sonraí

Irish Data Stewardship Network

Professionalisation of Research Data Stewardship in Ireland

Aoife Coffey, Junwen Luo & Sonraí Project Partners

December 2024

NORF FÓRAM NÁISIÚNTA NA HÉIREANN UM THAIGHDE OSCAILTE NATIONAL OPEN RESEARCH FORUM

HEA AN tÚDARÁS um ARD-OIDEACHAS HIGHER EDUCATION AUTHORITY

Rialtas na hÉireann Government of Ireland

Executive summary

The professionalisation of research data stewardship in Ireland is a critical enabler of advancing Open Research as outlined in the National Open Research Forum's Action Plan for Open Research[1]. FAIR (Findable, Accessible, Interoperable and Reusable) and Open Data are central to achieving this ambition with data stewardship and research data management (RDM) as the mechanism to turn these goals into reality. National and European funding agencies have acted as agents of change in our research landscape and now require formal data management plans. However, practical support for engagement with FAIR and Open Data, the prioritisation of data stewardship skills, professional roles and institutional support remains insufficient.

Current Landscape and Challenges

The limited number of skilled data stewards has been identified as a significant barrier to embedding RDM and implementing FAIR and Open Data. We highlight persistent issues such as inadequate training, insufficient funding and a general lack of awareness and recognition of data steward's role. Without a professional data steward, research quality, reproducibility and compliance suffer, potentially impacting Irelands competitiveness and research excellence. Stakeholders across the research landscape highlight the need for sustainable institutional support and professional pathway for data stewardship roles.

Recommendations

To advance data stewardship professionalisation in Ireland and support development of a sustainable FAIR and Open Data landscape Sonraí proposes the following recommendations

1. **Advocacy** for the critical role of data stewards in the research landscape.
2. **Build** partnerships with key stakeholders' researchers, funders, institutions, infrastructure providers and policy makers, to demonstrate the value and impact of data stewardship.
3. **Adopt** and adapt competency frameworks to standardise data stewardship roles.
4. **Develop** accredited training aligned with international standards to enhance professional skills and recognition.
5. **Engage** with policy bodies to formally recognise data stewardship roles within institutional structures.
6. **Provide** model job description to guide institutions in defining and implementing data steward roles.

Conclusions

A lack of professional recognition, sustainable funding and structured training currently limits the growth and impact of data stewardship in Ireland. Addressing these gaps and challenges through professionlisation, career pathway development and alignment with international standards is essential. By implementing these recommendations, we enhance our research ecosystem with robust and effective data stewardship. Ensuring the embedding and realisation of FAIR and Open Data across the landscape.

This project has received funding from Ireland's National Open Research Forum (NORF) under the 2022 Open Research Fund. NORF is funded by the Higher Education Authority (HEA) on behalf of the Department of Further and Higher Education, Research Innovation and Science (DFHERIS).

Context and Importance of Research Data Stewardship

The professionalisation of research data stewardship in Ireland is critical for advancing our commitment to Open Research as outlined in the National Open Research Forum's Action Plan for Open Research 2022-2030[1]. FAIR and Open Data are core components of realising this ambition, with data stewardship and research data management the mechanism to achieve these aims. Making data FAIR can also maximise the socio-economic impacts of public research funding through licensing and improved access to high quality data . National and European funding agencies all now require a data management plan and active engagement with FAIR and Open Data. It is widely recognised that availability of data stewardship skills across institutions and domains are critical enablers of FAIR and Open Data, and a must-have for the transition to an Open Research landscape.

While the vital need for data stewardship is recognised on the national and international stage, prioritisation of data stewardship skills, professionals and support within institutions is slow in coming. This issue was highlighted at the 'Data Management in Health Research' event held in University College Cork in October 2023 where participants cited lack of infrastructure, key supports including training and data stewards and funding as significant impediments to engagement with RDM, FAIR and Open Data[2]. This is surprising considering the growing body of evidence that when data stewardship is neglected, data

[1] https://doi.org/10.7486/DRI.ff36jz222
[2] https://osf.io/vax37

quality, reproducibility and compliance all suffer[3] [4]. Without availability of research data stewards, we risk our competency and competitiveness, and reputation for research excellence. This is also called out in the European Open Science Cloud Strategic Research and Innovation Agenda[5] they highlight the importance of skilled data professionals for the continued development of Open Research and research quality and competitiveness more broadly. However, it is positive to see that in 2024 a growing number of RPOs have specifically mentioned Open Research, some going as far as referencing to FAIR and Open in their strategic plans[6].

