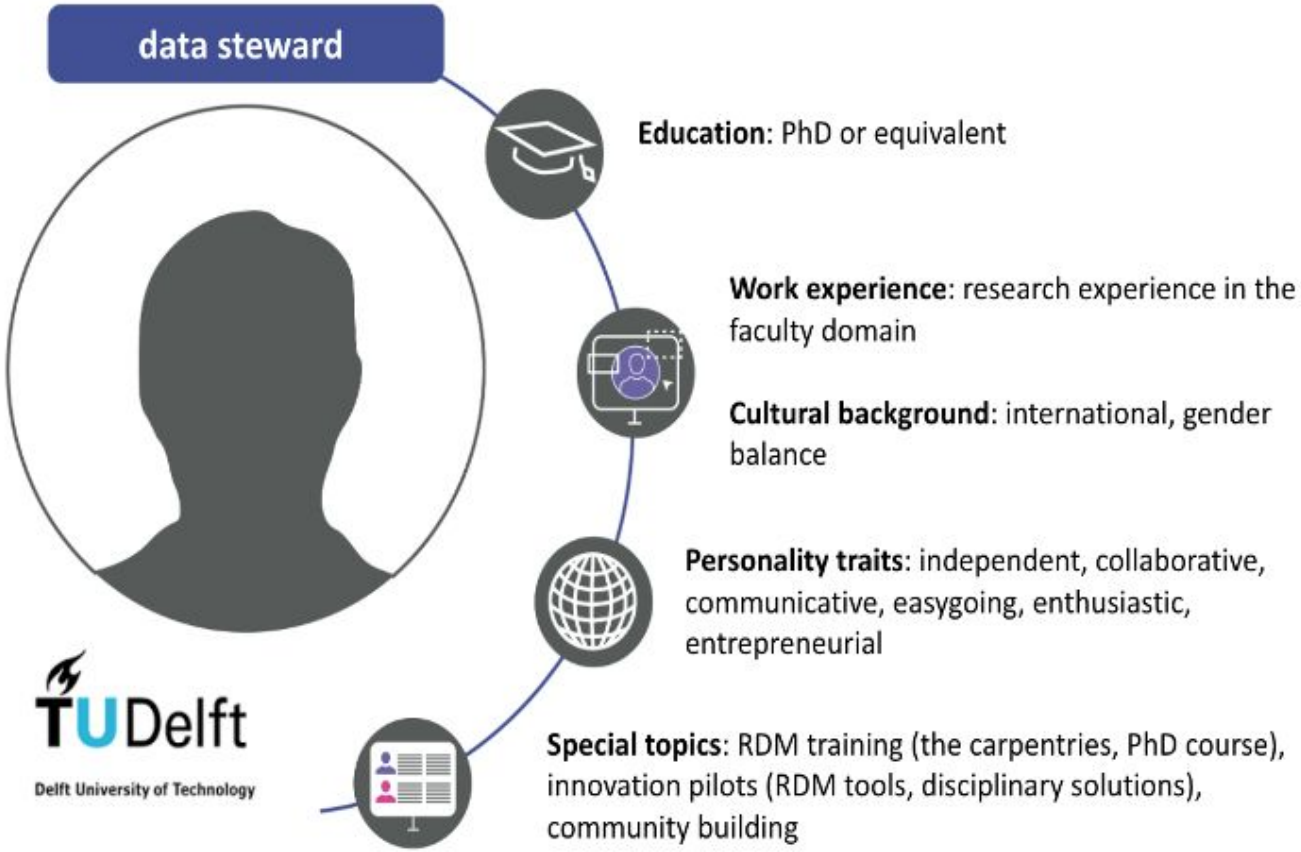
DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

nationaal programma open science

Landscape analysis on training and education

8 case studies (universities, UASs & UMCs) as a first fact check on data stewardship training and education in the Netherlands

Example Delft University of Technology data steward & case study reference card



Data steward in the organisation

1. Appointment explicit part of RDM policy
2. Positioned at the research institute
3. Additional RDM support for centralised tasks at library
4. No formal central coordination, library is informal linking pin

