IJDC | General Article

Data Stewardship through the Lens of Open Science Career Assessment Matrix

Antti M. Rousi
Aalto University

Xiaolong Liu
Aalto University

Lara Ejtehadian
Aalto University

Richard Darst
Aalto University

Pedro E.S. Silva
Aalto University

Udayanto Dwi Atmojo
Aalto University

Abstract

Data stewardship is a key expertise needed for the transformation towards more open and transparent science. This is particularly relevant in research institutions, where data stewards play a direct role in supporting research under open science requirements. However, the absence of established frameworks and merits for assessing this expertise has hindered recognition, professional development, and the integration of data stewardship into institutional practices. This work aims to examine how multidisciplinary data stewardship work transpires through the Open Science Career Assessment Matrix (OS-CAM); a tool designed to assess open science contributions across various dimensions. Using a case study approach, we report findings from a workshop where a multidisciplinary team of experts engaged in data stewardship described their work in relation to OS-CAM. This work presents a summary of the CV narratives and suggested merits for data stewardship developed in the workshop. Assessing data stewardship through OS-CAM provides a structured framework for evaluating, recognising, and rewarding these contributions, thereby increasing their visibility in academic and professional evaluations. However, our study also reveals notable gaps in OS-CAM's coverage of data stewardship, particularly the underrepresentation of infrastructure-related activities such as the management of data repositories. It is important to note that while OS-CAM may offer value in academic research settings, it is less applicable for data stewardship roles that extend beyond research or open science. Therefore, we recommend further research to include diverse institutions and participants, combined with other complementary frameworks, for a more comprehensive understanding of data stewardship's contribution to science and its recognition in or beyond academic communities.

Submitted 10 May 2025 ~ *Accepted* 31 October 2025

Correspondence should be addressed to Antti M. Rousi. Email: antti.m.rousi@aalto.fi

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

Introduction

Data stewardship is a core expertise needed to facilitate the transformation to more open and transparent science (Mons, 2018; Wendelborn, Anger, & Schickhardt, 2023). This expertise involves a wide range of activities, from managing research data to ensuring compliance with open science principles. Researchers and dedicated data stewards in higher education institutions and research organisations often perform these tasks, and their contributions may be presented as professional contributions in curricula vitae. However, due to both the diversity and evolving nature of the expertise, there is a lack of specific career merit frameworks on how their contribution should be assessed and recognised. Furthermore, many data stewardship activities are performed by researchers who are not formally identified as full-time data stewards, which often results in these contributions being poorly articulated or underreported. In this context, given the close relation of data stewardship and the principles of open science, established open science assessment frameworks offer a promising pathway to explicitly define and evaluate data stewardship activities while also highlighting their value in professional development and career advancement.

One of the key assessment frameworks of open science is the European Commission (EC)'s Open Science Career Assessment Matrix (OS-CAM) (EC et al., 2017). This framework, consisting of six open science activity categories, serves as a lens through which diverse research contributions can be evaluated. By offering practical guidance for researchers on how they could present their open science activities as a career merit, OS-CAM broadens traditional notions of research assessment. Applying this matrix to data stewardship could thus illuminate how such activities are manifested in research practice and how they may be better acknowledged in academic evaluation systems.

This work aims to elaborate on how multidisciplinary data stewardship practices can be assessed through the lens of OS-CAM. To investigate this, a case study was conducted during a workshop involving a multidisciplinary team of experts, including researchers serving as part-time data stewards, research software engineers, IT research data solution owners, and research data management specialists. Participants were tasked to write short, CV-style narratives describing their work, mapped to the topics and evaluation criteria found in OS-CAM. The findings highlight which OS-CAM topics received narratives and merits suggestions, as well as instances where data stewardship activities fell outside the existing OS-CAM topics. The following section introduces relevant background and the OS-CAM framework. Afterwards, a description of the case study is presented, followed by the main results from the workshop. The final part discusses the findings and elaborates on how OS-CAM may be used as a framework to assess data stewardship activities.

Background

Data Stewardship

Data stewardship refers to the supervision and accountability for data assets within an organisation, ensuring their quality, accessibility, and value throughout their lifecycle (Mons, 2018; Bardel et al., 2023). As the custodians of data, data stewards play a crucial role in promoting and implementing effective data management practices, which encompass a range of activities such as data creation, storage, sharing, archival, and disposal. These activities emphasise not only the technical aspects of data handling but also the legal and ethical considerations underlying the multifaceted nature of data stewardship (Wendelborn, Anger, & Schickhardt, 2023). The competences representing the crucial areas where data stewards should direct their expertise may be categorised as (1) Data Technical Competences, (2) Legal and Ethical Competences, (3) University Domain-Specific Competences, (4) Data Analysis and Interpretation Competences,

and (5) Communication, Collaboration, and Project Management Competences (Fitsilis et al., In press). Given the diversity of this expertise, no single person can undertake all responsibilities involved in comprehensive data stewardship (Wendelborn, Anger, & Schickhardt, 2023).

