Check for updates

RESEARCH ARTICLE

REVISED Data stewardship in Italy: Findings from the first national survey

[version 2; peer review: 2 approved, 3 approved with reservations, 1 not approved]
Previously titled: "Mapping Data stewardship in Italy: Findings from the first national survey"

Giulia Caldoni[1], Vittorio Iacovella[2], Emma Lazzeri[3], Liise Lehtsalu[4], Francesca Marchegiani[5], Mauro Paschetta[6], Valentina Pasquale[7], Giulia Pedonese[8], Andrea Tarallo[9]

[1]Research Division (ARIC), Alma Mater Studiorum - Università di Bologna, Bologna, Emilia-Romagna, 40126, Italy
[2]Universita degli Studi di Trento Centro Interdipartimentale Mente/Cervello, Rovereto, Trentino-Alto Adige/South Tyrol, 38068, Italy
[3]CNR - Ufficio Agenda Digitale e Processi, Rome, Lazio, 00185, Italy
[4]RDA Professionalising Data Stewardship Interest Group and RDA Data Steward Career Tracks Working Group, Brunico, Trentino Alto-Adige, 39031, Italy
[5]INFN - Laboratori Nazionali del Gran Sasso, Assergi, Abruzzo, 67100, Italy
[6]Politecnico di Torino, Turin, Piedmont, 10129, Italy
[7]Istituto Italiano di Tecnologia, Genoa, Liguria, 16163, Italy
[8]CNR - Istituto di Linguistica Computazionale "Antonio Zampolli", Pisa, Tuscany, 56124, Italy
[9]CNR - Istituto di Ricerca sugli Ecosistemi Terrestri, Lecce, Puglia, 73100, Italy

v2 **First published:** 13 Feb 2026, **6**:45
https://doi.org/10.12688/openreseurope.22183.1
Latest published: 31 Jul 2026, **6**:45
https://doi.org/10.12688/openreseurope.22183.2

Abstract

Research data management (RDM) plays a key role in ensuring research reproducibility and transparency by supporting data sharing in line with Open Science principles. Across the international landscape, professional figures supporting data management practices, such as data stewards (DSs), are emerging through a variety of career paths. The Italian Computing and Data Infrastructure (ICDI), together with the Horizon Europe Skills4EOSC project, conducted a national survey between April and December 2023 to map the presence, roles, and activities of data stewardship services across universities and research-performing organisations. The aim was to provide an initial overview of the Italian landscape, which is known to be fragmented and lacking of clear professional recognition for research support professionals. The survey, administered in Italian through the EUSurvey platform, consisted of 13 closed-ended questions addressing the professional background of DSs, their role and functions, and the structure of RDM support services. Data cleaning was performed on the 77 valid entries collected (out of 88 responses) before conducting the descriptive analyses. Results

Open Peer Review

Approval Status ✓ ? ✓ ? ? ✗

	1	2	3	4	5	6
version 2 (revision) 31 Jul 2026		? view				
version 1 13 Feb 2026	✓ view	? view	✓ view	? view	? view	✗ view

1. **Aoife Coffey**, University College Cork National University of Ireland University College Cork Boole Library (Ringgold ID: 494323), Cork, Ireland

Pádraig Mac Aodhgáin, University College Cork National University of Ireland University College Cork Boole Library (Ringgold ID: 494323), Cork, Ireland

indicate that most institutions employ staff with DS functions, although these professionals are seldom formally recognised as DSs. The variety of job titles highlighted by the survey reflects the absence of a standardised professional profile. Among the RDM support tasks performed by DSs, the most common include FAIR data management support, data management plan writing, researcher training, and policy consultancy. These findings provided a basis for the establishment of the Italian Data Steward Community, aiming to overcome the fragmentation through peer-to-peer knowledge sharing and support, and to advance the definition of a professionalisation pathway for DSs in Italy. The findings of this survey can serve as a case study to help DSs, research data professionals, and policymakers in other countries to map and understand their local landscape.

Plain Language summary

Good research depends on data that are well organised, well documented, and easy to share. Managing research data properly helps make research more transparent, trustworthy, and reusable, in line with Open Science principles. To support researchers with data management tasks, new professional roles are emerging. One of these roles is often called *data steward*, although this job title is not used consistently.

This study explores how data management support is organised in Italian universities and research institutions. Between April and December 2023, the Italian Computing and Data Infrastructure, supported by the Horizon Europe Skills4EOSC project, carried out a national survey to understand who provides data management support, what they do, and how they are recognised within their institutions.

The survey included 13 multiple-choice questions and was completed by 88 respondents, of whom 77 were suitable for analysis. The results show that many institutions already have staff members who support researchers with tasks such as writing data management plans, applying FAIR data principles, training researchers, and advising on data policies. However, these professionals often have different job titles and are rarely formally recognised as data stewards.

The study highlights a fragmented landscape and a lack of a clear professional profile in Italy. Based on these findings, the authors propose the creation of an Italian Data Steward Community to encourage knowledge sharing, reduce fragmentation, and support the development of a clear professional pathway. The results may also help other countries better understand and map their own data stewardship landscape.

Keywords

survey, data stewardship, Italy, data steward, professional profile, mapping

2. **Antti M. Rousi**, Aalto University, Espoo, Finland
3. **Louise Saul**, University of Southampton, Southampton, UK
4. **Angeliki Adamaki**, SciLifeLab Data Centre, Uppsala University (Ringgold ID: 463758), Uppsala, Sweden
5. **Jan Slifka**, Czech Technical University in Prague, Prague, Czech Republic
6. **Samsul Farid Samsuddin**, Universiti Malaya, Kuala Lumpur, Malaysia

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Horizon Europe gateway.

Corresponding authors: Giulia Caldoni (giulia.caldoni2@unibo.it), Francesca Marchegiani (francesca.marchegiani@lngs.infn.it)

Author roles: Caldoni G: Conceptualization, Data Curation, Formal Analysis, Writing – Original Draft Preparation; **Iacovella V**: Data Curation, Formal Analysis, Writing – Original Draft Preparation; **Lazzeri E**: Conceptualization; **Lehtsalu L**: Writing – Original Draft Preparation; **Marchegiani F**: Data Curation, Formal Analysis, Writing – Original Draft Preparation; **Paschetta M**: Writing – Original Draft Preparation; **Pasquale V**: Conceptualization, Data Curation, Formal Analysis, Writing – Original Draft Preparation; **Pedonese G**: Writing – Original Draft Preparation; **Tarallo A**: Formal Analysis, Writing – Original Draft Preparation

Competing interests: No competing interests were disclosed.