Training for data stewards

1. Training organised by the coordinator
2. Formal RDM training
3. Training on TUD research support

Learning on the job

1. Structured mentoring
2. Internal peer support (data stewards team)
3. External peer networks
4. Soft skills development: training and on the job
5. Gaining relevant research expertise

Strengths and challenges

1. Strength: well-organised (coordination)
2. Strength: institutional buy-in
3. Strength: sustainable
4. Strength: team-feeling among the data stewards
5. Challenge: expensive (9 FTE)

Landscape analysis on training and education

Challenges, based on the landscape analysis

[See Section 3.2 for details]

1. Formalisation: lack of formal education and training for data stewards, and unclear which skills and personal development are required for data stewardship
2. Defining needs: difficulty of defining training needs of data stewards - which will help trainers and educational experts to develop training programs
3. Existing training: lack of insight into existing training for data stewards: what is available, with which content, what knowledge gaps does it address, and how do learners value it?
4. Lifelong learning: lack of support for life-long learning for data stewards
5. Strategic vision: a bottom-up implementation of data stewardship - without a strategic vision and corresponding capacity, governance and funding in an organisation - is vulnerable and susceptible to frequently changing priorities
6. Awareness: the difficulty of raising awareness about data stewardship among researchers

CC BY

Landscape analysis on training and education

Recommendations, based on the landscape analysis

[See Section 3.3 for details, including match with challenges and addressed stakeholders]

1. Use case studies to plan training
2. Care for your data steward
3. Collaborate in training
4. Importance of community and networking
5. Coordinated approach to data stewardship
6. Flexibility in the job
7. Proximity of data stewards to peers

CC BY

Basic data steward job profile components

Challenges

[See Section 4.2 for details]

1. Profiles: lack of profiles that match responsibilities and tasks, which complicates recruitment
2. Career tracks: difficulty of defining clear career tracks for data stewards, including recognition and remuneration
3. Position: in the organisation, lack of attention for the distinctive position of a locally embedded data steward at the faculty or department level, and a centrally positioned generic data steward
4. Good practices: even if good practices and example job profiles exist, most organisations are not aware of these examples, as they are not shared
5. Capacity: lack of capacity in data stewardship, due to the demands for FAIR data. Formal job profiles will contribute to professionalising data stewardship and thus facilitate the process of recruitment

CC BY

Basic data steward job profile components

Various annexes with detailed information

- Domain areas, responsibilities, tasks, and competences of a data steward
- Basic components of the data steward job profile - for the UFO (universities) job classification system, as a first step
- Three local university data steward job profiles - for the FUWAVAZ (UMCs) job classification system
- Proposal for a job profile for data stewards in UASs
- Context of the Digital Competence Centers: additionally, the basic components of the research software engineer job profile

cc BY

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

nationaal programma open science

Basic data steward job profile components

Recommendations

[See Section 4.4 for details, including match with challenges and addressed stakeholders]

1. Formalise profiles
2. Adopt profiles
3. Create career perspectives
4. Allow diversity of roles and types
5. Adopt good practices
6. Secure positions

CC BY

Training, education and certification

Challenges

[See Section 5.2 for details]

1. Findability: lack of findable, adequate education and training for data stewards, or where they exist, missing clarity on how they relate to the competence development of data stewards
2. Competences: lack of agreement on responsibilities, tasks and competences of data stewards, which complicates developing education and training
3. Coordination: lack of alignment and coordination among and between local, national and international stakeholders on developing education and training related to data stewardship
4. Certification: lack of certification for data stewardship related education and training, training providers and trainers

CC BY

Training, education and certification

- Inventory of training resources and pilot annotation of courses based on competences
- Inventory of existing certification mechanism, for different categories: courses, trainees, trainers, and organisations
- Certification for data stewardship is still in its early days and needs to be done in alignment with similar activities in Europe and beyond

Recommendations

[See Section 5.4 for details, including match with challenges and addressed stakeholders]

