Embedding and Sustaining Research Data and Research Software Professionals in Research Teams and Institutions

Introduction

Embedding and Sustaining Research Data and Research Software Professionals in Research Teams and Institutions

Marta Teperek
Programme Leader FAIR Data at Open Science NL

Lieke de Boer
Scientific Community Manager at the Netherlands eScience Center

Fieke Schoots
Training coordinator Health-RI & Community Coordinator TDCC-LSH

Yan Wang
Head Research Data and Software, TU Delft Library

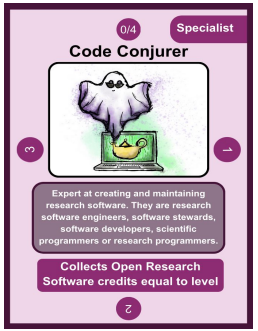
Celia van Gelder
Training Programme Manager at Health-RI, Network Manager of TDCC-LSH

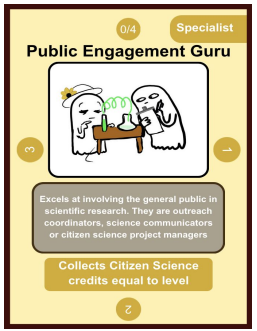
Jeremy Cohen
Advanced Research Fellow and Director of Research Software Engineering Strategy at Imperial college

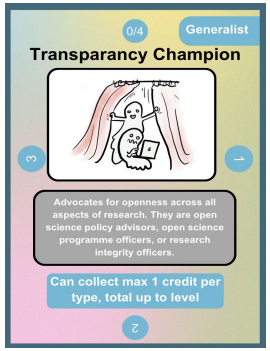
Sarah Coombs
Content Coordinator Digital Competency for Practice-Oriented Research, Research Support Advisor Saxion, Open Science Advisor Vereniging Hogescholen

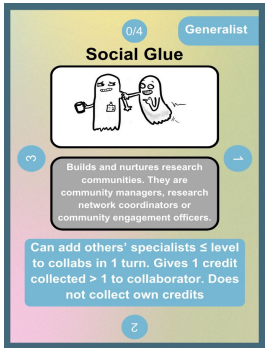
Introduction

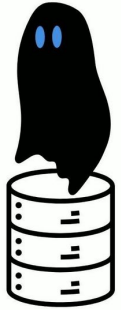
Introduction











https://ghostcollective.github.io

Introduction

Event

Recognition & Rewards

I like what you did there.

8 8 8 8

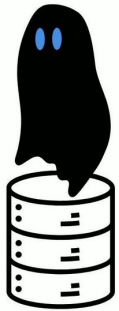
Event

Major Squeeze

Major funding cut-backs.

Lose one role card.

3 3 3 3



https://ghostcollective.github.io

Introduction

Junior data specialist 'cluster Panel- en Survey Onderzoek

Geplaatst	Deadline	Locatie
26 okt	10 nov	Utrecht

Ben jij onze nieuwe data specialist (16 uur per week) die van onze schat aan data nieuwe onderzoeksbestanden maakt waar onze onderzoekers analyses op kunnen loslaten?

Introduction

Data Steward Netherlands Institute for Neuroscience Amsterdam

Published	Deadline	Location
14 Oct	17 Nov	Amsterdam

Introduction

Medior Low Code Ontwikkelaar

Geplaatst	Deadline	Locatie
31 okt	30 nov	Tilburg

Tilburg University | Divisie Library & IT Services is op zoek naar een Medior Low Code Ontwikkelaar, afdeling: Softwareontwikkeling & Labsupport (Development), locatie: Tilburg.