Grant information: This work has been partially funded by the European Union's Horizon Europe framework program under grant agreement No 101058527 (Skills4EOSC - Skills for the European Open Science Commons: creating a training ecosystem for Open and FAIR science). Andrea Tarallo has been funded by EU - Next Generation EU Mission 4, Component 2 - CUP B53C22002150006 - Project IR0000032 – ITINERIS - Italian Integrated Environmental Research Infrastructures System. Francesca Marchegiani has been funded by EU - Next Generation EU Mission 4, Component 2 - CUP I19D22000090007 - Project ID: IR0000024 "LNGS - Future Upgrade To Unveil Rare Events"
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Caldoni G *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Caldoni G, Iacovella V, Lazzeri E *et al.* **Data stewardship in Italy: Findings from the first national survey [version 2; peer review: 2 approved, 3 approved with reservations, 1 not approved]** Open Research Europe 2026, **6**:45 https://doi.org/10.12688/openreseurope.22183.2

First published: 13 Feb 2026, **6**:45 https://doi.org/10.12688/openreseurope.22183.1

REVISED **Amendments from Version 1**

In response to the reviewers' comments, the manuscript has been substantially revised to improve clarity, methodological transparency, consistency, and alignment between the results and conclusions. The *Introduction* was reorganised and expanded to provide clearer definitions of key concepts, harmonise terminology related to data stewardship and data stewardship services, and explicitly state the study aims and research questions. The literature review was strengthened through the inclusion of additional references and broader engagement with the international literature. The *Materials and Methods* section was revised to better explain the survey design process, the rationale for reopening the survey, the construction of the list of data stewardship activities, and the dissemination strategy. Throughout the *Analysis of Survey Results* section, denominators and analytical subsets are now clearly specified, inconsistencies and typographical errors have been corrected, and findings have been reorganised around the research questions to improve readability. Several figures were revised or replaced with tables, city labels were added to the geographic visualisation in Figure 1, and data presentation was simplified by reporting absolute numbers alongside percentages. The *Discussion* section was extensively revised to ensure that interpretations remain proportional to the descriptive nature of the data, and to acknowledge limitations more explicitly. Statements that could be perceived as overinterpretations were removed or rephrased using more cautious language, terminology was standardised throughout the text, and conclusions were streamlined to focus on the main findings. Finally, references, data citations, acronyms, figure captions, and language were carefully reviewed and corrected.

Any further responses from the reviewers can be found at the end of the article

Introduction

Data stewards (DSs) are a cornerstone of the research data ecosystem. Born at the intersection of data management, research support, and policy development (policies regarding research data management, Open Science, open access, personal data, knowledge management, etc.), this profession has developed (and continues to develop) in response to the growing complexity of managing, sharing, and preserving digital research outputs. As research institutions have embraced the principles of Open Science and FAIR (Findable, Accessible, Interoperable, and Reusable) (Wilkinson *et al.*, 2016), the need for dedicated professionals who can bridge the gap between researchers, IT specialists, and policy-makers has become evident. In this article, we consider DSs as key actors in ensuring that research data are not only technically well-managed but also ethically, legally, and contextually sound, and that research-performing institutions have in place policies and workflows that enable FAIR research data management (RDM); DSs can be in advisory roles, or embedded in research groups with more hands-on roles (The Turing Way Community, 2025). The specific role(s) and function(s) of DSs are not unequivocally defined, and often depend on organisational and national context. The same applies to data stewardship services, i.e. which functions are performed by DSs in a specific context utilising the resources made locally available (Ayres *et al.*, 2022). For this reason, throughout this article, the terms "DS", "DS functions" and "data stewardship" are used in a broad sense to refer to all activities related to data stewardship, as well as to all professionals whose roles can be fully or partially associated with this function, acknowledging that these terms may encompass a variety of roles, activities, and organisational interpretations.

This study aims to investigate how data stewardship services are organised in Italian Universities and Research Performing Organisations (RPOs), by applying descriptive statistics to the responses collected from the national DS community via a survey (see Materials and Methods).

The analysis addresses the following research questions:

- RQ1: To what extent are DSs formally integrated in research performing organisations in Italy?
- RQ2: What are the typical training pathways of Italian DSs?
- RQ3: What activities do Italian DSs perform in their organisations?
- RQ4: How are the RDM support services organised in different research performing organisations across Italy?

The article is structured as follows: the introduction provides a brief overview of the emergence of DSs as professional figures in Europe and in Italy; the section 'Materials and methods' outlines the methodology; the section 'Analysis of Survey Results' presents and discusses the findings; and the section 'Discussion' offers final reflections on the findings.

The emergence of research data stewardship in Europe and beyond

DSs as professionals have emerged only in the last 10 years, at least in research contexts (Wildgaard *et al.*, 2020). Following the publication of the FAIR principles in 2016 (Wilkinson *et al.*, 2016), the *Turning FAIR into reality* report by

the European Commission's Expert Group on FAIR Data recognised data stewardship as a key service for enabling FAIR data. The report also highlighted the need for new job profiles in data science and data stewardship to support researchers throughout the research data lifecycle (European Commission. Directorate General for Research and Innovation, 2018). Around the same time, universities in the Netherlands developed their first dedicated data stewardship programmes, including the 3-year Data Stewardship Project in TU Delft that introduced faculty-level DSs at the university (Turkyilmaz-van der Velden, 2018). Following the report dissemination, several European countries (e.g. Denmark, Austria, the Netherlands, Germany, and Poland) carried out national-level projects that surveyed local data stewardship landscapes, defined DS roles and service models, identified skills and competencies, and outlined training programmes (Gruber *et al.*, 2021; Jetten *et al.*, 2021; Seidlmayer *et al.*, 2023; Wałek, 2023; Wildgaard *et al.*, 2020). More recently, a national-level survey has also been completed in Portugal, following the example and adapting the national-level survey in Italy discussed in this article (Boavida *et al.*, 2025). These national-level projects helped to scope data stewardship in Europe and highlighted similarities and differences in data stewardship service models as well as in DS roles and functions across national contexts.

National-level work in the Netherlands and in Denmark identified three key functional areas for data stewardship: policy, research, and infrastructure, and defined DS roles in these areas (Schmitz, 2020). The German study identified five DS profiles: (1) generalist, (2) advisory RDM, (3) discipline-specific, (4) coordinating, and (5) infrastructure-focused DS (Seidlmayer *et al.*, 2023). Although these national-level studies evidence an overlap between DS roles and functions in different national contexts, they also highlight that DS role(s) and function(s) depend on the local context in which data stewardship services are initiated. The latter is confirmed by a recent comparative case study of data stewardship services in universities (Basalti *et al.*, 2024; Landrgottová, 2026; Rousi, Boehm and Wang, 2024). Moreover, in some (national) contexts the functions of a DS overlap with, or are subsumed by, data librarian or data curator roles (Ashiq and Warraich, 2023; Federer, 2018; Tammaro *et al.*, 2019). The Austrian study directly links DS roles to their organisational context by introducing DS profiles and related skills and competences through the definition of three data stewardship service models (Gruber *et al.*, 2021). National-level studies of data stewardship – like the Italian survey presented in this article – contribute to a better understanding of DS roles and how they are carried out in specific organisational contexts, thus supporting a more nuanced understanding of data stewardship services and their development.

Efforts to professionalise research data stewardship

A better understanding of DSs in various organisational and national contexts also allows for the convergence of DS roles across national, European and international contexts. Several efforts have aimed at aligning DS roles and data stewardship service models. Within the Research Data Alliance (RDA), the Professionalising Data Stewardship Interest Group[1] (endorsed in 2020) has completed two global surveys. The first one in 2021 on data stewardship service models (Ayres *et al.*, 2022) found that most respondents considered their service provision still in the early stages of maturity. A second one, administered in 2022 and focusing on DS career tracks (Newbold *et al.*, 2024) highlighted that professionals who self-identify as DSs have many different job titles – only 33% of the European respondents in DS roles at the time of the survey (n = 205) actually had the job title "data steward" – and underscored the lack of clearly defined career paths for DSs. An OpenAIRE survey completed in 2023 (Gurdal and Yavuz, 2023) found that 67% of the responding institutions did not have formal job descriptions for their RDM roles and the professionals in those roles had a variety of job titles (e.g. 'data steward', 'data manager', 'project employee'). The work by RDA and OpenAIRE highlights that lack of agreed-upon definitions of DS limits the recognition of DS roles (Kalová *et al.*, 2024). The Horizon Europe Skills4EOSC project[2] built on the DS roles, skills and competences defined in the various national contexts to deliver Minimum Viable Skills profiles for "coordinating data steward" and "embedded data steward" that are intended to have European-level applicability (Hansen *et al.*, 2025). National-level studies – like the survey results presented in this article – continue to provide valuable input to further define and recognise DS roles at national, European and international level and contribute to the continued professionalisation of data stewardship.