The NORF Landscape Report[7] established that data stewards were a relatively rare feature in our research landscape, although numbers were growing due to the addition RDM requirements placed by funders on the research landscape. To gain better insight into the circumstances in Ireland in 2024, Sonraí undertook a series of interviews with data stewards, librarians, project managers, infrastructure providers, funders and researchers across career stages and domains. A total of 21 interviews were conducted, comprised of 14 female participants and 7 males. Throughout this report excerpts and perspectives from those interviewees have been included.

What is a research data steward? Why is it important?

Data stewardship can be defined as the tasks and responsibilities that relate to the management, sharing, and preservation of research data and other

[3] https://doi.org/10.1038/s41562-016-0021
[4] https://doi.org/10.1186/s13041-020-0552-2
[5] https://eosc.eu/sria-mar/
[6] https://www.ucc.ie/en/president/strategy2028/
[7] https://doi.org/10.7486/DRI.5q485c938

outputs. A research data steward is an emerging professional role that connects researchers with policy, legislation, infrastructure and infrastructure providers through research data management. They can help ensure transparency, accountability and reproducibility of the research outputs and aid researchers to successfully implement the FAIR Principles. Through implementation of FAIR and data sharing they can also play a role in aiding researchers and institutions to manage their intellectual property and developing impact from it. Data stewards can play a significant role in reducing research waste by improving data quality. They also help compliance with the GDPR and other related legislation and policies by translating those policy requirements into actions. They provide a vital link between the researchers and effective implementation of study design, ethics, legislation, storage, infrastructure and post project archiving.

The focus and remit of a particular data steward varies depending on the unit in which they are based. Data stewards based within central units tend not to be experts in a particular data type or discipline but use their skills to support policy, infrastructure development, integration with broader institutional research management, as well as supporting skills and training opportunities for the wider research community. Those data stewards embedded within colleges, schools and research units in general are more focused on the needs of disciplines and often work directly on research data collection and organisation alongside the researchers. However, it is evident that there often overlaps between these task areas[8] and when there is only one data steward

[8] https://doi.org/10.5281/zenodo.4623713

within an institution which is often the case in Ireland key aspects of data stewardship are unsupported and left entirely to the researcher or the level of support required overwhelms the service offered.

> **Interviewee** *"there's definitely kind of a lack of awareness among those doing research data management that they are actually data stewards and therefore they're not valuing it as well as they should."*

> **Interviewee** *"I think the biggest challenge I think Sonrai is addressing this is the invisibility of the topic. It's the hidden work. It's intangible work, even to researchers."*

We know from several studies such as the international State of Open Data survey[9] the national TROPIC survey[10] and through the Sonraí networking activities that resources, rewards, knowledge and confidence in engaging with RDM, FAIR and Open Data are low across the research landscape. This is against a background, a national and international context in which data stewardship skills and data stewards are a critical need rather than a nice to have but in the Irish research landscape they are often absent.

> **Interviewee** *"lack of knowledge … that's the first challenge".*

During the Sonraí interviews several of those interviewed emphasised the potential efficiencies that effective data stewardship can bring to the research system and workflow. They noted that those efficiencies are critical for maintaining competitiveness in funding applications. For researchers' data stewards help save time, resources and improve research outcomes. For

[9] https://doi.org/10.6084/m9.figshare.27337476.v1
[10] https://osf.io/ad4he

institutions they reduce the risk of non-compliance, enhance research process and culture, and help maximise research outputs. For research funders they increase funding impacts by improving the quality and in some cases the number of reusable outputs from funded research.

> **Interviewee** *"it's just critical to like competitiveness and it helps with funding applications as well, because you can say in funding applications that the university already provides these solutions"*

> **Interviewee** *"I think it is a benefit in terms of economizing as well in that if you have specialist like this, they can be used by multiple different research groups within the university so you're kind of potentially saving them duplication of efforts across project"*

What does professionalisation mean?