To tackle challenges arising from the demand for diverse expertise, researchers in higher education institutions and research organisations, coming from disciplines such as computer science, life sciences, social sciences, and humanities, frequently undertake data stewardship tasks that significantly enhance the quality and integrity of scientific research (Fitsilis et al., In press). Their ability to integrate technical skills, legal and ethical awareness, and domain-specific knowledge allows them to manage data effectively within and across disciplines.

Recognising the importance of these contributions, guidelines such as the Finnish National Board on Research Integrity (TENK)'s Template for Researchers' Curriculum Vitae recommend that these activities be explicitly documented as professional merits (TENK, 2024). This practice not only validates the critical role of data stewardship in fostering transparency and reproducibility but also ensures that such contributions receive appropriate recognition in career evaluations and academic promotions. Despite its importance, the recognition of data stewardship as a professional merit within the academic and research communities is still evolving (Demchenko & Stoy, 2021; Fitsilis et al., In press; Mons, 2018; Wendelborn, Anger, & Schickhardt, 2023). This evolution is partly hindered by the lack of standardised framework for assessing and acknowledging the researchers' diverse data stewardship activities as part of their professional contributions. Building on the TENK guidelines, this study conceptualises data stewardship merits as documented instances of such activity, characterised by quantifiable indicators of quality.

The Role of Data Stewardship in Open Science

Open science advocates for a broad range of practices aimed at making scientific research more accessible, reproducible, transparent, and collaborative (EC, 2015). These practices include open data, open software, open methods, open access publishing, open peer review, and open educational resources, among others (Pontika & Knoth, 2015). At the heart of many of these practices lies the effective management and sharing of data, which is where data stewardship becomes integral. Given the critical role of data in contemporary research, data stewardship has emerged as a pivotal expertise in facilitating the transition to open and transparent science (Aksenova et al., 2024; Fitsilis et al., In press). The responsibilities of data stewards closely align with open science principles, establishing data stewardship as a cornerstone of open science. Through their expertise in managing, documenting, and sharing data, data stewards substantially contribute to the realisation of open science principles. Recognising and evaluating the contributions of data stewardship is crucial for promoting professional growth and sustaining open science practices (EC et al., 2017). Additionally, the intersection of data stewardship and open science presents a valuable opportunity to assess their combined impact through the use of established open science evaluation frameworks.

Open Science Career Assessment Matrix

Research assessment is moving away from publication and journal-based metrics to acknowledging the diversity of research practices and outputs (Arentoft et al., 2022). A key theme of the Coalition for Advancing Research Assessment (CoARA) is that the quality of research is demonstrated through transparency of used methods, reproducibility of results, and openness of research (Arentoft et al., 2022). Furthermore, CoARA suggests that research assessment should be based on qualitative review of candidates, which is supported by quantitative indicators and metrics, if appropriate. Importantly, one of CoARA's commitments is that future assessments would recognise and appreciate diverse careers in research, including working as a data steward or research software engineer (Arentoft et al. 2022).

Research funders are developing novel assessment frameworks to aid this transition in research assessment. One of these novel assessment frameworks is the EC's OS-CAM (EC et al., 2017). OS-CAM offers a more comprehensive approach to evaluating research activities, placing emphasis on open science principles. It outlines specific criteria for assessing research output,

research processes, and teaching-related open science practices, with the goal of providing practical guidance to researchers on how to present their open science contributions as part of their professional contributions or career achievements. The OS-CAM framework consists of six key open science activity categories, each subdivided into subtopics (EC et al., 2017). Table 1 summarises these open science activity categories, highlighting the various areas where researchers can demonstrate their commitment to open science practices.

Table 1. Open Science Career Assessment Matrix's open science topics (EC at al., 2017).

OS-CAM activity categories	OS-CAM subtopics
Research output	Research activity
	Publications
	Datasets and research results
	Open source
	Funding
Research process	Stakeholder engagement / citizen science
	Collaboration and interdisciplinarity
	Research integrity
	Risk management
Service and leadership	Leadership
	Academic standing
	Peer review
	Networking
Research impact	Communication and dissemination
	IP (patents, licences)
	Societal impact
	Knowledge exchange
Teaching and supervision	Teaching
	Mentoring
	Supervision
Professional experience	Continuing professional development
	Project management
	Personal qualities

Case Study Design

Case Study Participants

Data collection of our case study was based on materials produced during a workshop that involved a multidisciplinary team of experts specialising in data stewardship. The workshop took place at Aalto University, Finland. Formed in 2010, Aalto University has four schools that focus on technical sciences, one school for business and management, and one for arts and design. In

2023, total personnel count was c. 4,900 and degree student count c. 15,000.[1] The case team of experts that took part in the workshop were all employed by Aalto University. Table 2 presents the backgrounds and roles of the workshop participants.