Data stewardship in Italy

In Italy, similarly to other European countries, the DS role is yet to be formalised. Following the increasing importance of professional RDM activities to support the research lifecycle, several Italian universities and research institutions have progressively introduced data stewardship roles, training opportunities, and governance models. In 2019, a group of research and e-infrastructures in Italy signed the "Italian Computing and Data Infrastructure" (ICDI) *memorandum of understanding*. ICDI is a forum created to promote synergies at the national level and to optimise the Italian participation

[1]RDA Professionalising Data Stewardship Interest Group, https://www.rd-alliance.org/groups/professionalising-data-stewardship-ig/activity/, URL checked on 2025/11/28

[2]Skills for the European Open Science Commons: creating a training ecosystem for Open and FAIR science, https://www.skills4eosc.eu/, URL checked on 2025/12/04

in European and global challenges in the field of Open Science and FAIR data (*Italian Computing and Data Infrastructure*, no date). By including several Italian research institutions and universities, ICDI progressively became the forum where Open Science and RDM experts could exchange ideas, start collaborations, and propose solutions to policy-makers with a bottom-up approach. The monthly online "Open Science Cafè" webinars, on major Open Science topics and updates, exemplify ICDI's efforts in community engagement and building in the Italian landscape[3].

One of the leading organisations in ICDI is Consortium GARR[4], which also coordinated the aforementioned Horizon Europe Skills4EOSC project. Among its activities, this project defined DS roles and training, and supported the development of professional networks of DSs across Europe. The survey that we discuss in this article emerged from the collaboration between ICDI and Consortium GARR in Skills4EOSC and aimed at understanding better DS roles and services in Italian universities and research-performing organisations. The survey also served as the first step towards developing a national community of practice for DSs and research data management roles. This article presents the survey results, shedding light on DS roles and data stewardship service models in the Italian context. This is the first attempt to study data stewardship in Italy.

Materials and Methods

Survey development and distribution

The "Survey on Data Stewards in Italy: a mapping of professionals supporting research data management" was developed as a bottom-up initiative by Consortium GARR, coordinator of the Skills4EOSC project, in collaboration with the University of Bologna (UNIBO) and the Italian Institute of Technology (IIT), under the umbrella of the Italian Computing and Data Infrastructure (ICDI).

The survey was developed as a completely anonymous web-based set of questions using EUSurvey, an open-source platform by the European Commission. It was distributed in Italian to facilitate completion and improve response rates, but an English version was also prepared.

To maximise participation, we initially targeted existing communities interested in RDM topics. We shared the survey via two national mailing lists focused on Open Science ("oa-italia@lists.icdi.it" and "icdi@lists.garr.it"), and expanded the audience through word-of-mouth, institutional social media channels (e.g., Skills4EOSC LinkedIn page[5]), and an article in the GARR Magazine (Colcelli, 2023). We acknowledge the limitations of this approach, as it may have introduced sampling bias (see Discussion section).

The survey was officially launched in April–May 2023 (from 13/04/2023 to 31/05/2023) and reopened in November–December 2023 (from 07/11/2023 to 01/12/2023), aiming to increase participation and expand the geographical coverage. We acknowledge that collecting data during two distinct time periods may represent a limitation of the study, as temporal differences could have influenced respondents' perceptions and answers.

Survey and question structure

The introduction of the survey provided a general definition of DS to ensure a shared understanding among respondents, as survey completion was not limited to professionals with the formal job title of DS.

The first section of the survey collected information about the respondents' organisations and roles. Respondents were also asked to report whether they were completing the survey as individuals in a personal capacity or as representatives of their organisation. This distinction was intended to differentiate responses provided under an official institutional mandate from those based on the respondent's own knowledge and experience. Regardless of the selected category, respondents reported information about their organisation.

Only those respondents who reported DSs in their organisations could proceed to the second part of the survey; for all others, the survey ended after collecting profile information.

The second section of the survey included 13 closed-ended questions that explored the presence of DSs in Italian research performing organisations, their training pathways, their institutional activities, and the organisation of data stewardship services, with the aim of addressing the research questions presented in the introduction.

[3]https://www.icdi.it/it/attivita/competence-centre/open-science-cafe
[4]https://www.garr.it/it
[5]https://www.linkedin.com/company/skills4eosc/, URL checked on 2025/12/11

To structure data collection and facilitate analysis, the questionnaire consisted mostly of closed-ended questions. We included single-choice questions where respondents were required to select one mutually exclusive option, as well as multiple-choice questions that allowed the selection of more than one response. Open-ended questions were included to capture information that could not be reliably categorised in advance.

We designed the survey to collect only anonymous data from respondents.

The questions were developed through an iterative design process informed primarily by the authors' direct professional experience, which allowed to capture the context-specific dynamics and practical considerations. This firsthand expertise was complemented by insights drawn from international reports, particularly those developed in the Netherlands and Denmark. As an example, this approach was adopted to compile the list of data stewardship activities presented in the question "What kind of support do the "data stewards" or equivalent figures provide within your organisation?" We archived survey templates, in English and Italian, as well as the survey output dataset, in Zenodo (Caldoni, Iacovella, *et al.*, 2025).

Data cleaning and visualisation

A total of 84 anonymous responses were collected and exported as a.csv file, on which data cleaning was performed.

OpenRefine[6] () was used to split multiple-choice responses, correct inconsistencies and cluster respondent roles into three macro-categories: "researcher", "technologist" and "technical-administrative staff". These three macro-categories were chosen as they represent distinct and formalised functions in Italian universities and research performing organisations. Namely, "researchers" are directly responsible for producing scientific knowledge and conducting original research, while "technologists" provide advanced technical and specialist support to research. This second role typically involves managing laboratories, digital infrastructures, data platforms, scientific instruments, or complex technical workflows. Finally, "technical-administrative staff" ensures the day-to-day functioning of the organisation by handling administrative, financial, and organisational tasks.

Manual data cleaning operations consisted of:

- adding ROR identifiers when not provided by respondents,
- removing responses not from Italy (used for functionality tests),
- adding a column for organisational acronyms,
- classifying organisations (universities vs. research entities),
- eliminating duplicate answers.

Lastly, two responses referring to the affiliated organisation of the respondents were manually inspected and excluded because the authors could unequivocally verify that these records were incorrect.

After cleaning, 77 valid responses remained. The final dataset was reviewed by all authors to ensure consistency and accuracy.

Given the limited sample size, the analysis was restricted to descriptive statistics, performed using Microsoft Excel[7].

R was used to produce a map of the geographical distribution of the respondents (Figure 1). The resulting R-plot was replicated, for visualisation purposes, into a multi-level image using the vector graphics editor Inkscape[8]. To improve readability, city-centred pie-charts reflecting geographical provenance were subdivided into 4 categories, according to the number of respondents. Cities with a single respondent were chosen as the unit size. Pie charts for cities with 2–5 respondents were visualised as 1.5 times the size of the unit; for cities with 6–14 respondents, the pie chart size was scaled at 2 times the size of the unit. The only pie chart corresponding to a city with more than 14 respondents was resized to 2.5 times the size of the unit.