Introduction

Introduction

Research Technician(s) for Spinoff on Robotic Manufacturing of Optical Circuits

Published	Deadline	Location
31 Jul	27 Jan	Amsterdam

Introduction

Team Lead Talent Attraction Team (0.8 - 1.0 fte)

Published	Deadline	Location
31 Oct	15 Nov	Eindhoven

Some observations

- There are many different names for professionals who may do similar things
- There are similar names used for professionals doing widely different things
- Postings suggest some of the following:
  - Job Profiles
  - Appreciation and career perspectives
  - Training & skills
  - Structural embedding
  - Sustainable funding
- Each of these are actually big challenges within hiring institutions that we will discuss today

Data-related roles

Role	Description
Data steward	A person responsible for keeping the quality, integrity, and access arrangements of data and metadata in a manner that is consistent with applicable law, institutional policy, and individual permissions. Data stewardship implies professional and careful treatment of data throughout all stages of a research process. A data steward aims at guaranteeing that data is appropriately treated at all stages of the research cycle (i.e., design, collection, processing, analysis, preservation, data sharing and reuse).
Data analyst	This is someone who knows statistics. They may know programming, or they may be an Excel wizard. Either way, they can build models based on low-level data. Most importantly, they know which questions to ask of the data.
Data engineer	Operating at a low level close to the data, they are people who write the code that handles data and moves it around. They may have some machine learning background.
Data manager	A data manager is a person responsible for the management of data objects including metadata. These people think about managing and preserving data. They are information specialists, archivists, librarians and compliance officers.
Data scientist	A practitioner of data science. It is a generic term that encompasses many fields of specialised expertise. In the current report, data analysts, data stewards and research software engineers are considered as sub-groups of data scientists. In certain contexts, data scientist is also sometimes used in a more limited way that makes it equivalent to either the data analyst or software engineer roles.
Research software engineer	A growing number of people in academia combine expertise in programming with an intricate understanding of research. These Research Software Engineers may start off as researchers who spend time developing software to progress their research or they may start off from a more conventional software-development background and be drawn to research by the challenge of using software to further research.
Research support professional	In the context of digitalisation, these are the people who support scientific researchers conducting data-intensive science. They are not necessarily part of a research team and might be considered as service providers. This is a broad category that can include data stewards, RSEs, data managers, librarians and archivists.

Table 2.1 Roles in the data professional landscape[43]

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

Table from our NPOS report (page 20-21)

nationaal programma open science

Workshop outline

- Opening and introduction
- Pitches: what are the challenges?
- Group discussion
- Recap on the group discussion
- Workshop close

Let's get to know each other

Go to menti.com and enter

1639 8176

Pitches on Current Landscapes

Perspectives:

- Appreciation, recognition & career perspectives
- Job Profiles
- Training & skills
- Structural embedding
- Sustainable funding

Appreciation, recognition & career perspectives (Jeremy)

- Existing role structures in many research-performing institutions follow traditional models

- But research is changing – data, software and computing infrastructure roles represent key contributors to research outputs in the modern era of large-scale multi-disciplinary research and “team science”

- These newer (or previously hidden) roles don’t always align with current institutional role structures

Appreciation, recognition & career perspectives (Jeremy)

- Role structures in many research-performing institutions follow traditional models - newer (or previously hidden) roles don't always align

Grassroots Activities vs. Institution-led Strategies

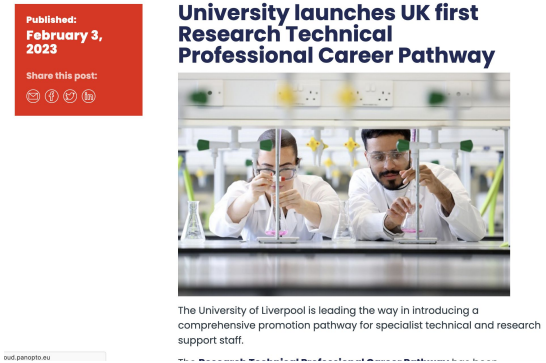
- **Community:** Grassroots communities are vital: **bring people together**, **demonstrate demand** for improved career support and **advocating for change**
    - I lead the local Imperial College London research software community and the regional Research Software London community - both have proved valuable in developing appreciation and recognition

- **Institutional strategy:** In order to effectively support technical careers and underpin future research quality, institutions need to:
    - **Recognise the contribution** of technical roles; **embrace the changing research landscape**; update structures.