Table 2. Backgrounds and roles of the workshop participants.

Role	Description	Number of participants
Researchers as part-time data stewards	Aalto University has a programme of researchers functioning as part-time data stewards. The data stewards are a multidisciplinary group representing different departments[2]	9
Research software engineers	Research software engineers are full-time service personnel who support software, computing, and data[3]	3
Research data specialists	Research data specialists are full-time service personnel residing in Research Services who work on research information systems, open access publishing, and data management planning	3
IT research data solution owners	Full-time IT expert personnel for research data management solutions	2

Case Study Protocol and Analysis

The participants were divided into field-specific groups consisting of computational research and data science, electrical engineering, materials science, medical information and communication technology (ICT), social and behavioural sciences, and water and civil engineering based on the expertise of the workshop attendees. Within their field-specific groups, participants were asked to choose one to three subtopics from the OS-CAM matrix and write short open science narratives, including applicable merits, understood here as professional contributions corresponding to the evaluation criteria in OS-CAM, with a focus on data stewardship. These merits are seen as documented instances of data stewardship activity, characterised by quantifiable indicators of quality. Alongside the OS-CAM matrix, the work of Torres-Salinas et al. (2024) on narrative bibliometrics was presented as background for the workshop participants.

After the narratives and merits were collected, all the workshop materials were collected for analysis. Although individual narratives were initially categorised under specific OS-CAM topics, one collected narrative could also be interpreted to represent many of the OS-CAM subtopics. The findings present our best interpretation of the most applicable subtopic for each narrative. In some instances, the merits and the narratives were divided into several OS-CAM topics and are reported in the applicable sections. The OS-CAM subtopics present in the original framework

[1] Aalto University key figures and annual reports: https://www.aalto.fi/en/aalto-university/key-figures-of-2024-and-annual-reports
[2] Aalto Data Agents: https://www.aalto.fi/en/services/data-agents
[3] Aalto Research Software Engineers: https://scicomp.aalto.fi/rse/

that were not referred to in the narratives were excluded from the findings. The analysis also revealed data stewardship activities that extend beyond OS-CAM's current scope, which are reported in a dedicated section within the findings.

Selected CV narratives generated in the workshop are published as examples illustrating how data stewardship activities can be represented in researchers' or data stewards' CVs (see Appendix). Following the principles of personal data minimisation, all unnecessary identifiers were removed from the narratives.

Findings

Table 3 summarises all suggested merits per OS-CAM topics that could be used for assessing data stewardship. What follows are summaries of the workshop materials categorised into OS-CAM topics.

Table 3. Summary of suggested merits related to OS-CAM matrix for assessing data stewardship. OS-CAM activity categories and subtopics with no suggested merits omitted.

Research output	Merits
Research activity	Number of publications and other outputs investigating research data and software management, open science policies and guidelines, etc.
Publications	Number of accesses Number of downloads Number of citations (e.g., Scopus) Citation percentile within the database (e.g., Scopus) Altmetric Attention Score Prereview.org reviews Speaker invitations (based on publications)
Data sets and research results	Number of published data sets Number of accesses Number of downloads
Open source	Number of published software Release platform (e.g., GitHub, PyPI) statistics, e.g., GitHub stars Unique contributors outside the primary team Number of opened issues Number of libraries.io software dependencies (e.g., to three other software packages)
Research process	Merits

Stakeholder engagement / citizen science	Research data management in projects with business collaboration
Collaboration and interdisciplinarity	Research data management in projects with interdisciplinary collaboration
Research integrity	Position as Research Ethics Advisor Amount of teaching and training given in research integrity Number of conference presentations in specialised conferences (e.g., ENRIO) Partaking in ReProhacks and work to reproduce scientific results of others Number of studies with pre-registrations (registered reports, clinical trials) Number of studies with PRISMA literature reviews
Risk management	Number of projects with, e.g., sensitive personal data or environmentally hazardous materials Laboratory and instrument administration

Service and leadership	Merits
Academic standing	Position as a data steward within one's higher education institution
Leadership	Partaking in forming and developing institutional and national research data management policies, frameworks and guidelines. Level and amount of policy work
Peer review	Number of reviews for software projects
Networking	Participating in national and international networks related to data stewardship

Research impact	Merits
Communication and dissemination	Number of projects with patient and public involvement in research design
Societal impact	Altmetric score News coverage

Teaching and supervision	Merits
Teaching	List of courses or lectures given on research data management or related topics

	Invited talks
	Number of participants
	Number of open materials
	Number of downloads on open materials
Mentoring	Mentored research groups and colleagues

Research Output

Some workshop participants actively pushed the boundaries of open science by turning their focus toward studying and publishing on open science itself, for example, with the topic of scientific journals' research data policies. In parallel, researchers working on a wide range of topics are also integrating open science principles into their work. Several participants highlighted the importance of disseminating research articles openly, either through publishers' open-access options or institutional repositories. Quantitative merits such as citations (e.g., Scopus), Altmetric scores, and download counts were highlighted as indicators of impact.
In addition to publications, the participants emphasised the role of openly accessible datasets, often hosted on platforms like Zenodo. The link between research papers and datasets was particularly notable, as seen in multiple narratives where datasets were explicitly linked to published articles. This practice supports data transparency and reuse, not only reinforcing OS-CAM's recognition of open data as an essential research output but also seeing open data sets as a demonstration of practical and impactful data stewardship.