[6]OpenRefine, v3.8.7, https://openrefine.org/, URL checked on 2026/07/21
[7]Microsoft® Excel® per Microsoft 365 MSO (Version 2606 Build 16.0.20131.20154, 64 bit)
[8]Inkscape 1.3 (0e150ed6c4, 2023-07-21), https://inkscape.org, URL checked on 2025-24-10

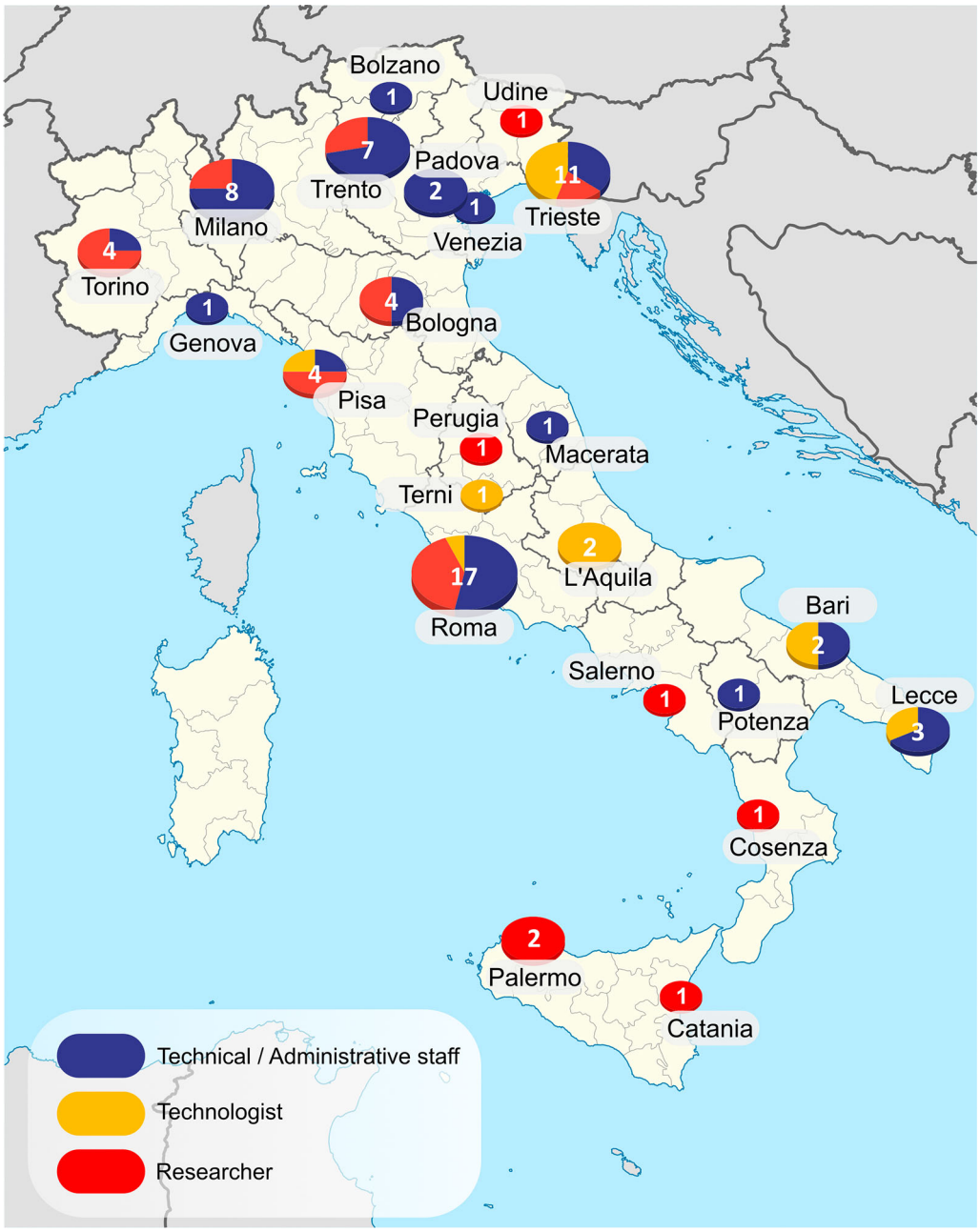

Figure 1. Map of institutional roles represented among respondents to the national survey.

The results of the analysis are presented in graphs, maps and tables to visualise the diversity and geographical coverage of the survey.

Analysis of Survey Results

Institutional roles and personal interests

First, we asked respondents to indicate their affiliation and specify their role within the organisation. This granted a better picture of the professionals involved in the study, which was useful in the qualitative analysis of the data.

We collected responses from 48 different institutions (based on unique ROR identifiers), reflecting broad institutional representation. As presented in Figure 1, these institutions are geographically distributed throughout Italy, although regional representation is not uniform.

Out of the total 77 respondents, 33 are affiliated with universities, while 44 are affiliated with public research performing organisations.

Regarding the professional role of the respondents, we considered the three main types of employees within the Italian public research performing organisations: Technical/Administrative staff, Technologist, and Researcher. To highlight possible geographical differences within the respondents' population, we collated responses from different research performing organisations in the same city. We mounted single-city pie charts on a map of Italian administrative subdivisions[9].

The majority of the respondents belong to the Technical/Administrative support staff (n = 38), while another significant portion of respondents identified themselves as a researcher (n = 27). Only a small number of technologists (n = 12) answered the survey, consistent with the uneven distribution of this professional framework across the national landscape (see the section "Data cleaning and visualisation" in the "Materials and Methods" chapter for the definition of these three different professional roles).

As participation was not restricted to individuals formally designated as DSs, the results capture perspectives from a broader community involved in data-related activities. While responses likely reflect participants' professional experiences, the survey did not explicitly ascertain the perspective from which answers were provided.

Then, we asked whether the respondents were participating in the survey in a personal capacity or whether they were officially representing their organisation. Most of the respondents participated in a personal capacity (57 out of 77, 74%), while 26% responded on behalf of their institution (20 out of 77).

RQ1: To what extent are DSs formally integrated in research performing organisations in Italy?

We started our analysis by examining the presence and the role of DSs or analogous figures in the organisations that participated in the survey. In this section, we focus our analysis solely on the responses provided on behalf of organisations (n = 20).

About 65% (13 out of 20) of the respondents state that there are professionals with DS functions (Figure 2).

Interestingly, only one out of 13 respondents reported that the DS role is formally recognised (Figure 3).

Table 1 shows that the job titles used to identify professional figures working in DS roles can vary considerably, even within the same organisation. Respondents were allowed to select as many job titles as necessary to describe the reality of

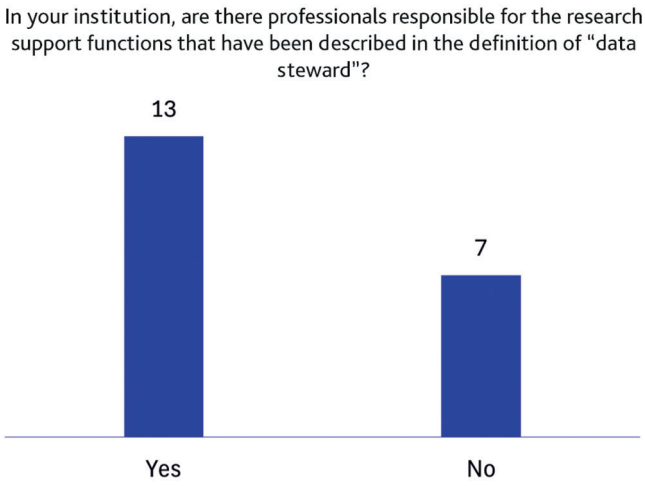

Figure 2. Presence of professional figures performing DS functions. (n = 20 respondents, answering on behalf of their institutions.)

