Data Stewardship Training: Current Landscape and Insights from the Community

Aoife Coffey, Junwen Luo & Sonraí Project Partners

December 2024

Executive Summary

This report highlights the urgent need for skilled data stewards to support Irelands transition to an Open Research landscape. Aligning with the National Action Plan for Open Research[1] and the priority actions outlined in the European Open Science Cloud Strategic Research and Innovation Agenda[2] we identify a critical skills gap in data stewardship and propose actionable steps to address the challenge.

Key Findings

1. There are insufficient data stewards to meet national demand echoing similar European concerns.

2. Current training is fragmented, hidden, high level and does not fully address needs.

3. Training needs encompass two cohorts: data stewards (new and existing) and upskilling the wider research community in data stewardship.

4. Interviews across the research community highlight the need for structured, practical and discipline specific training.

5. Interviews also emphasised the importance of early career training, peer learning and mentorship while voicing concerns about the overwhelming demands on researchers.

6. Opportunities to address training are available through the TROPIC project[4] and collaboration with stakeholders and Skills4EOSC[21].

Recommendation and Next Steps

1. Sonraí proposes developing a microcredential in data stewardship as a flexible, accredited, and job ready training option to address the shared needs of researchers, research support and data stewards.

2. Sonraí recognises the transformative power of community networks for knowledge sharing and professional growth and will continue to build a diverse and vibrant community nationally.

Conclusions

Professionalising data stewardship is essential for Ireland to remain competitive in an Open Research environment. By addressing the skills gap through accredited training, fostering peer learning and strengthening the data stewardship network, Sonraí aims to support a sustainable and vibrant data steward ecosystem that aligns with national and European Open Research landscape.

Introduction

The National Action Plan for Open Research[1] calls for a mature ecosystem to support the full research data lifecycle and highlights the critical role that data stewards and data stewardship skills play in turning this into a reality. The Sonraí Professionalisation of Data Stewardship report highlights that there are not sufficient skilled data stewards nationally to meet current demand. This is echoed in the European Open Science Cloud Association Strategic Research and Innovation Agenda (EOSC SRIA)[2] which states that "at the moment, there are not enough adequately trained people to meet current demand for open and data-intensive science needs, let alone to meet increasing demand and diversity goals". It goes on to say that recognition and defined career paths for data professionals are essential to make those roles a viable choice[2]. They are clear that a coordinated approach is needed and highlight that the fragmented training landscape is a barrier and reduces the impact of current offerings[2]. Under the priority action development of the next generation of Open Science and data professionals 4 out of 6 activities are focused on skills and training[2]. These include developing data skills profiles; recognition of data skills; a quality assurance framework for trainers and trainees; and facilitating and simplifying lifelong learning mechanisms for up-skilling[2]. The NORF Action Plan Goal 5.2 calls for the provision of training for researchers and research

---

[1] https://norf.ie/wp-content/uploads/2022/11/National-Action-Plan-for-Open-Research-webversion.pdf
[2] https://eosc.eu/sria-mar/

support personnel and Action 5.1 call for a national approach to data stewardship training which addresses needs at all levels[3].

These aims are also core to the Sonraí mission which is to enable the development of data stewardship skills across the national research landscape through raising the profile of data stewards, greater recognition of the need for data stewardship, professionalisation of the role of data steward and skills and knowledge development throughout the emerging community. The EOSC SRIA[2] also stated that "It is therefore important to overcome existing gaps and barriers in the necessary skills and training quickly, to reduce the risk of Europe losing a leading position in Open Science". This must be a consideration for Ireland too, we do not want to be left behind in the transition to an Open Research landscape. It is a certainty that a data skills deficit is going to make competing on a European stage difficult for Irish research and will limit the benefits gained from national research investment.

> **Interviewee:** *"good data management is good research. And then if you have good research, you make an impact in the world"*

Sonraí Interviews

To better understand the training needs of data professional and researchers Sonraí conducted a series of interviews with research data stewards, related roles and researchers to explore data stewardship as an emerging profession and identify the support needed for those working as data professionals and those transitioning into the role. The interviews were conducted with data

---

[3] https://norf.ie/wp-content/uploads/2022/11/National-Action-Plan-for-Open-Research-webversion.pdf

stewards, librarians, project managers, infrastructure providers, funders and researchers across career stages and domains. A total of 20 interviews were conducted, comprised of 14 female participants and 6 males. Excerpts from these interviews are used throughout this report.