1. Standardise metadata for training
2. Develop a training annotation process
3. Create curated resources
4. Align with international certification initiatives
5. Identify a certification provider

CC BY

Data stewards skills tool

Challenge

[See Section 6.2 for details]

Single point of reference (tool): up-to-date information on competences, profiles, training, and allowing for (self)assessment and identification of career development, for:

- Data stewards
- Employers looking at recruitment
- Employers and employees wanting to make an assessment of skills and development needs
- Training providers wanting to develop new training opportunities

CC BY

Data stewards skills tool

This single challenge summarises several challenges mentioned before

- Compare data steward roles to clarify responsibilities and ensure that teams contain complementary skills
- Individual data stewards have no benchmarks against which they can assess their skills and knowledge
- Data stewards have varied backgrounds and therefore different development needs
- Employers have to rely on previous examples of job vacancies to describe the competences for new job profiles and assess potential candidates
- Data stewards have no clarity in potential career progression
- Data stewards and employers often rely on serendipity to discover training opportunities
- Data stewards and their employing institutions frequently rely on national networks to share ideas and information, but with the growing population of data stewards it is becoming increasingly difficult to keep track of all opportunities and examples of best practice

CC BY

data steward

Education and qualification: Ma or PhD in a relevant scientific domain

Work experience and background: research experience in university or business; speaks the language of the researcher; aware of researchers' needs; embedded in specific projects

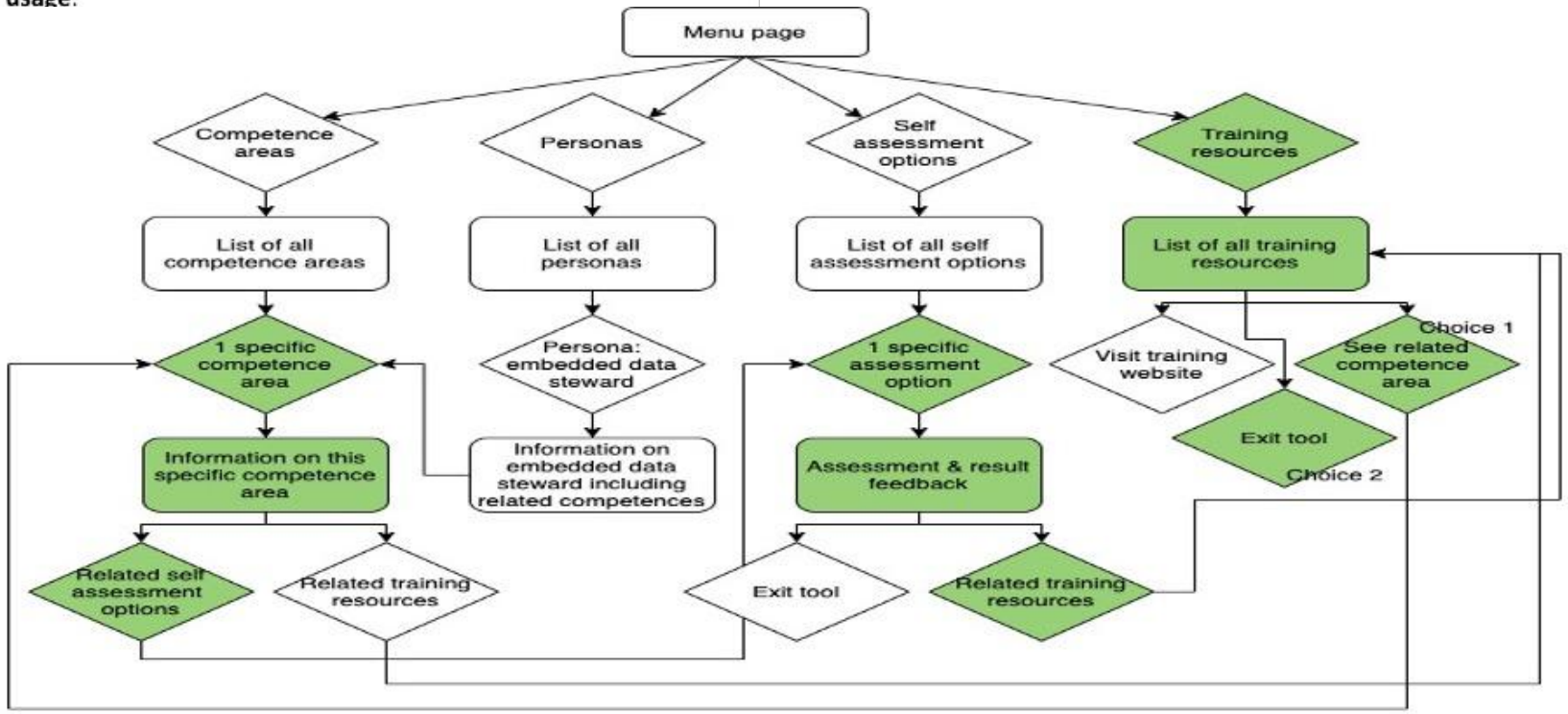
Responsibilities and tasks: efficient and quick in handling data; analytical focus; responsible for data processing and analysing; manages big data; takes a consultancy role