[9]https://commons.wikimedia.org/wiki/File:Italy_provincial_location_map_2016.svg, URL checked on 2025/10/24

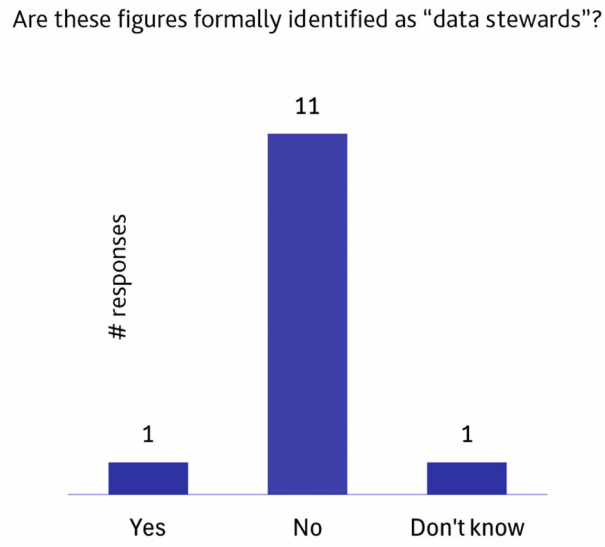

Figure 3. Presence of professional figures supporting research management formally identified as DSs. (n = 13 respondents, answering on behalf of their organisations and reporting that the organisation has staff dedicated to DS functions.)

Table 1. Breakdown of the formal designation of professional figures with DS functions within organisations.

Job title	Number of responses
Research support staff	2
Data manager	2
Research Data Management Specialist	5
Data scientist	2
Data expert	2
Information specialist	1
Other	3

n = 11 respondents, answering on behalf of their organisations and reporting that the organisation has staff dedicated to DS functions. Multiple (non-mutually exclusive) responses are allowed.

data stewardship in their organisation. Respondents were provided with seven possible, non-mutually exclusive response options, partially derived from the roles most represented in the international landscape (Ayres *et al.*, 2022; Jetten *et al.*, 2021) and had the chance to add more options in a free text box if necessary.

Further evidence of the lack of a clear professional framework for DSs is seen in the distribution of the roles for those performing data stewardship activities, including research roles (post-doc, research fellow, etc.), technical roles (e.g. IT support, data processing technician, etc.) and administrative roles (e.g. librarian, research support, project manager and others). As shown in Table 2, no significant differences were observed across these categories.

RQ2: What are the typical training pathways of Italian DSs?

Considering the lack of any officially recognised national training pathway leading to the title of "data steward" in Italy, we asked about the educational background of DSs. For this analysis, we included all respondents who reported the presence of DSs in their organisations (n = 38), without dividing responses given on behalf of the organisation or in a personal capacity. Respondents were allowed to select multiple response options, to allow us to capture the nuances of a single DS training path.

Figure 4 shows the prevalence of specific disciplinary backgrounds. Indeed, the largest group (18 responses) reported a master's degree or PhD in a relevant disciplinary field (e.g., biology, physics, social sciences). The second largest group

Table 2. Breakdown of the professional roles of DSs within organisations.

Roles	Number of responses
Technical staff	5
Administrative staff	6
Research staff	6
Don't know	1
Other	2

n = 13 respondents, answering on behalf of their organisations and reporting that the organisation has staff dedicated to DS functions. Multiple (non-mutually exclusive) responses are allowed.

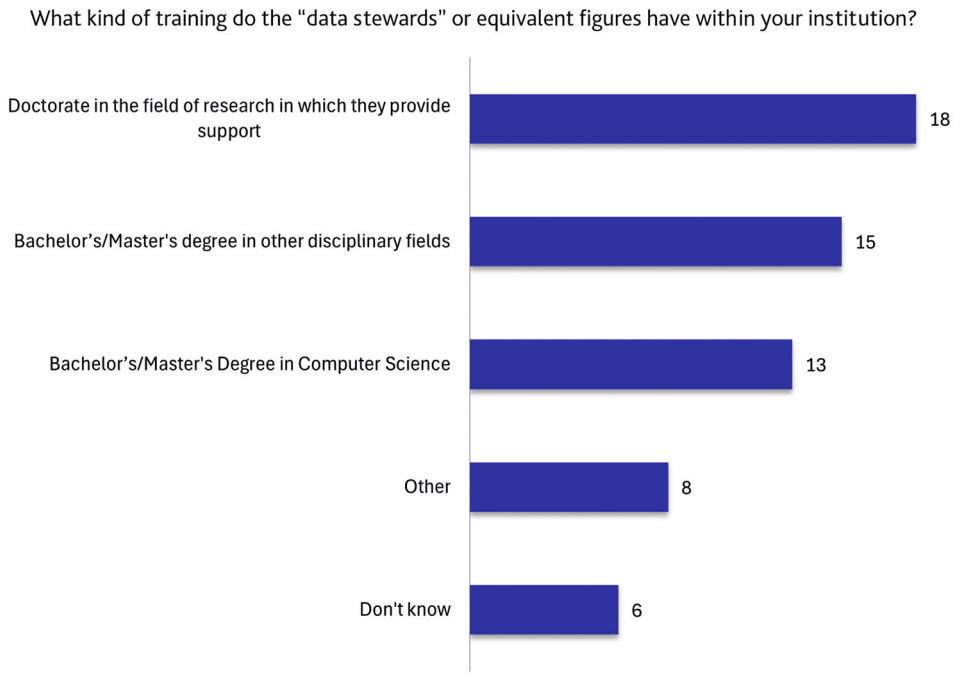

Figure 4. Breakdown of the types of training possessed by DSs before their recruitment within organisations. (n = 38 respondents, answering either on behalf of their organisation or personally, reporting that the organisation has staff dedicated to DS functions. Multiple (non–mutually exclusive) responses are allowed.)

(15 responses) indicated degrees in related or complementary fields (e.g., library science, archiving, information science, statistics). A total of 13 respondents reported degrees in Computing or Computer Science. Finally, 8 responses selected "Other" and 6 selected "I don't know".

RQ3: What activities do Italian DSs perform in their organisations?

After detailing the DSs' background, we investigated their activities, focusing mainly on support and data management (Ayres *et al.*, 2022; Basalti *et al.*, 2024; Jetten *et al.*, 2021). Namely, we asked what kind of activities the DSs perform in their organisation, providing seven possible response options as well as the option to indicate more activities. The list of the response options included:

- "support for research data management and the application of FAIR principles, identification of disciplinary repositories, storage and backup solutions, metadata, etc.;"
- "support for writing the data management plan; assessment of possible legal issues (e.g., privacy, IP and licences, etc.);"
- "drafting of institutional policies for research data management;"

- "advice on national, international and funding body policies and mandates for data management and compliance with any obligations;"
- "training of researchers on Open Science, FAIR data and research data management issues;"
- "support for the use of data management tools made available by their institution (e.g. institutional repository, institutional cloud, etc.);"
- "support for the management and maintenance of the institutional repository (e.g. care and validation of deposited datasets)."

We considered for our analysis here all respondents who indicated that their organisations had DS (n = 38), without restricting our analysis to responses provided on behalf of an institution only. Respondents could select more than one response option to capture the extent to which DS activities are characterised by specialisation or encompass multiple DS functions.