Who needs training?

When addressing training and skills development in data stewardship, a fundamental question for the Sonraí project has been identifying the primary audience for such training. We find that those with immediate training needs can be grouped into two broad categories. The first includes those already working in data stewardship or related roles such as research support staff, librarians, data managers, lab managers, project managers and others who seek professional development to expand their skillset. The second group comprises researchers at all levels from early career stage to senior researchers, who also require skills development in their fields as the data requirements from publishers, research funders and other stakeholders become increasingly necessary, urgent, and complex. Academics need to be able to efficiently take an active role in research data management (RDM) and data stewardship, so they can ensure their outputs comply with best practice and requirements of FAIR (Findable, Accessible, Interoperable, Reusable) and Open Data.

We know that effective RDM improves research quality, saves time and reduces waste surely these are outcomes we wish for our researchers. Open Research is the new normal and preparedness for an Open Research ecosystem is a necessary skillset. We also find that RDM and stewardship skills

are highly transferable across roles and even fields. We can see that in the varied career paths described by those data professionals interviewed. Their backgrounds include research across a range of disciplines, data infrastructures, librarianship, IT and industry. Implementing mandatory training at a national level could present logistical challenges. However, the Epigeum[4] platform provides a precedent with its research integrity training. When discussing training during the Sonraí interviews several interviewees pointed to the need for training for early career researchers (ECRs). The development of the TRaining for OPen research in an Irish Context (TROPIC) project[5] and Sonraí ambitions suggest there maybe potential for a mandatory data stewardship training program targeting ECRs. This would ensure that our ECRs are well versed in RDM, and data stewardship enabling them to capitalise on a highly transferrable skillset in academia, industry or the public sector. In an area with a precarity issue this is not something that should be dismissed easily but should be seriously considered as a future path for our graduates and those who wish to move on from academia, as reflected by the comments below.

> **Interviewee** *"I feel it should be bundled up in training people how to be good researchers. So, if you know if you're training people how to use particular methodologies, you should also be training them, how to manage, hold their data"*

---

[4] https://courses.epigeum.com/programme?id=107

[5] https://norf.ie/national-open-research-training-programme/

sonraí Irish Data Stewardship Network

> **Interviewee** *"The key is to get people early to understand early in their career what is needed. What the regulations are, what the best practice is so that they can continue"*

> **Interviewee** *"to have something that you think, wait, this is a real skill that I've been doing for years, that I can transfer over to a potential permanent role"*

Training and Professional Development

Data Stewardship Certificate Program – University of Vienna

The Data Stewardship Certificate Program – University of Vienna[6] was developed in the university library in 2022 and is one of the first to offer certification for data stewards. The program offers research performing organisations an opportunity to upskill current research support staff as well as offer an alternative career path for researchers. It's described as helping participants to acquire knowledge, expertise and key skills required to act as data stewards. Modules include the basics of RDM and Open Science, basics of IT and data science, the FAIR data lifecycle, RDM support and project work on data stewardship in practice. Taught through English in a hybrid format part-time it has broad appeal. The first cohort completed the course in June 2023, with 25 representatives from 10 countries, the second cohort of 26 complete their studies in Oct 2024 with 26 participants from 12 countries[6]. The coordinators for the certificate program hope that lessons learned will inform the development of a national certified data stewardship program[6].

---

[6] https://rdm.univie.ac.at/data-stewards-at-the-university/become-a-data-steward/

sonrai Irish Data Stewardship Network

There are other data stewardship programs including a national one in development in Finland[7], though not many are taught through English nor are they available online. The Netherlands has an extensive network of training offerings including Essentials 4 Data Support[8] course, DCC spring training days[9] and Digital Competence Centres[10] as part of the National Coordination Point RDM[11]. But they have yet to establish certified training for data stewards and consider there is still work to be done as discussed at a workshop on data stewardship capacity in April 2023[12] although data steward has been recognised as a formal job classification by the Universities of the Netherlands[13].