One narrative detailed the release of research software on GitHub, including quantitative metrics such as stars, external contributors, issue tracking, and software dependencies. While these indicators provide insights into software adoption, the participant noted that collecting such metrics remains a manual process. As with open data, the release of research software with usage metrics can be also seen as a demonstration of practical data stewardship.

No narratives explicitly addressed funding in the context of open science, such as grants awarded for open research initiatives. While open science initiatives increasingly receive dedicated funding, researchers may not explicitly link their funding sources to open research outputs in CV-style narratives. Moreover, grants often support multiple aspects of research, making it difficult to isolate open science-specific funding contributions in a concise narrative format. Similarly, while OS-CAM includes research activity as a subtopic, the collected examples focused on tangible outputs rather than broader research processes or methodological openness.

Research Process

The research process related to the CV narratives highlighted the importance of engaging stakeholders in research, particularly in materials science, where involving citizens through peer review and feedback ensures the research is aligned with societal needs. Citizen science was emphasised to democratise research, making it more relevant and credible. This reflects OS-CAM's focus on the societal impact and inclusiveness of research. As inclusive data collection and citizen science expand, the associated data stewardship skills will likely be regarded as valuable academic competencies.

Materials science was seen to benefit from interdisciplinary collaboration, extending across not only technical fields like physics and chemistry but also arts and humanities. Proper data stewardship plays a crucial role in fostering this interdisciplinarity by ensuring that data are well-managed, accessible, and documented in a way that supports seamless collaboration across different disciplines. This collaborative approach enriches the field's societal contributions by integrating diverse perspectives and expertise, which in turn drives innovation and leads to solutions that are more responsive to complex societal challenges.

Participants emphasised the importance of research integrity, particularly in medical sciences, through adherence to established guidelines on reviewing and reporting like Equator (Altman et

al., 2008) and the PRISMA 2020 statement (Page et al., 2021). Efforts like pre-registration and clinical trial transparency were mentioned as measures to reduce risks such as falsification and plagiarism. The challenge of balancing open science with data protection was also noted, especially when dealing with sensitive data (e.g., study participants).

Risk management narratives focused on the ethical and legal aspects of data handling. In materials science, managing risks such as environmental impact and data misappropriation was seen as essential for maintaining research integrity through practical measures, such as chemical-inventory control, equipment maintenance, and researcher training. Similarly, the social and behavioural sciences focused on data protection, including Data Protection Impact Assessments and securing permissions for data sharing. In general, the well-documented link between data stewardship and data protection and responsible research (e.g., Wendelborn, Anger, & Schickhardt, 2023) was also visible within the materials produced in the workshop.

Services and Leadership

The narratives on service and leadership emphasised academic standing in promoting open science and research data management at institutional levels. Participants took active roles in guiding younger researchers, co-authoring open science handbooks, and contributing to university-wide open science policies. Service work in committees and working groups further demonstrated leadership in shaping institutional strategies for open research.

Data stewards play a role in creating and implementing institutional data management policies. The findings suggest the importance of encouraging institutions to involve data stewards in relevant policy formation and decision making. This involvement can support robust data governance and compliance, enhancing the quality and integrity of research through well-formulated policies and guidelines.

Participants mentioned contributing to open peer review by participating in the Research Data Alliance and Aalto University's open science policies specifically on software-related results, though specific narratives on peer review engagement were limited.

The importance of networking was evident through active participation in national and international working groups and collaborative initiatives. These actions demonstrate how networking can support data stewardship by connecting researchers across disciplines and institutions.

Research Impact

In medical ICT, research impact narratives highlighted the importance of public engagement, especially through patient involvement and public collaboration in study design. Involving participants early in the study design process fosters open communication, a core principle of open science. However, measuring impact through patient involvement remains challenging, as typical metrics like media attention (e.g., Altmetric scores, news coverage) are difficult to standardise and track.

Intellectual property (IP) was not mentioned, which could indicate either the research areas of participants were not patent-driven or that the focus on open science practices overshadowed commercialisation aspects. Similarly, knowledge exchange beyond academia was not directly addressed, which reflects the fact that the participants were primarily university employees who have been more focused on internal institutional activities rather than broader industry partnerships or non-academic collaborations.