The results (Figure 5) show that DSs or equivalent figures are prevalently involved in RDM support, particularly in applying FAIR principles (34 responses) and assisting with DMPs (31 responses). A second cluster of activities included drafting institutional policies (24 responses), training on Open Science and RDM (25 responses), and support for institutional data management tools (26 responses). More specialised activities like providing advice on national, international and funding body policies and mandates for data management and compliance with any obligations (19 responses), as well as legal assessment (18 responses) and repository maintenance (18 responses) were less prevalent.

As shown in Figure 5, most respondents (95%) selected at least 1 support activity out of the 8 we proposed in the survey. Only 6 respondents (16%) selected 1–3 related activities. 30 out of 38 (more than 75% of the cohort) selected at least 4 activities. This multi-task related partition shows a substantial equality between the number of activities: the number of people selecting 4 to 8 activities is, for all the 5 categories (4,5,6,7,or 8 activities), always about 15% (respectively, 6,5,5,7, and 7 people). By focusing on the 11 (about one third of the cohort) respondents selecting 4–5 activities, there is a solid pattern (10 out of 11) of selecting both RDM and DMP. This often (7 people) occurs together with training duties. A

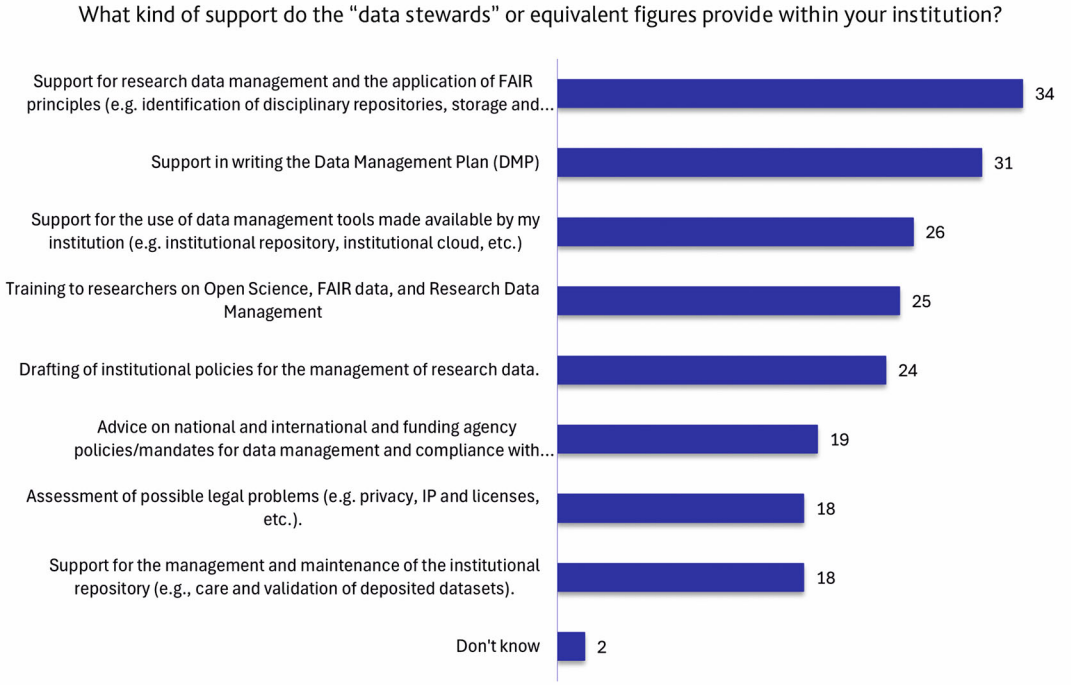

Figure 5. Breakdown of the support functions carried out by data stewards or analogous figures within organisations. (n = 38 respondents, answering either on behalf of their institution or personally, reporting that the organization has staff dedicated to DS functions. Multiple (non–mutually exclusive) responses are allowed.)

substantial proportion of respondents of the entire cohort (19 out of 38, 50%) selected 6 or more different activities. They covered most of the proposed activities, with slightly fewer mentions (13 out of 19) of the "Advice on local and international policies" and the "Institutional repository" activities.

After mapping the type of activities provided, we decided to investigate the target groups of these services. In particular, we asked whether these services were only available to researchers with project-related data management obligations. 29 (out of 38) respondents (76%) indicated that services are available to all research staff, regardless of project obligations.

RQ4: How are the RDM support services organised in different research performing organisations across Italy?

In the last RQ, we proceeded to investigate how DS activities are organised. We based the analysis on 38 out of 77 respondents indicating the presence of DSs in their organisations.

As shown in Table 3, most respondents (26 out of 38) reported that their organisation had two or more professionals performing roles equivalent to DSs. The presence of a single individual in this role was reported in seven organisations.

Figure 6 shows that in 42% of cases (16 out of 38 responses) DSs are embedded within research facilities, while 26% (10 out of 38) frame these professional figures within central administrative areas. Interestingly, 21% (8 out of 38) of respondents reported a hybrid situation, with DS roles established at both levels.

To further characterise the organisational placement of DSs, we combined responses indicating a central administrative location with those reporting a hybrid model, resulting in a total of 18 cases. The distribution of these cases across specific central organisational units is presented in Table 4.

Since this question allowed for multiple, non-mutually exclusive responses, three out of 18 respondents indicated more than one area of the central administration, in particular, research support, IT, and administrative divisions. This may reflect the interdisciplinary and cross-functional nature of the DS role, even within centralised structures.

Table 3. Count of DSs within organisations.

Number of DS or equivalent figures	Number of responses
Don't know	5
N > 10	2
2 ≤ N ≤ 10	24
N = 1	7

N = number of DSs; n = 38 respondents, answering either on behalf of their organisations or in a personal capacity, reporting that the organisation has staff dedicated to DS functions.

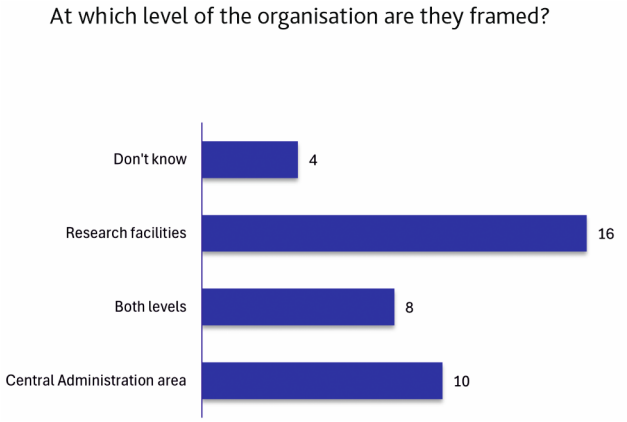

Figure 6. Level of the organisation where "data stewards" or equivalent figures are framed within organisations. (n = 38 respondents, answering either on behalf of their organisation or personally, reporting that the organisation has staff dedicated to DS functions.)

Table 4. Breakdown of the central administration areas where DSs are framed within organisations.

Central administration area: services or units involved	Number of responses
Research support	12
Research organisation	4
IT	4
Library	1
Don't know	2
Other	2

n = 18 respondents, answering either on behalf of their organisations or personally, reporting that the organisation has staff dedicated to DS functions and employed, either exclusively or partially, in a central administration area. Multiple (non–mutually exclusive) responses are allowed.