Skills4EOSCFellowship program – professional development for data professionals

The Skills4EOSC Fellowship Program[14] was a unique program which aimed to enable learning and knowledge sharing through exchanges and experiential learning between data professionals across different institutions. This call opened for the second time in 2024, both calls targeted data stewards and other data professionals. Fellowships were funded up to 3 months and a total of 8 individuals participated over the duration of the program. The program focused on career development aiming to enhance participants skills in their current role. The fellowship encouraged collaboration, peer learning and

---

[7] https://www.tuni.fi/en/news/university-tampere-receives-funding-training-data-stewards
[8] https://researchdata.nl/en/services/cursus/
[9] https://lcrdm.nl/dcc/
[10] https://lcrdm.nl/dcc/
[11] https://lcrdm.nl/
[12] https://zenodo.org/records/8135134
[13] https://tdcc.nl/building-capacity-for-data-stewardship-in-the-netherlands-formal-job-profiles-training-and-career-perspective/
[14] https://www.skills4eosc.eu/participate/fellowship-programme

engagement in professional networks which may not be available in the participants home institutions. Overall, it had a positive impact with 7 of the 8 finding it beneficial for their career paths[15].

Unfortunately, despite the program's success and strong support from both host and home institutions and its positive impact on competency and skills development for data stewards and data professionals there are no plans to run this program again.

Current National Landscape

The TROPIC Project[5] funded by the National Open Research Forum aims to develop and pilot a national open research training programme to upskill researchers in the fundamentals of Open Research practices. To further this aim, it conducted a national survey on Open Research practices in 2024.

The TROPIC survey[16] found a gap between awareness and implementation in Open Research practices. Further evidence from the survey suggests that the gap is influenced by factors such as skills, knowledge, time, incentives and recognition. But despite the challenges the majority of respondents view Open Research as a positive addition to the research ecosystem. The report also highlights that most respondents have never had formal training in Open Research practices, something that is echoed in the Sonraí interviews but 80% express a strong desire to do so with Open Data one of areas eliciting the highest level of interest.

---

[15] Unpublished D6.4 Skills4EOSC Fellowship Programme Report
[16] https://osf.io/ad4he

Authors of the TROPIC survey[16] concluded that while a standardised approach to Open Research training might seem appealing, those devising training programs need to consider the diverse methodologies within the Irish research ecosystem. Effective programs and strategies need to be put in place to ensure that Open Research practices are embedded across disciplines.

In the Sonraí project we also recognise that data stewardship requires different skills and expertise depending on the role, discipline or organisational unit. Methodological differences have significant impacts on the approach taken to RDM and potential for FAIR and Open Data sharing. Training for these roles must not only align with best practice and principles but also deliver practical, relevant and actionable guidance to support data stewards and researchers effectively in their day-to-day responsibilities.

> **Interviewee** *"So you can make assumptions about transferability that actually don't hold true, and I see that a lot within the qualitative data sciences and particular the way Open Science has been framed at the moment, and it's often framed in ways that actually are in conflict with some of the epistemological underpinnings of qualitative data science"*

The TROPIC Survey[16] also found that librarians and those in other research support roles have a greater level of awareness of Open Research practices compared with researchers. Interestingly this was considered a barrier when it comes to data stewardship by one of those interviewed by Sonraí suggesting that breaking down the silos and embedding these skills across the research landscape is a work in progress.

> **Interviewee** *"it's siloed in the library…. but then librarians often don't have research experience, which I think is a barrier"*

Almost all those interviewed by Sonraí expressed a desire for training that went beyond the current offering either for themselves, their past experiences or for the communities that they support. Several stated that it would give them confidence in their abilities and skills or would help with recognition of the role and progression. They also expressed concerns about limited depth, accessibility and alignment of training with career prospects.