Teaching

Participants integrated data stewardship and open science principles into teaching, curriculum development, and mentoring. For example, lecturers in research methods emphasised the interconnection of research ethics, data management, and data protection, while webinars on practical open data management attracted significant viewership. Open learning materials (e.g.,

under CC-BY licences) and hands-on support for early-career researchers further illustrate how data stewards promote transparent research practices.

The aspect of mentoring was also included in the CV narratives on data stewardship. Experienced researchers can actively promote good data management practices and transparent research methods within their teams and encourage early-career researchers to openly share their research data.

Professional Experience

Although some of the produced narratives were categorised under the OS-CAM topic of professional experience, their scrutiny revealed that they were more collections of actions that could be categorised under the other OS-CAM topics, such as Service and leadership and Teaching, and are reported in the corresponding sections. Some of the content classified under professional experience, such as repository and infrastructure work, had no clear place in the OS-CAM activity classification. What follows is the reporting of topics found in the data stewardship narratives that had no clear place within the OS-CAM matrix.

Topics Outside of OS-CAM

Several narratives described activities that extend beyond OS-CAM's current scope. One participant detailed the creation and maintenance of a dedicated research software repository, which has archived multiple research outputs since 2018. Another contributor, acting as an application owner, emphasised developing and supporting laboratory information management systems (LIMS) and research data management tools such as the Aalto Materials Digitalization Platform[4] and Aalto Electronic Laboratory Notebook,[5] while also providing technical training and support for cloud-based data storage and computing.

Another area not captured by OS-CAM is the development and implementation of data management plans. Narratives highlighted activities such as reviewing and commenting on data management plans for major funding applications (e.g., Research Council of Finland, Horizon Europe) and formulating institutional guidelines to govern the research data lifecycle.

Discussion

This study's findings have important implications for recognising and promoting the professional development of data stewards and researchers engaged in data stewardship activities. The participating multidisciplinary team of experts linked their data-related activities to a majority of OS-CAM topics, including research outputs, research processes, service and leadership, research impact, and teaching. Assessing data stewardship through OS-CAM thus provides a structured framework for evaluating, recognising, and rewarding these contributions, thereby increasing their visibility in academic and professional evaluations. This approach supports career advancement by offering tangible merits to document and showcase efforts. By mapping these activities to OS-CAM, our analysis also demonstrates that robust data stewardship significantly enhances open science practices—ranging from improved data sharing and collaborative research to the creation of open educational resources. Consequently, integrating such a framework can promote more rigorous data management practices and adherence to open science principles, ultimately benefiting the entire research community.

While OS-CAM provides a valuable lens for assessing and recognising data stewardship, our analysis also revealed notable gaps in its current scope. Infrastructure-related activities—such as the

[4] Aalto Materials Digitalization Platform: https://www.aalto.fi/en/services/aalto-materials-digitalization-platform-amad

[5] Aalto Electronic Laboratory Notebook: https://www.aalto.fi/en/services/aaltos-electronic-laboratory-notebook-aalto-notebook

management of data repositories and the development of research data management tools—are vital to the backbone of open science yet remain underrepresented within the current OS-CAM topics. Acknowledging infrastructure-related contributions of data stewards can incentivise them to more actively engage in the development and maintenance of research data management infrastructures. Similarly, the formulation and implementation of data management plans, which are critical for ensuring data quality, compliance, and integrity throughout the research lifecycle, are not fully acknowledged by OS-CAM. Moreover, certain data stewardship tasks, such as internal data governance and compliance with institutional or funding body regulations, may not align strictly with open science principles. As a result, relying solely on OS-CAM for the assessment of data stewardship might overlook significant contributions that are essential to effective research data management, even if they fall outside the conventional open science paradigm.

Beyond these thematic gaps identified from the workshop, another emerging area insufficiently captured by OS-CAM is the impact of artificial intelligence (AI) on research. AI developments can potentially have a significant impact on the field of data stewardship (Azeroual, 2024). AI can support data stewardship work by automating the collection and processing of large and/or complex datasets, making it easier to identify and correct errors, duplicates, and inconsistencies in datasets and helping ensure compliance with legal and ethical standards (Azeroual, 2024). One major area that AI could positively impact is automated data management planning, from creation to review of data management plans. Also, AI tools can mentor new researchers in data management practices and assist with routine tasks such as metadata generation, cataloguing, and documentation, therefore giving data stewards more time to focus on specialist work. However, most of the above (AI) applications are not yet widely available, proven at scale, or routinely implemented in institutional workflows. By securely and consciously integrating AI, research institutions could support their data stewardship practices in alignment with OS-CAM and advancing open science principles. Given the rapid pace of progress, institutions should track developments and experiment cautiously with secure, well-governed integrations that align with OS-CAM and open science principles, while further research clarifies benefits, risks, and appropriate safeguards.