- **Community-led initiatives:** Community-led initiatives are key in changing perspectives:
    - **COARA**: Coalition for Advancing Research Assessment (https://coara.eu/)
    - **DORA**: The Declaration on Research Assessment (https://sfdora.org/)
    - Working Groups such as RDA/ReSA convened **PRO4RS**: Policies in Research Organisations for Research Software

Supporting and recognising technical careers: A UK perspective

- **Research Software Engineering (RSE) groups and teams**
  - Formed in 2012 - currently >40 groups in the UK[1]
- **Recognising technical roles**
  - Universities of Liverpool and Warick are leading in developing technical career pathways[2]
- **Stronger recognition driven by funding and clear identifiers**
  - Term “Research Technical Professional” (RTP) increasingly used in the UK community
  - Funders now actively encouraging RTPs as PIs in some calls - DRI programme



[1] https://society-rse.org/community/rse-groups/
[2] https://news.liverpool.ac.uk/2023/02/03/university-launches-uk-first-research-technical-professional-career-pathway/

Towards an official job profile for data stewards

In 2021 the National Programme Open Science commissioned a report. Capacity building for data stewardship is hampered by lack of:

- consensus on responsibilities and tasks
- formal profiles (including knowledge, skills and abilities (KSAs))
- (tailored) education & training

Conclusion: (Inter)national alignment and coordination needed to achieve coherent training and education, accompanied with a consistent human resource (HR) policy

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

Where are we in 2024 with job profiles for data stewards & RSEs

- First definition of responsibilities and tasks of data stewards & RSE (2019-2021) and a national roadmap defined (NPOS report)

- **Universities:** Formal (UFO) job profile for data stewards adopted at Dutch Universities in 2021

- **UMCs**: Radboudumc has local FUWAVAZ implementation of profiles for data steward, data analist & RSE

- **UAS**: DCC-PO has taken job description further, positive advise of VH

- In 2023 the Netherlands eScience Center launched Job Profile and Role Description for Research Software Engineers (RSEs)

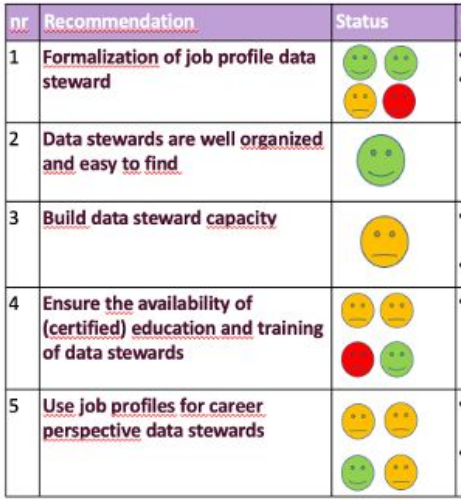
Opportunities & Challenges

Opportunities:

- Contribute to clear(er) task descriptions
- Contribute to transparent and justified grading
- Make position more visible and accountable (also for HR / management)
- Allows career path in organisations

Challenges:

- How to ensure the adoption of the official job profile?
- How to attract and retain talented staff?
- How to overcome the divide between academic and support staff which limits the opportunities for contracts and for getting recognition for contributing to research?
- How does the role fit in the Recognition & Rewards programme?
- No formal training or qualifications (yet)



nr	Recommendation	Status
1	Formalization of job profile data steward	
2	Data stewards are well organized and easy to find	
3	Build data steward capacity	
4	Ensure the availability of (certified) education and training of data stewards	
5	Use job profiles for career perspective data stewards	

Workshop on the implementation of official data stewards job profiles for 3 types of research organisations in the Netherlands (April 2023) Report, Blog

COLUMN *Celia van Gelder*

Data Stewardship: Nederland loopt vooraan, maar we zijn er nog niet

Het wordt me regelmatig gevraagd in de Europese wandelgangen: "Hoe krijgt Nederland dat toch voor elkaar?" Of het nu gaat over FAIR data, Open Science, of de nieuwe rol van data stewardship— Men ziet wat er in Nederland gebeurt en wil graag van onze aanpak leren. Maar hier, in eigen land, heb ik soms het gevoel dat niet iedereen weet heeft van de enorme stappen die gezet zijn.