Discussion

The findings of this survey provide an overview of the emerging professional landscape of data stewardship in Italian universities and public research performing organisations.

We observed that only a minority of respondents participated in the survey on behalf of their organisation. Consistent with the European trends reporting a correlation between institutional maturity in RDM and the adoption of DS roles (Gruber *et al.*, 2021; Jetten *et al.*, 2021; Seidlmayer *et al.*, 2023; Wildgaard *et al.*, 2020) it is reasonable to infer that organisations with higher levels of awareness and engagement in RDM are more likely to be developing or implementing structured RDM support services and to have designated a representative for survey completion.

The results collected with RQ1 indicate that the DS role is increasingly recognised as essential in the research landscape, yet still characterised by heterogeneous role definitions and limited institutional recognition. We decided to consider only responses provided on behalf of the different organisations (n = 20) as our interest was focused on capturing the formal and institutional definition of the DS role.

We observed that approximately 65% (13 out of 20) of the respondents reported the presence of professionals performing DS functions (Figure 2). As shown in Figure 3, the DS role is formally recognised only in one case. On the other hand, we observed a variety of job titles associated with DS roles, even within the same organisation (Table 1). Taken together, these findings show that research performing organisations in Italy recognise the relevance of introducing DS functions (e.g. support to RDM and DMP drafting/updating, promote the adoption of FAIR principles, facilitate data sharing in compliance with ethical, legal, and security requirements, training, etc.). However, this awareness does not lead to the formal recognition or standardisation of DS role(s), a finding which aligns with the international landscape (Ayres *et al.*, 2022; Hansen *et al.*, 2025; Newbold *et al.*, 2024). The fact that DSs appear to be equally employed either as technical, administrative or research staff (Table 2) further builds on these observations, suggesting that DS functions are often assigned to previously existing roles, without providing adequate formalisation, visibility, or clear career pathways.

We also note that 35% of the respondents (7 out of 20, Figure 2) reported the absence of DSs in their organisation, suggesting uneven levels of awareness on this topic.

Hence, and in line with previous surveys in other national contexts (Jetten *et al.*, 2021; Seidlmayer *et al.*, 2023; Wildgaard *et al.*, 2020), our results show that the definition of the DS role depends heavily on the organisation, underlining the absence of a coordinated national framework.

With RQ2 and RQ3, we aimed to characterise the professional identity of DSs, specifically in terms of their educational background and the activities they perform. In the majority of cases, DSs hold advanced degrees in disciplinary fields related to the scientific discipline in which they provide data stewardship (Figure 4). Domain-specific expertise is a key asset to provide effective data stewardship services, as it enables DSs to understand disciplinary research practices and facilitates communication with researchers (Gruber *et al.*, 2021; Jetten *et al.*, 2021; Seidlmayer *et al.*, 2023; Wildgaard *et al.*, 2020). Interestingly, an educational background in computer science is frequently reported, thus suggesting a growing emphasis on the technical dimension of the DS role. Nonetheless, the results reflect a lack of consensus regarding the educational pathways necessary for DS role.

When looking at DS functions, according to the responses gathered in Figure 5, a DS is a professional with at least one or more cross-cutting skills in RDM (e.g. disciplinary, IT and technical, legal). Among the eight activities included in the survey, support for RDM and for DMP drafting are the most frequently reported. Interestingly, these core activities are often carried out alongside two or more additional activities, such as support for the use of RDM tools provided by the organisation, training and capacity-building efforts. These results highlight that a single DS should be able to perform multiple DS functions, which fall between those specific to the research domain and those typical of administrative functions.

Overall, these findings depict a complex landscape in which DSs are expected to acquire and develop a broad range of skills to perform effectively, while the corresponding training pathways remain unclear and insufficiently aligned, often relying on traditional learning models that do not fully address all relevant skills and competencies. This gap highlights the need for more structured, comprehensive, and role-specific training frameworks to better align professional preparation with the evolving demands of the field.

The final part of our survey aimed to portray how DS services are organised. Most respondents reported that their organisation employs between two and ten professionals performing the DS role (24 out of 38 respondents, Table 3). While the range is too broad to accurately characterise the typical number of DSs in research performing organisations, it suggests that the preferred strategy is to employ more than one person performing such functions. The second most common response indicates the presence of only one DS in the organisation, showing different approaches to the delivery of data stewardship services in different research performing organisations.

This diversity is further sustained when investigating at which level DSs are introduced in research performing organisations (Figure 6): according to our results, we observe that DSs are more frequently employed in research facilities (n = 16) than in central administration (n = 10), with 8 cases reporting their presence at both levels. This distribution is coherent with the description of two different DS profiles presented by the Skills4EOSC Minimum Viable Skills profiles (Hansen *et al.*, 2025). Specifically, the project identified the "Coordinator" DS as the one providing support across an organisation's research domains and units, while the "Embedded" DS operates close to a research team and to its domain-specific practices. The data collected with our survey seem to support the presence of these different profiles in Italy as well. Interestingly, around 20% of the respondents (8 out of 38) reported that DS are framed at both levels, suggesting that some organisations are still experimenting with hybrid service models that can benefit the general academic community while also providing specialised support.

Overall, this study provides initial insights into the evolution and development of the DS role in Italian research performing organisations. The findings suggest a growing recognition of DS functions and their gradual, though heterogeneous, integration into institutional practices and service frameworks, which appear to be diversified. The DS role is intrinsically interdisciplinary, combining research, technical and support competencies, yet the results highlight the need for clearer role definition, more structured training and stronger national coordination.

We acknowledge the limitations of this analysis. First, the analysis is based on data collected in 2023 and therefore may not fully reflect the current state of the Italian DS landscape. Second, the dissemination strategy reflected the bottom-up nature of the initiative and relied primarily on channels associated with the RDM communities. While this approach was effective in reaching professionals likely to be involved in data stewardship activities, it may also have introduced a sampling bias, potentially overrepresenting research performing organisations with established RDM services. Finally, although the survey allowed us to gather responses from across the country (Figure 1), we recognise that they represent only a limited fraction of Italian research performing organisations, with responses obtained from 17 out of the 104 universities and research centres operating in Italy[10]. This should be considered in light of the study's primary aim, which was to provide a snapshot of the existing practices and roles rather than to develop a comprehensive representation of the entire national context. A further consideration is that participation was open to professionals with diverse roles and was not restricted to individuals formally serving as data stewards. Consequently, the findings reflect perspectives from a broader data stewardship community. While respondents likely drew on their own professional experiences, the survey did not explicitly determine whether answers referred to personal practice or broader organizational observations.

Despite these limitations, our findings offer a starting point to support institutional decision-makers not only in defining and structuring data stewardship services within an increasing number of universities and research performing

[10]https://www.mur.gov.it/it/aree-tematiche/universita/le-universita, URL checked on 2025/12/02 https://www.mur.gov.it/it/aree-tematiche/ricerca/il-sistema-della-ricerca/enti-di-ricerca-pubblici, URL checked on 2025/12/02

organisations, but also in advancing the professionalisation of the DS role and the development of dedicated career paths. Beyond its immediate results, the survey represents a valuable output in itself, as its methodology can be readily adopted in other national contexts seeking to undertake a similar landscape analysis. This has been recently demonstrated in Portugal (Boavida *et al.*, 2025).