In addition to AI-related developments, data stewardship also overlaps with other emerging professional roles, such as research software engineers (RSEs). In practice, RSEs frequently take on responsibilities beyond software development, and contribute to data stewardship through practices that involve data management, enabling data sharing and reproducible research practice in their projects. While an RSE typically has a broader, more technical mandate focused on implementation, a data steward often concentrates on data content, interoperability, and long-term management. Our analysis suggests that OS-CAM can also be useful when assessing RSE activities. Within the OS-CAM context, research outputs and research process were especially relevant topics for RSEs. For example, RSEs may have a larger role in producing and releasing datasets, and a very large role in all open-source merits. RSEs assist in collaboration and interdisciplinarity by bringing computational tools into new research fields. They also contribute to research integrity; not only by ensuring legal and ethical compliance for software-based tools, but also by promoting the general reproducibility of computational science.

While these discussions highlight the multifaceted relevance of OS-CAM, our findings are based on a single case study conducted at Aalto University with a specific group of participants in a few specific research fields, which limits the diversity of our sample and the generalisability of the results. These limitations indicate that the framework requires refinement to comprehensively capture the full spectrum of data stewardship activities. Future studies should include a broader range of institutions, geographical regions, and participant profiles to validate and refine these insights. Additionally, the reliance on self-reported narratives may introduce biases such as social desirability or selective memory. Future research could incorporate objective, quantifiable measures alongside qualitative assessments, and longitudinal studies to capture the evolving nature and long-term impact of data stewardship roles. An expanded and refined OS-CAM that incorporates detailed merits for data stewardship would provide clearer guidelines for presenting these contributions as career achievements. Such a framework could encourage institutions to formally recognise and reward data stewardship activities, paving the way for dedicated career paths and enhanced professional growth. By capturing a broader range of contributions beyond

traditional research outputs, the framework would also promote more rigorous data management practices and deeper adherence to open science principles. Furthermore, these improvements are particularly relevant for part-time data stewards who primarily engage in research yet perform substantial data stewardship work. A refined framework that acknowledges both research and data stewardship contributions could help ensure that these dual roles are recognised in career evaluations, funding applications, and promotion criteria, thereby supporting a more balanced workload and clearer career advancement opportunities.

Finally, looking beyond OS-CAM, several complementary initiatives and frameworks provide useful comparators and potential implementation pathways for OS-CAM. The EOSC Skills and Capabilities Framework (Whyte et al., 2018), "Professionalising Data Stewardship in the Netherlands: Competences, Training and Education" (Jetten et al., 2021), and "Data Stewardship on the Map" (Verheul et al., 2019) are largely role- or task-based and do not incorporate specific career merits, making them complementary to OS-CAM's output- and contribution-focused perspective. In parallel, the EU-funded Open Research Assessment Dataspace (GraspOS)[6] and Open Universal Science (OPUS)[7] projects are developing responsible, open assessment frameworks, indicators, and infrastructures that make contributions beyond traditional publications—such as data stewardship, software, and open practices—visible and citable. Together, these resources can inform how OS-CAM-aligned data stewardship activities are documented, evaluated, and rewarded in hiring, promotion, and funding; a future study could map competencies to OS-CAM categories and test GraspOS/OPUS tools for integrated assessment.

Conclusion

Assessing data stewardship through OS-CAM provides a structured framework for evaluating, recognising, and rewarding these merits, thereby increasing their visibility in academic and professional evaluations. This approach supports career advancement of data stewards by offering tangible contributions to document and showcase their activities. By mapping data stewardship activities to OS-CAM, our analysis also demonstrates that robust data stewardship significantly enhances open science practices. However, to fully capture the scope of data stewardship, OS-CAM must be refined to address identified gaps such as infrastructure-related activities, data management plans, and AI's impact on data management. It is important to also note that while OS-CAM may offer value in academic research settings, it is likely less applicable for broader data stewardship roles that extend beyond research or open science. We recommend further research to include diverse institutions and participants, combined with other complementary frameworks, for a more comprehensive understanding of data stewardship's contribution to science and its recognition in or beyond academic communities.

Acknowledgements

The authors thank the Aalto Open Research Network members that participated in the workshop and produced the CV narratives analysed in this study. The authors also thank Anna-Kaisa Hyrkkänen and Tomi Toikko from the Management Information Systems team at Aalto University.

[6] Open Research Assessment Dataspace (GraspOS): https://graspos.eu/
[7] Open Universal Science (OPUS): https://opusproject.eu/about/

Data Availability

Selected CV narratives are included in the Appendix.