Een goed voorbeeld is de professionalisering van data stewardship. In steeds meer onderzoeksorganisaties zijn data stewards inmiddels essentiële partners voor de onderzoekers. Ze ondersteunen onderzoekers bij duurzaam en verantwoord databeheer en -deling, zowel binnen onderzoeksgroepen als in ondersteunende afdelingen (bijvoorbeeld in Local Digital Competence Centers (LDCCs)). Waar het domein-specifieke uitdagingen betreft werken zij samen in de context van de Thematic Digital Competence Centers (TDCCs). Gezamenlijk dragen zij bij aan een sterk research data support landschap in Nederland.

Deze vooruitgang is mede mogelijk gemaakt door financiers als NWO en ZonMw, en het Nationaal Programma Open Science (NPOS). Sinds 2019 werken we aan de professionalisering van data stewardship. Samen met belangrijke stakeholders hebben we een competentieprofiel voor data stewards opgesteld, wat heeft bijgedragen aan het tot stand komen van het universitaire UFO-profiel in 2021. Ook UMC's en hogescholen hebben inmiddels flinke stappen gezet.

Naast de betrokkenheid op organisatieniveau is er ook veel activiteit voor en door de data stewards zelf. Sinds 2017 is de Data Stewards Interest Group (DSIG) een levendig platform waar zij elkaar ontmoeten en kennis delen. Sinds 2023 hebben de DSIG en de drie TDCCs de handen ineengeslagen en krijgen domein-specifieke thema's expliciet aandacht. Samen werken we voortdurend aan het verbeteren van de ondersteuning van wetenschappelijk onderzoek.

Maar, we zijn er nog niet. In 2023 hebben we met alle betrokkenen opnieuw de balans opgemaakt. We zijn trots op wat bereikt is, maar de uitdagingen blijven. Er zijn nog steeds te weinig data stewards, en het ontbreekt aan een helder carrièrepad, wat er ook toe leidt dat ze vertrekken naar het bedrijfsleven. Bovendien is er nog steeds een tekort aan adequate training, en moeten data stewards nog regelmatig onderzoekers overtuigen van de noodzaak voor goed data stewardship - iets dat anno 2024 eigenlijk niet meer nodig zou moeten zijn.

Op het gebied van training gebeurt er gelukkig veel. Zo verzorgt RDNL al meer dan tien jaar de cursus Essentials 4 Data Support (E4DS), een instapcursus voor research data support professionals. Vanaf 2025 wordt dit verder uitgebreid onder het Open Science NL programma, en zullen de RDNL-partners een nationaal trainings- en communityplatform ontwikkelen. Hierbij worden uiteraard alle stakeholders betrokken.

Onze grootste troef? De mensen. Data stewards zijn van onschatbare waarde, en het is onze gezamenlijke verantwoordelijkheid om ervoor te zorgen dat we voldoende experts met de juiste vaardigheden en kennis op de juiste plekken krijgen. Bovendien moeten we hen een aantrekkelijk carrièrepad bieden binnen de academische wereld, zodat ze hun expertise blijven inzetten voor het wetenschappelijk onderzoek.

Kortom, we hebben grote stappen gemaakt, maar er is ook nog werk aan de winkel. De ingrediënten zijn er. Laten we samen de handschoen oppakken.

Celia van Gelder is Training Programme Manager bij Health-RI en tevens Network Manager van TDCC-LSH. Health-RI werkt aan het realiseren van een geïntegreerde data-infrastructuur voor data gedreven onderzoek, beleid en innovatie en bouwt ook een Human Capital Agenda op. Health-RI is partner in RDNL, samen met DANS, 4TU.ResearchData en SURF.
www.health-ri.nl/, https://tdcc.nl/

E-DATA & RESEARCH

https://edata.nl/2024/10/21/data-stewardship-nederland-loopt-vooraan-maar-we-zijn-er-nog-niet/