Professionalisation is the process by which a role or occupation transforms itself into a recognised profession of the highest integrity and competence. It involves the development and recognition of defined skills, competences, identities, norms and values which are identified as core attributes and characteristics of a role. Without professionalisation or the naming of a role you risk tasks falling to a wide range of roles and individuals, making it difficult to identify under whose remit and responsibility a task falls. This can make career paths and progression difficult to define, or even to decide what professional competences, training and skills development are required. In organisations, recognising and rewarding efforts in data stewardship practices on voluntary basis can be challenging when contributions go unnoticed or

unacknowledged, particularly when they fall outside the formal scope of defined roles. This will demotivate researchers and research support staff conducting or contributing to the actual RDM work. For data stewards lack of defined roles or inclusion in formal role descriptions can often mean that the work undertaken is undervalued or under resourced in terms of skills, and capacity.

> **Interviewee** *"I didn't set out to be a data steward, I just ended up as data steward by default because no one else had those competencies within our team"*

> **Interviewee** *"I don't think they really realize that they are data stewards and that they have these skills because it's something that tends to be kind of tacked on to the bottom of the job description"*

> **Interviewee** *"it isn't kind of a distinct or kind of definitive part of my role as such. But it's just because I ... tend to be that person who kind of raises that flag about data"*

In the NORF policy brief 'Coordinated support for data stewardship'[11] it calls for "professionalisation of the data steward role within the Irish research landscape and outline pay scales, competencies and career paths and trajectories for data stewards to enable the sustainable development and retention of core data expertise". Ireland is not unique in considering the professionalisation of data stewardship as critical for the development of FAIR and Open Data ecosystem. Some have move faster than other in this space, in

[11] https://doi.org/10.7486/DRI.js95m610b

the Netherlands data steward is a recognised role within national pay scale[12]. Recently, Open Science NL launched a 1.2-million-euro project to recognise and reward Open Science in recognition that the success of Open Science is not the responsibility of the individual but must be integrated into institutional policies[13]. The European Open Science Cloud Association Strategic Research and Innovation Agenda[5] also recognises that "at the moment, there are not enough adequately trained people to meet current demand". It goes on to say that for data professionals "it is necessary to provide recognition for these roles and define career paths that make them a viable choice" and list 'developing the next generation of Open Science and data professionals' as a priority action. The Research Data Alliance (RDA) Professionalisation of Data Stewardship Interest Group[14] a grass root organisation with over 300 members representing data stewards globally formed in 2021 and in its landscape report also identified that professionalisation was critical to fulfilling the potential of an Open Research landscape. The need for more data stewards and professionalisation as a key enabler was called out as a priority by those interviewed by Sonraí across multiple interviews.

> **Interviewee** *"I think ...every institution will have to make sure they have people and I think plural doing [data stewardship] to a high standard"*

[12] https://tdcc.nl/building-capacity-for-data-stewardship-in-the-netherlands-formal-job-profiles-training-and-career-perspective/
[13] Knowledge institutions take action on open science in recruitment and promotion policies | Open Science NL
[14] https://www.rd-alliance.org/groups/professionalising-data-stewardship-ig/members/all-members/

Career pathways

The RDA Professionalisation of Data Stewardship IG conducted a survey[15] in 2022 which aimed to understand the current and future perspectives of those working as data stewards or in related roles. When respondents were asked how their career will progress several respondents expected their role to evolve over time or to build career progression for themselves. Several respondents also replied that they would make lateral moves highlighting that for some data stewardship is parallel to an existing professional role or an academic or research career path.

The authors of the study also noted that two respondents expressed concern about the future potential of data stewardship as a career path and that those who suggested they would move into specialized roles were often already in professional fields such as librarianship, digital preservation or data governance[15]. Suggesting that while career paths are opening there is still work to be done to define data stewardship as a professional role with potential progression and career opportunities.

We asked those interviewed their perspectives on career paths and prospects for the data steward role, and whether they consider data stewardship as a professional role. The consensus was that there is no established career path for data stewards outside of existing professional roles. Comments from those interviewed were quite compelling on the challenges faced they underscore the urgent need for institutional support, career progression, and recognition

[15] https://doi.org/10.15497/RDA00102

and reward of data stewardship as critical enabler and valued role within the research landscape.