RESEARCH

User story (example) for tool usage:
- As a research oriented data management is an intrinsic p
- As a research oriented data help my researchers with ma
- As a research oriented data data management, so I know

Research data steward persona (individual perspective)

Training resources, example learner journey of a research data steward



CC BY

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

Data stewards skills tool

- Sketch for a data steward skills tool as a single point of reference for data steward competences, training and education
- 5 data steward personas, to show how stewards could benefit from the tool, and possible pathways for these personas in such a tool
- In collaboration with ELIXIR, as a pilot, content was added to the EBI Competency Hub tool: expertise areas, responsibilities, tasks, KSAs, and learning objectives. The Competency Hub is interested to expand their tool for our use case and discuss adaptations needed

Recommendations

[See Section 6.4 for details, including match with challenges and addressed stakeholders]

1. Competency Hub integration
2. Committee of stakeholders for development process
3. Working group for content
4. Potential owner inventory

CC BY

Implementing data stewardship

Summary of recommendations

- Analyse local data stewardship **case studies** to clarify the needs of your organisation
- Formalise the **data steward profile** and stimulate local organisations to adopt these
- These profiles should reflect the required competences and be sensitive to diversity in background, expertise, roles, types and organisations. This should result in **secured positions and career tracks**
- Collaborate locally and nationally in **organising training and develop a certified curriculum** to meet the required expertise level and capacity
- Standardise **metadata for training and education** and create curated training resources
- **Create a data steward skills tool** as a single point of reference for data steward competences, training and education

First steps towards implementation in the Netherlands

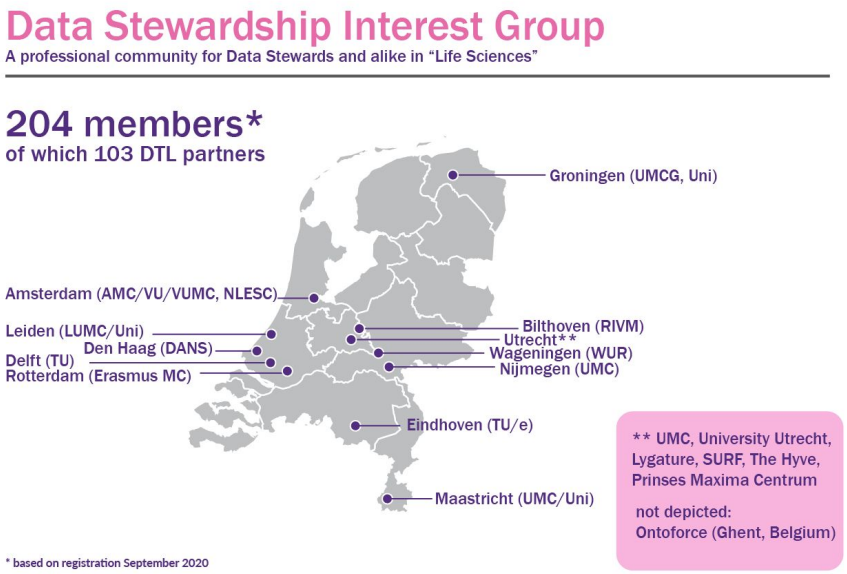
- NPOS-F recommendations will be taken further by the **Dutch NPOS 2021-2030 FAIR Data Programme**
- The data stewardship competency framework has been published in the **EBI Competency Hub**
- The competency work is/will be taken along into **various Dutch training efforts**, e.g., Helis (DTL), DCC Implementation Network (LCRDM), RDNL (E4DS training)
- The basic job components have been transformed into a **formal UFO (university) data steward profile**, which is expected to be formalised in August 2021. Similar discussions are ongoing with the FUWAVAZ (UMCs) working group