References

Aksenova, A., Johny, A., Adams, T., Gribbon, P., Jacobs, M., & Hofmann-Apitius, M. (2024). Current state of data stewardship tools in life science. *Frontiers in Big Data* 7, Article 1428568. doi:10.3389/fdata.2024.1428568

Altman, D. G., Simera, I., Hoey, J., Moher, D., & Schulz, K. (2008). EQUATOR: reporting guidelines for health research. *Open Medicine* 2(2), e49-50. https://pmc.ncbi.nlm.nih.gov/articles/PMC3090180/

Arentoft, M., Berghmans, S., Borrell-Damian, L., Bottaro, S., Faure, J.-E., Gaillard, ... Stroobants, K. (2022). *Agreement on Reforming Research Assessment.* Zenodo. doi:10.5281/zenodo.13480728

Azeroual, O. (2024). Smart Data Stewardship: Innovating Governance and Quality with AI. In L. Gruenwald, E. Masciara, & J. Bernardino (Eds.), *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, Vol. 3: KMIS* (pp. 187–196). doi:10.5220/0012918200003838

Bardel, A., Hasani-Mavriqi, I., Jean-Quartier, C., Schaffer, P., Macher, T., Kalová, T., ... Piroi, F. (2023). Data Stewardship—Austrian National Strategy and Alignment. *Zeitschrift für Hochschulentwicklung* 18 (Sonderheft Forschung), 65–88. doi:10.21240/zfhe/SH-F/05

Demchenko, Y., & Stoy, L. (2021). Research data management and data stewardship competences in university curriculum. In T. Klinger, C. Kollmitzer, & A. Pester (Eds.), *2021 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1717–1726). doi:10.1109/EDUCON46332.2021.9453956

European Commission Directorate-General for Research and Innovation. (2015). *Open innovation, open science, open to the world: a vision for Europe.* Publications Office. Retrieved from: https://data.europa.eu/doi/10.2777/061652

European Commission Directorate-General for Research and Innovation, O'Carroll, C., Rentier, B., Cabello Valdes, C., Esposito, F., Kaunismaa, ... Wouters, P. (2017). *Evaluation of research careers fully acknowledging Open Science practices: rewards, incentives and/or recognition for researchers practicing Open Science.* Publications Office. Retrieved from: https://data.europa.eu/doi/10.2777/75255

Finnish National Board on Research Integrity (TENK). (2024). *Template for researcher's curriculum vitae.* Retrieved from: https://tenk.fi/en/advice-and-materials/template-researchers-curriculum-vitae

Fitsilis, P., Damasiotis, V., Dervenis, C., Kyriatzis, V., & Tsoutsa, P. (In press). Effective Data Stewardship in Higher Education: Skills, Competences, and the Emerging Role of Open Data Stewards. *International Journal of Changes in Education.* doi:10.47852/bonviewIJCE52025166

Jetten, M., Grootveld, M., Mordant, A., Jansen, M., Bloemers, M., Miedema, M., & Van Gelder, C. W. G. (2021). *Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship* (1.1). Zenodo. doi.org/10.5281/zenodo.4623713

Mons, B. (2018). *Data Stewardship for Open Science: Implementing FAIR Principles*. Boca Raton, FL: CRC Press.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 71, Article 372. doi:10.1136/bmj.n71

Pontika, N., & Knoth, P. (2015). *Open Science Taxonomy*. FOSTER. doi:10.6084/m9.figshare.1508606.v3

Torres-Salinas, D., Orduña-Malea, E., Delgado-Vázquez, Á., Gorraiz, J., & Arroyo-Machado, W. (2024). Foundations of Narrative Bibliometrics. *Journal of Informetrics* 18(3), Article 101546. doi:10.1016/j.joi.2024.101546

Verheul, I., Imming, M., Ringerma, J., Mordant, A., van der Ploeg, J.-L., & Pronk, M. (2019). *Data Stewardship on the map: A study of tasks and roles in Dutch research institutes.* Zenodo. doi:10.5281/zenodo.2669150

Wendelborn, C., Anger, M., & Schickhardt, C. (2023). What is data stewardship? Towards a comprehensive understanding. *Journal of Biomedical Informatics* 140, Article 104337. doi:10.1016/j.jbi.2023.104337

Whyte, A., De Vries, J., Thorat, R., Kuehn, E., Sipos, G., Cavalli, ... Ashley, K. (2018). *D7.3: Skills and Capability Framework* (1.0). Zenodo. doi:10.5281/zenodo.5094800

Appendix: Selected Narratives from the Workshop

Example Narratives on Research Outputs

"Our 2021 article on scientific journals research data policies is disseminated Open Access by the publisher ([Insert DOI]) and the study's research data are available from Zenodo ([Insert DOI]). As of 7.8.2024, these outputs have been accessed over [Insert number] times and the article has been cited 24 times (91st percentile) as counted in the Scopus database. With the Altmeric Attention Score of 34, this article is also within the top 5% of all works scored by Altmetric." (Publications, research data specialist, social and behavioural sciences)