> Interviewee *"I think the reason you hit the ceiling is because…you're not really anywhere or you're in between things"*

> Interviewee *"that trajectory and that career path is something which is missing although it is recognized and there's a lot of talk … about recognition of the value of that role, but it's just not there yet"*

> Interviewee *"there isn't a data steward career path at the moment. So you kind of have to be careful that you don't get siloed into a role that has nowhere to go and isn't recognized"*

Mapping to existing roles and structures

In a recent RDA Plenary *Birds of a Feather* session[16], it was highlighted that it is also not yet clear where within organisational structures data stewards fit or how the role maps to existing professional roles. However, mapping data stewardship against research roles or existing research-support roles may not be advisable as those cohorts have their own issues with recognition and progression. Equally many data stewards and research data support services are currently based in libraries which means for those without a library qualification progression is also difficult. These issues were also highlighted across several of the Sonraí interviews. Suggesting that nationally as we define what a data steward is we also need to figure out where it fits into existing

[16] https://www.rd-alliance.org/session_entry/birds-of-a-feather-session-application-03-07-2024-2/

structures in a sustainable way which enables career progression and development.

> **Interviewee** *"the IUA policy is that postdoc researcher is by its very nature a temporary job that you will eventually grow out of and get a real job … it would be really dangerous for data stewards to map onto those jobs".*

> **Interviewee** *"… either the researcher pay scale or most likely the administrative pay scale, because that gives opportunities for advancement based on how they do their work … they would not be able to advance within the library one because having [a Library] degree is essential"*

Core Competencies

Defining the core competencies for research data stewardship across diverse disciplines is a complex but essential task. As we have discussed, the role of a data steward is often shaped by the specific context of the unit or institution in which they operate. However, there is a foundational set of knowledge and skills common to all data stewards, regardless of their domain. This is also a key step towards professionalisation.

Significant efforts have been made to outline these core competencies. The European Open Science Cloud (EOSC) Skills and Training Working Group[17] has identified a minimum viable skillset for data stewards[18] and the Skills4EOSC

[17] https://eosc.eu/eu-project/skills4eosc-project/
[18] https://doi.org/10.5281/zenodo.14006764

initiative is set to release a comprehensive curriculum[19] further consolidating the key competencies for the field. Similarly, the EOSC report D7. 3: Skills and Capabilities Framework[20] outlined a comprehensive list of competences and capabilities for data stewards, providing valuable insights into essential skills and knowledge areas. The framework identifies nine categories related to research data management, FAIR and Open Data sharing which span the full data lifecycle from planning and design to data publication and preservation, but also includes governance, policy, infrastructure, training and support[21]. The Open Science NL programme in the Netherlands has also undertaken significant work to codify data stewardship competences and skills recognizing that without a common understanding of the role it will be difficult to recruit and retain skilled professionals[21]

As a shared understanding of core competencies is emerging, it is shaping a cohesive vision and laying the foundation for professionalisation of data stewardship grounded in consistent, relevant supports and services across disciplines. In line with this Sonraí is not advocating for a one size fits all version of data stewardship or data stewards but aims to move the discussion forward to a space where a broad church of data professionals and data stewards flourish underpinned by a common understanding of the role, responsibilities and a range of competencies. This is also the approach being called for in the EOSC Strategic Research and Innovation Agenda[5] which aims to establish data

[19] https://www.skills4eosc.eu/news/skills4eosc-project-harmonises-training-curriculum-for-data-stewards
[20] https://doi.org/10.5281/zenodo.5094800
[21] https://doi.org/10.5281/zenodo.4320504

stewardship competency centres formed of cross-disciplinary experts working through networks.

The Role of Training

Dedicated accredited training plays a vital role in professionlisation. It empowers individuals to feel confident in their abilities, enhances recognition and reward, supports professional growth and development and positively influences the perception of others. Training for the wider research community provides a basis to appreciate the activities of data stewards, contribute to them, and to enter the profession. Relevant training is available from many libraries and research support units across Ireland, but it is often high level or introductory and does not meet the needs of those who need to learn more than the principles and basics. In the Sonraí report on training and skills development it is clearly identified that structured accredited training in data stewardship would be welcomed by data professionals but also by the wider research landscape.

> **Interviewee** *"I think having a more kind of structured approach to training would be really useful, just even just to give you the kind of confidence that you know what you're doing"*

> **Interviewee** *"...the specifics for your area, you kind of either try and learn it yourself or you have to try and find somebody else in the area that will help you with this"*

Funding Data Stewardship

It is accepted that funding of data stewards is critical to embedding data stewardship and by extension FAIR and Open Data. While grant funding was mentioned as a potential source by those interviewed many acknowledged relying on it risked another precarious role in the research landscape.