> **Interviewee** *"there's an argument to be made for doing a formalised training course or taking a formalised training course like essentials for research support, which is one thing but actually paying to do a full masters course when there are still, you know, perhaps limited career opportunities available to you. It might be a different prospect."*

> **Interviewee** *"if they exist, the kind of more formal approaches to train and kind of knowledge in this area are at an early stage and you know perhaps we haven't seen the kind of benefits of that yet"*

The challenge of securing buy-in from researchers, institutions and wider landscape on the need and value of training was also expressed.

> **Interviewee** *"convincing people that they need the training might be a real problem"*

A common theme across the interviews was while existing training material were valued and consider of a good standard there was also frustration with the lack of specialised training resources and materials and concern over the

difficulty in translating general content into discipline specific expertise. Some domain specific materials are available examples include the RDMkit[17] for life sciences or the DARIAH-Campus[18] for arts, humanities and social science but there are still gaps and online self-guided resources do not always hit the mark, nor do they align with local circumstances. It is also difficult to find those relevant training resources, almost requiring specialist knowledge to direct researchers to materials, the fragmentation of the training landscape is a key issue called out in the EOSC SRIA[2].

> **Interviewee** *"one of the things I really noticed when I was working in the area was that there was loads of introductory training like you could do lots of kind of introduction to FAIR principles like I did some really good training with GoFAIR and OpenAire and places like that. But it was very kind of surface level and introductory and I kind of felt like I didn't have any option for getting more in-depth training or like filling in knowledge gap"*

> **Interviewee** *"I would say the training is very good for top level understanding....converting that into your area can be quite tricky"*

As we transition to an Open Research landscape the fragmentation of research support and lack of clear roles and responsibilities in relation to data stewardship was highlighted. While rigid boundaries for RDM across a project are likely to be counterproductive; clear understanding of roles and responsibilities are critical.

---

[17] https://rdmkit.elixir-europe.org/your_domain
[18] https://campus.dariah.eu/

> **Interviewee** *"[I'd] like to know what my responsibilities like if I was doing training, what your responsibilities are, the shared responsibilities, like the responsibility of the university at large, this shared responsibility of the different departments in terms of data management, the personal responsibility of the researcher"*

Concerns were raised about data stewardship engagement and the acquisition related skills becoming overwhelming for researchers. Interviewees highlighted the increasing demands on researchers from multiple sources, particularly in a context where there is a shortage of data stewards and other research support professionals. This lack of support can make growing data stewardship requirements feel even more burdensome.

All interviewees emphasised the need for sustained funding and a greater number of skilled professionals to effectively embed data stewardship across the research landscape. This highlights that any training and skilled development initiatives must consider current capacity for engagement at design stage. This includes leveraging the flexibility of technology incorporating best practices in learning and teaching and offering diverse modes of delivery.

> **Interviewee** *'So we are supporting them by telling them, what to do at least, but there is a fine line there between us being part of the solution, being part of the problem"*

sonrai Irish Data Stewardship Network

> **Interviewee** *"I do also feel …… the research job is so difficult at the moment. Do you want to be adding more things that people have to do to get a research job? " ……..I know from talking to some colleagues that and they find it a little bit scary and a little bit overwhelming"*

Another suggestion which came up in multiple interviews is the importance of community for peer learning and support. Creating opportunities for more informal networking and workshopping events should not be undervalued and were a priority activity for some.

> **Interviewee** *"more conversations about Open Science and qualitative research"*

> **Interviewee** *"just having a network with people and being able to facilitate communication between people would be a great start"*

This is reflective of the fact that most interviewees had no formal training in data stewardship specifically but found great benefit in peer learning and mentorship. As Sonraí matures we hope that we can facilitate the growing data stewardship community and continue to build on the successful events over the last two years to create more opportunities for networking, peer learning and mentorship.

> **Interviewee** *"a huge amount of what I've learned was through a more experienced colleague who I was working with. So to have …. some kind of a like a mentoring program or something I think would be really useful"*

sonraí Irish Data Stewardship Network

Conclusions

Both nationally and internationally, there is a pressing need for flexible practice orientated training in data stewardship. However, training alone is not enough, building a strong community, increasing recognition, and fostering networking opportunities are equally essential for the growth and support of the data stewardship community. Sonraí also has a pressing need for a sustainable mechanism to continue its activities.