CC BY

Community building

- Facilitating communities is an essential element of professionalising data stewardship and FAIR data stewardship capacity building. These communities exchange experiences and good practices, and jointly tackle various data challenges

- For many years, DTL facilitates the Data Stewards Interest Group (DSIG), with regular meetings and a vibrant (slack channel- 372 members) community for data stewards ar like-minded in the Netherlands **and beyond (welcome to join!)** to share experiences and foster the (Dutch) national implementation of data stewardship



CC BY

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

nationaal programma open science

Acknowledgements

Reference

Mijke Jetten, Marjan Grootveld, Annemie Mordant, Mascha Jansen, Margreet Bloemers, Margriet Miedema, & Celia W.G. van Gelder. (2021). Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship. Zenodo. https://doi.org/10.5281/zenodo.4320504

Contact

For more information about the project, you may contact Mijke Jetten (mijke.jetten@dtls.nl) or Celia van Gelder (celia.van.gelder@dtls.nl)

Authors

Mijke Jetten (Dutch Techcentre for Lifesciences, DTL), Marjan Grootveld (Data Archiving and Networked Services, DANS), Annemie Mordant (Maastricht University), Mascha Jansen (GO FAIR Foundation, GFF), Margreet Bloemers (ZonMw), Margriet Miedema (National Coordination Point Research Data Management, LCRDM), Celia van Gelder (Dutch Techcentre for Lifesciences, DTL)

CC BY

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

nationaal programma open science

Acknowledgements

Contributors

Daily project lead: Mijke Jetten (DTL) and Celia van Gelder (DTL)

Participating umbrella organisations: VSNU (Darco Jansen), VH (Jort Diekerhof), NFU (Petra Drankier), PNN (Tess van Doorn), SURF (Annette Langedijk, Claudia Behnke), ZonMw (Margreet Bloemers), LCRDM (Margriet Bloemers) and Elevate Health (Nienke Verdonk)

Participating organisations: Erasmus MC (Sabrina Gunput), Radboudumc (Mirjam Brullemans-Spansier, Peter-Bram 't Hoen), UMCG (Salome Scholtens), UMC Utrecht (Jasmin Böhmer), Eindhoven University of Technology (Iza Witkowska), Leiden University (Kristina Hettne, Joanne Yeomans), Maastricht University (Erik Jansen, Annemie Mordant), Radboud University (Inge Slouwerhof), TU Delft (Marta Teperek, Yan Wang), VU (Brett Olivier, Lena Karvovskaya), Avans University of Applied Sciences (Cora Bijlsma, Miaomiao Zhou), Fontys University of Applied Sciences (Masha Boosten-Ovtchinnikova), HAN University of Applied Sciences (Ingrid Busser), Saxion University of Applied Sciences (Renate Mattiszik), University of Applied Sciences Leiden (Mischa Barthel), DANS (Marjan Grootveld), GO FAIR Foundation (Mascha Jansen), Netherlands eScience Center (Mateusz Kuzak, Carlos Martinez), RDNL (Ellen Verbakel) and SURF (Carlos Teijeiro Barjas)

CC BY

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

nationaal programma open science