The role of research funders is widely acknowledged as critical in supporting and enabling data stewardship through their programs. Indeed, much of the national progress achieved so far can be attributed changes driven by updates to funder policies both national and European. However, it is also recognised that funders need to move beyond requiring a DMP to incentivizing and reward data stewardship activities and outputs. There is some evidence that this is changing with the use of narrative CVs and a move towards more qualitive research assessment metrics as part of the Coalition for Advancing Research Assessment[22] and national initiatives such as the Roadmap to Embedding Open Research Practices in Ireland (ABOARD)[23].

> **Interviewee** *"[funders] need to continue to with the work that they're doing, but also they probably need to help fund the supports for it and they also need to be pretty clear that what they are actually rewarding is the process and it's not the outcomes, that in fact is research fails and researchers should be allowed fail"*

[22] https://coara.eu/
[23] https://norf.ie/aboard-project/

> **Interviewee** *"So we have to find ways to show the impact datasets. We'll have to think about that and how can we gain, how can we explain impact of the process"*

Providing clear guidance on data stewardship costs and associated roles within funding applications is also an important yet under explored avenue. There is significant variance across the landscape regarding what is allowed under the terms of the various funders. This is coupled with the immaturity of RDM and data stewardship costing models[24] [25] available which are too general to be effectively applied across disciplines and local contexts. A clear understanding and articulation of competencies tailored to the Irish context is essential. Without this it becomes challenging to integrate data stewardship task and responsibilities into roles in a meaningful way often leaving them feeling like an afterthought or being omitted entirely.

> **Interviewee** *"when they have funding calls like, you know, you've got a funding call for 250,000, they could, they could say you have to spend 10% of it or 20% on archiving the data set"*

> **Interviewee** *"putting the jobs out there at the grant stage, understanding how to design data stewardship into jobs"*

However, this implies a continued dependence on short term funding of roles which are tied to specific grants and funding opportunities and is countered across many of the interviews by a recognised need for long term approach

[24] https://www.openaire.eu/estimating-costs-rdm-tool
[25] https://doi.org/10.5281/zenodo.3837717

and sustained support for data stewards and related roles. The tension between trying to build data stewardship into funding call and defining it as an independent role may lie with the ongoing work needed to define the role of a data steward, the data stewardship responsibilities of a researcher and necessary intersection and crossover between the two, which will vary depending on discipline and project. But it is very clear that there is a need for both data stewards and data literate researchers to bring a FAIR and Open Data landscape into reality.

> **Interviewee** *"Start to embed it into the system so that you move away from the dependence on opportunistic funding"*

> **Interviewee** *"we need to find pathways to and find funding so that these roles can be sustained"*

> **Interviewee** *"data stewards as a recognized career located in an institution that supports and nourishes them for the long term"*

Conclusions and the Future

As evidenced by the expert contributions captured by Sonraí from the research community there is still no consensus on competencies and skillsets required for data stewards, nor is their position within organisational structures or career pathways well defined, making career progression difficult. To improve the current situation and build a robust future for research data through effective data stewardship in Ireland an integrated and structured approach is essential. This approach must prioritise sustainable funding, support from funders and institutions, professional recognition and community wide engagement. There is clear agreement across the research landscape that there are not enough data stewards currently. To address this gap and attract skilled professionals, data stewardship in Ireland must be established as a recognised and viable career pathway.

These are the key recommendation:

1. **Recognise** the critical contributions of research data stewards in enhancing research data quality, research integrity and the culture of the research ecosystem and continue to advocate for greater recognition of data stewards nationally.
2. **Build** relationships with key stakeholders such as researchers across disciplines, funding agencies, academic institutions, research infrastructures and industry leaders in the research, public and private sectors to highlight the value and impact of data stewards across domains.

3. **Adapt** and utilize the minimal viable skillset and competency frameworks such as those developed by European Open Science Cloud and the National Open Science Community in the Netherlands to standardise the approach to data stewardship roles.
4. **Provide** accredited training on data stewardship to data stewards and the wider community, which leans into existing resources and international standards ensuring alignment with European benchmarks and fostering professionalisation of data stewardship in Ireland.
5. **Engage** with policy makers, funders and institutions to integrate data stewardship as a recognised and essential role within the research ecosystem.
6. **Create** and disseminate model job descriptions for various types of data steward roles, offering guidance to institutions seeking to define these positions within their organisation.