October 2024

Workshop on the implementation of official data stewards job profiles for 3 types of research organisations in the Netherlands (April 2023) Report, Blog

nr	Recommendation	Status	Next steps
1	Formalization of job profile data steward		• Create clarity on the data stewardship tasks • Make open science achievements such as open data or software competitive to make the case for investing in data stewardship for (faculty) boards
2	Data stewards are well organized and easy to find		
3	Build data steward capacity		• Team up with all stakeholders in your organization, including HR staff and senior researchers or deans who can act as ambassadors for data stewardship • Find funds for structural and sustainable data stewardship support
4	Ensure the availability of (certified) education and training of data stewards		• Set up acknowledged professional training for data stewards and explore if a professional association is desirable (cf. software engineering societies or privacy officers associations)
5	Use job profiles for career perspective data stewards		• Make FAIR and data stewardship meaningful by focusing less on compliance and more on societal responsibility • Allow data stewards time to be part of the research team and recognize their efforts as part of the team's work

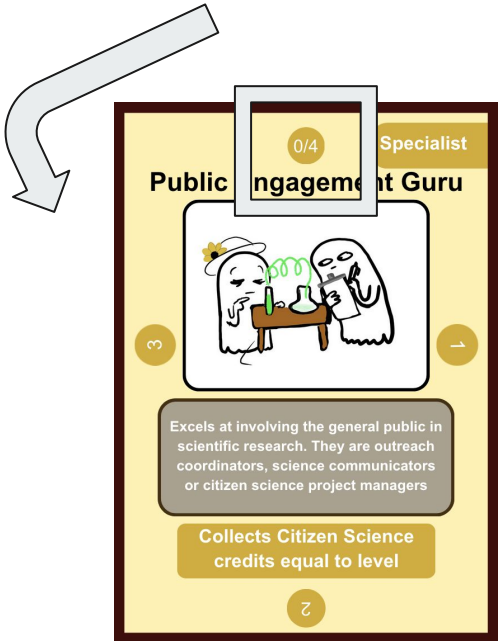
Training and skills

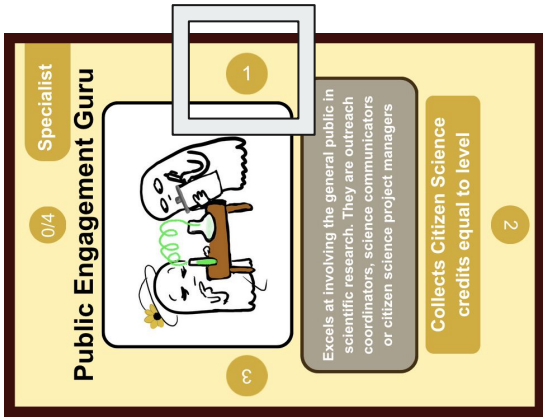
Junior data specialist 'cluster Panel- en Survey Onderzoek

Geplaatst	Deadline	Locatie
26 okt	10 nov	Utrecht

Ben jij onze nieuwe data specialist (16 uur per week) die van onze schat aan data nieuwe onderzoeksbestanden maakt waar onze onderzoekers analyses op kunnen loslaten?

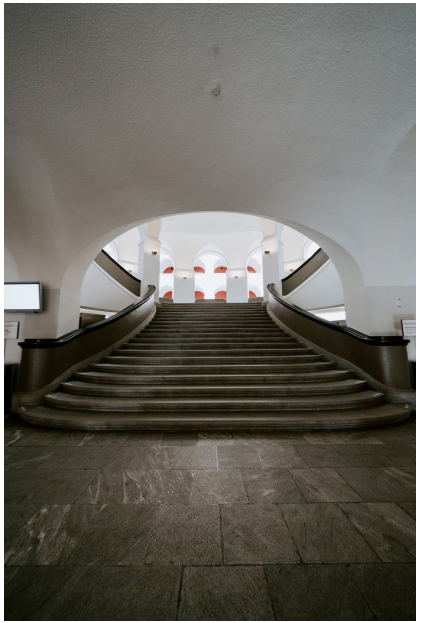
Training and skills





Training and skills

- Embedding professionals in an organization means providing them with a long-term career perspective
- In order for professionals to grow within their careers (up a ladder with increasing expertise and responsibility) we need to be able to train these professionals
- In order to make the right hiring choices at any level, candidates should be able to acquire skills required for the role inside and outside the hiring organization