Next steps

Accredited training

When designing a learning pathway or action plan there are several key considerations which must be clear from the outset these include learner needs, alignment with wider landscape or organisational goals and enablement of effective learning and skills development. As shown in the interviews conducted by Sonraí, the TROPIC survey[16] and priorities set at a European level in the EOSC SRIA[2] it is evident that there is a clear gap in data stewardship training and skills development nationally and in the wider landscape.

As previously discussed, we propose that support and training are needed for two key groups, data stewards new and existing and researchers seeking to develop data stewardship skills. It is our belief that the initial steps in training should not diverge into separate paths for researchers and those wishing to follow a data stewardship career path. Instead, we believe that a unified foundation can address the shared needs of both groups fostering

collaboration and common understanding of the cross over and shared space between researcher and data steward. The availability of formal accredited training is also crucial step towards professionalisation and establishing data stewardship as a viable and recognised career path. For researchers' data stewardship training opens a new potential career path which leans into their disciplinary expertise and capitalised on transferrable data stewardship skills. For data stewards it will standardise core competencies and give them confidence in their role. For institutions it will ensure that critical skills are available and acquired by their community.

As a first step Sonraí Irish Data Stewardship Network proposes to develop a microcredential in data stewardship, strategically addressing the escalating demand for these critical competencies within the national context. We aim that this will be a first point in a data stewardship learning pathway which will branch out into discipline and specialist offerings in the future and is bolstered by networking and peering learning opportunities facilitated by Sonraí. This is also an opportunity for Sonraí to move towards sustainability as the provider of data stewardship training nationally.

Microcredentials offer a flexible means for individuals to upskill, reskill, or specialise in particular areas of interest, contributing to lifelong learning and professional development. The MicroCreds[19] project is a national project backed by the Irish Universities Association which aim to develop a coherent framework of microcredentials to support lifelong learning. They offer learners an opportunity to gain practical job ready skills in a way which is widely

---

[19] https://microcreds.ie/about

recognised and accredited by a university. Accreditation provides validation of competency and expertise, instilling trust among stakeholders such as funding agencies, institutions, and the public. By developing a microcredential in data stewardship, we can aid professional recognition of the role and equip professionals with the necessary skills to responsibly capture, manage, curate, and share data throughout its lifecycle.

To achieve this, we formed a Task and Finish Group which is a time-limited group focused on delivering specific outputs. This group brought together key stakeholders and experts in research, RDM, the FAIR principles, and data stewardship and representative from all research career stages, data professional, librarians and funders including collaboration with the Skills4EOSC Data Stewardship Curriculum[20] working group. Through a series of workshops and sprints, we will collaboratively adapt the Skills4EOSC data stewardship curriculum for this vital microcredential. The group met for the first time in October 2024 and with an aim of October 2025 for the launch of the first microcredential in data stewardship.

Networking and Community Building

Sonraí recogises the importance of creating opportunities for meaningful connections between individuals and organisations. We know that a strong supportive community is essential for long term impact, for promoting shared goals, supporting domain specific and grassroots movements and to recognise and celebrate the growth of data stewardship nationally. We also know that peer learning has transformative power fostering shared experiences,

---

[20] https://www.skills4eosc.eu/news/skills4eosc-project-harmonises-training-curriculum-for-data-stewards

knowledge and growth. Through our activities we commit to acting as a bridge between stakeholders, a uniting network which will continue to build a diverse, vibrant and resilient data stewardship community positioning Sonraí as a champion of data stewardship, a collective voice of data stewardship expertise and a provider of sustainable data stewardship training and development nationally.

To enable these aims we will;

- Continue to promote and grow the Sonraí mailing list and social media channels.

- Facilitate networking event and peer learning opportunities such as workshops, webinars and seminars.

- Develop a master class series for continual professional development.

- Recognise data champions within our community and celebrate their achievements through our website and social media platform.

- Collaborate and encourage partnerships across diverse stakeholder groups to foster cross pollination of ideas, solutions and inspiration.

- Encourage local solutions, leadership and innovation in data stewardship.

Data Availability Statement

The data underlying this report will not be shared as FAIR and Open Data as it is not possible sufficiently anonymise the data to protect the identity of participants.