"The software for our paper [Insert title] is released on GitHub, it has [Insert number] stars and over the last year, [Insert number] unique contributors outside of our team and [Insert number] issues opened. It has also been released on PyPI and Conda, with a total of [Insert number] downloads since release [Insert number] months ago. According to libraries.io, the software is a dependency of [Insert number] other software packages. Currently, these metrics are visible but collecting the metrics is a fully manual process. Note that most projects are not heavily used, so these measure popularity more than quality." (Open source, research software engineer, computational research and data science)

Example Narratives on Research Process

"In my position at [Insert supervisor]'s group at [Insert university and department], I had the responsibility of ordering the chemical supplies and keeping updated the group's chemicals database. Additionally, I was responsible for providing training on the use of the thermal camera model [Insert model number], including giving an introduction to thermal imaging and skills to use software [Insert model number]. From 2014 to 2017, I was responsible for the organization of the Lab [Insert lab number], where I ensured an appropriate utilization of the space. This includes maintaining a clean working environment and establishing protocols for safety of both users and equipment, while taking care of the functioning condition of tools and ensuring apparatus were stored properly when not in use." (Risk management, researcher as a part-time data steward, materials science)

"As a [data steward] I have advised projects in my school on managing the risks related to their research data. This has included advising on Data Protection Impact Assessments and Privacy Notices, as well as helping plan the lifecycle of confidential data. This has required good understanding of the legal and ethical framework that governs how confidential research data is handled as well as ability to seek out appropriate solutions for risk management, e.g. safe storage. Much of this risk management has also been geared towards securing appropriate permissions that allow the data to be opened." (Risk management, research data specialist, social and behavioural sciences)

Example Narratives on Services and Leadership

"10% of my work time is allocated to service work in the Research Data Management team in improving research data management in [Insert university], particularly at the [Insert department]. My roles and contributions so far include (but not limited to):

- Contribute in formulating [Insert university] Open Science Policy, particularly on the aspect related to research software and software-related results.
- Contribute in Research Data Alliance recommendation [Insert name of the recommendation] [Insert DOI]." (Leadership, narrative shortened, researcher as a part-time data steward, electrical engineering)

Example Narratives on Research Impact

"Patient and public involvement is a mandatory requirement in medical research. If we extend this into a more broad perspective of designing studies with individuals, the involvement of the study population before the actual study is important to ensure alignment between the researchers' needs and the societal expectations towards researchers. Add here for example the designing of past studies with musicians, or with theatres - so less "medical" for this case, but more related to designing studies with the participants. This also touches on aspects of citizen science. It is difficult to create metrics for measuring involvement of patients and subjects. Metrics could be of media interest after a study is published e.g. altmetric examples or news coverage." (Research impact, researcher as a part-time data steward, medical ICT)

Example Narratives on Teaching

"I have been involved in the development of open learning materials on themes such as data protection, best practices for opening your code, reusing your data, artificial integrity especially in relation to research integrity. Open science drives my teaching not only because I try to make all learning materials openly available under CC-BY license, but also because all materials teach about topics related to transparency: transparency of the procedures used for collecting and analysing the data, academic transparency in the research process, and transparency towards the

study participants and the scientific community. By teaching open science practices we can ensure that transparency is addressed in all the research processes, for more ethical and sustainable research outcomes." (Teaching, researcher as a part-time data steward, medical ICT)

"As a [data steward] I have actively promoted and implemented best practices in research data management and open science. I provided hands-on support to research groups within my department, guiding them in adopting transparent research methods and effective data management strategies. Under my initiative, my research group has committed to publishing open datasets alongside our freely accessible articles, reinforcing our dedication to open science." (Mentoring, researcher as a part-time data steward, electrical engineering)

Example Narratives on Repository and Infrastructure Work

"I created and maintained [Insert project name] research software repository in [Insert university] Git version control, which has supported collection and archival of at least 8 research results done at [Insert project name] since 2018." (Researcher as a part-time data steward, electrical engineering)

"As the Application Owner, by developing and maintaining infrastructures and tools that support data management and data sharing. Examples include: [Insert university] Materials Digitalization Platform, [Insert university] Electronic Laboratory Notebook. Also providing training and especially technical support to users in using these tools and gathering feature development requests." (IT research data solution owner)

"By providing support to researchers in using cloud infrastructure and applications which can be used for data storage and computing. IT Services provides secure and ready to use cloud providers for storing and sharing research data including Microsoft Azure, Google Cloud Platform, AWS." (IT research data solution owner)

Example Narratives on Data Management Plans

"Solve research data management related cases in the department and also [Insert university] wide community, ranging from reviewing and commenting on the Data Management Plan ([Insert funder names] issue research work documentation and licensing, and provide practical advice on various topics, for example on use of version control." (Researcher as a part-time data steward, electrical engineering)

"Formulate [Insert project name] data management plan document to provide principle guidelines for researchers on steps and practicalities in their research data lifecycle, from collection until archiving phase, taking into account [Insert University] Open Science Policy and well known data management practices." (Researcher as a part-time data steward, electrical engineering)