Claudio Schwartz on Unsplash

Training and skills - Challenges

- It can be unclear what training is available and what training is needed
- A lack of clear expectations from professionals (what do I need to know to grow?) exacerbates this issue
- Even if it is clear what training is needed, there may not be the capacity to deliver training

remi skatulski on Unsplash

Possible ways to address these challenges

Which networks and platforms exist to give an overview of what training is available?

How can professionals join forces so that trainings are optimally delivered across institutes?

Which skills frameworks may be useful in determining what training is needed for different professionals?

Museums Victoria on Unsplash

Structural embedding of Research Support Professional

Institution vs cultural embedding

Initutional:

- Profiles exist in both UAS's & Unie's but are not institutionally standardised across institutions
- Big challenge lies in the discussion around central/decentralised staffing, services, policy, tooling, etc.

Structural embedding of Research Support Professional

Cultural:

Structural solutions for embedding are not there yet and go beyond Rewards and Recognition as they are a reflection of how professional support staff are viewed within both projects and institutions

Can only truly be embedded when both institutional and cultural embedding are achieved and this can only be accomplished when the value of support staff is acknowledged at both project and institutional level

Good Research Support is the difference between Research and Good Research

Sustainable funding

There is some direct funding for roles of data stewards and RSEs.

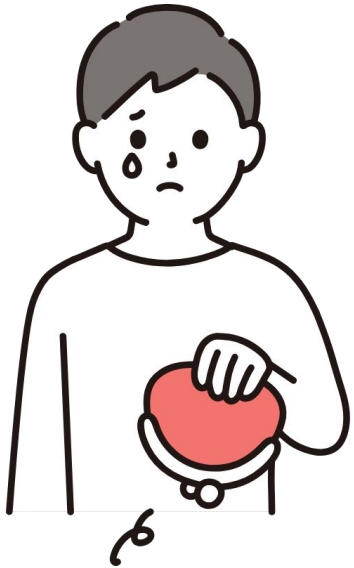
Researchers can also ask for data/software capacity when applying for grants.

So what are the problems?

Even direct funding for such roles (e.g. NWO and OSNL) funding is **short-term**. So institutional backing is essential.

Research funding is **project-based** - often limited to temporary appointments and also at limited salaries, causing j**ob insecurity** and **limiting career growth** of such professionals.

Not all researchers are keen to budget for an RSEs/data manager instead of another researcher.

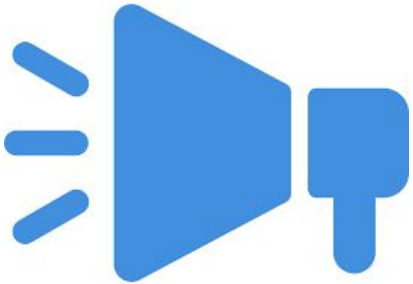


Possible pathways to address these challenges

Getting structural funding from institutions:
How to get the buy in and commitment from the institution to secure structural funding? How to scale?

Changing funding rules
How to focus on research quality and on contributing to a more sustainable ecosystem?

Ensuring that data managers/RSEs are budgeted on grants
How to convince researchers that hiring a data manager or an RSE for a project instead of another PhD student will lead to efficiency gains?



Menti continued!

Go to menti.com and enter

1639 8176

Group Discussions

Menti continued!

Go to menti.com and enter

1639 8176

Next steps

Thank you

Embedding and Sustaining Research Data and Research Software Professionals in Research Teams and Institutions

Marta Teperek
m.teperek@nwo.nl

Lieke de Boer
l.deboer@esciencecenter.nl

Celia van Gelder
celia.vangelder@health-ri.nl

Jeremy Cohen
jeremy.cohen@imperial.ac.uk

Sarah Coombs
s.k.coombs@saxion.nl