Moreover, the survey served as the initial step to kick off the Italian Data Stewards Community (Comunità Italiana Data Steward, CIDS) in 2023. The Community has since articulated three overarching objectives, formalised in its 2025 manifesto (Caldoni, Coppini, *et al.*, 2025), which are aligned with the main observations reported in this study: (i) to create opportunities for sharing and developing new skills among DSs; (ii) to support the adoption of best practices in RDM; and (iii) to increase awareness of the DS's role within the research ecosystem. As a bottom-up initiative, the Community now includes more than 150 members and promotes networking and knowledge exchange through regular meetings, publications, and collaborative activities. The survey thus played a foundational role in fostering the Community's self-recognition and professional consolidation within the Italian context.

Ethics and consent

The survey was voluntary and anonymous. Participants were informed that "the survey is anonymous and therefore none of the answers will be related to the participant's personal data." No personal identifiers or sensitive personal data were collected. The questionnaire focused on institutional practices concerning the presence and activities of data stewards and equivalent research-support professionals in Italian universities and research organisations. Participation in the survey implied informed consent. Given the anonymous and minimal-risk nature of the study and the absence of identifiable personal data, the study was considered exempt from formal ethical review under the applicable institutional guidelines.

Data and software availability

Underlying data

Zenodo: Data and code for Mapping Data Stewardship in Italy: Findings from the First National Survey (1.0). https://doi.org/10.5281/zenodo.17907703 (Caldoni, Iacovella, *et al.*, 2025).

The project contains the following underlying data:

Data_Figure1.csv (Portion of the whole dataset, integrated with geographical information related to the institutions of the respondents, ready to be used by "Make_Figure1.R script).

Data_Survey.csv (File containing the whole dataset acquired through the survey).

DS_Survey_Template_EN.pdf (English version of the text of the survey presented to the participants).

DS_Survey_Template_IT.pdf (File containing the text of the survey presented to the participants).

Make_Figure1.R (R code to produce Figure 1 on "Mapping Data Stewardship in Italy: Findings from the First National Survey" manuscript).

Readme.pdf (Readme file describing the content of the dataset).

Acknowledgements

The authors thank Chiara Basalti (Alma Mater Studiorum - Università di Bologna), Sara Di Giorgio (Consortium GARR) and Monica Forni (Alma Mater Studiorum - Università di Bologna) for their contribution in the initial conceptualisation of the survey.

References

Ashiq M, Warraich NF: **A systematized review on data librarianship literature: Current services, challenges, skills, and motivational factors.** *Journal of Librarianship and Information Science.* 2023; **55**(2): 414–433.
Publisher Full Text

Ayres B, *et al.*: *RDA Professionalising Data Stewardship - Current Models of Data.* Stewardship: Survey Report; 2022.
Publisher Full Text

Basalti C, *et al.*: **Recommendations for Data Stewardship Skills.** *Training and Curricula with Implementation Examples from European Countries and*

Universities. Zenodo. 2024.
Publisher Full Text

Boavida CP, *et al.*: **Inquérito nacional sobre o papel dos profissionais no apoio à gestão de dados de investigação.** *Zenodo.* 2025.
Publisher Full Text

Caldoni G, Iacovella V, *et al.*: **Data and code for Mapping Data Stewardship in Italy: Findings from the First National Survey.** *Zenodo.* 2025.
Publisher Full Text

Caldoni G, Coppini S, *et al.*: **Manifesto of the "Italian Data Steward Community".** *Zenodo.* 2025.
Publisher Full Text

Colcelli MF, Basalti C, Caldoni G, *et al.*: *Il ruolo dei data steward nella ricerca e la comunità italiana.* GARR News; 2023.
Reference Source

European Commission. Directorate General for Research and Innovation: *Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data.* LU: Publications Office; 2018.
Publisher Full Text

Federer L: **Defining data librarianship: a survey of competencies, skills, and training.** *Journal of the Medical Library Association.* 2018; **106**(3): 294–303.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Gruber A, Schranzhofer H, Knopper S, *et al.*: **Kompetenzen von Data Stewards an österreichischen Universitäten.** *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare.* 2021; **74**(1): 12–32.
Publisher Full Text

Gurdal G, Yavuz B: **OpenAIRE Survey: Who is Managing Research Data?** *OpenAIRE.* 24 February. 2023. (Accessed: 27 April 2026).
Reference Source

Hansen KK, *et al.*: **Data Steward: Minimum Viable Skills Profile.** 2025.
Publisher Full Text

Italian Computing and Data Infrastructure (no date). (Accessed: 11 December 2025).
Reference Source

Jetten M, *et al.*: *Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship.* Zenodo. 2021.
Publisher Full Text

Kalová T, *et al.*: **Data Stewardship Career Paths: Recommendations of the EOSC Task Force Data Stewardship Curricula and Career Paths.** 2024.
Publisher Full Text

Landrgottová A: *Comparative Analysis of the Data Stewardship role at Charles University and the Member Universities within the 4EU+.* MA Thesis. Charles University; 2026.
Reference Source

Newbold E, *et al.*: *RDA Professionalising Data Stewardship IG- What does a career track for data stewards look like?* Research Data Alliance; 2024.
Publisher Full Text

Rousi AM, Boehm RI, Wang Y: **Data stewardship: case studies from North American, Dutch and Finnish universities.** *Journal of Documentation.* 2024; **80**(7): 306–324.
Publisher Full Text

Schmitz F: **The roles of data stewards in the data stewardship landscape identified in Denmark and the Netherlands.** 2020.
Publisher Full Text

Seidlmayer E, *et al.*: *Forschung unterstützen - Empfehlungen für Data Stewardship an akademischen Forschungsinstitutionen: Ergebnisse des Projektes DataStew.* 2023.
Reference Source

Tammaro AM, Matusiak KK, Sposito FA, *et al.*: **Data Curator's Roles and Responsibilities: An International Perspective.** *Libri.* 2019; **69**(2): 89–104.
Publisher Full Text

The Turing Way Community: **The Turing Way: A handbook for reproducible, ethical and collaborative research.** *Zenodo.* 2025.
Publisher Full Text

Turkyilmaz-van der Velden Y: **Data Stewardship at Delft University of Technology.** *International Data Week/SciDataCon.* 2018.
Publisher Full Text

Wałek A: **Data Stewardship in Poland – Curricula and Career Paths – The Very Beginning of the New Profession.** *Przegląd Biblioteczny.* 2023; **91**: 63–81.
Publisher Full Text

Wildgaard L , *et al.*: **National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management.** 2020.
Publisher Full Text

Wilkinson MD, Dumontier M, Aalbersberg IJJ, *et al.*: **The FAIR Guiding Principles for scientific data management and stewardship.** *Scientific Data.* 2016; **3**(1): 160018.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✔ ? ✔ ? ? ✘

Version 2

Reviewer Report 07 September 2026

https://doi.org/10.21956/openreseurope.26746.r78363

© **2026 Rousi A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Antti M. Rousi
Aalto University, Espoo, Espoo, Finland

Dear Authors,

Thank you for the revised version. I genuinely appreciate the amount of work you put into the revision. Thank you also for the detailed replies to reviewers. The well-crafted replies signal that you value our work. Despite major improvements, I think the paper still needs to be finalized for better clarity before it can be indexed. Please see my detailed suggestions below. I hope the authors find my comments helpful.

General remarks

Although the authors have put much effort into revising the article, the current revision still displays carelessness in the form of, e.g., discrepancies between reply to reviewers and article execution and in my understanding misplaced Tables. The authors need to still review the text thoroughly for more consistent terminology and presentation.

In my understanding, the findings could be summarized as “There is a growing recognition of the importance of data stewardship among Italian universities and research institutions. However, there is no common professional profile for the personnel currently undertaking data stewardship activities in Italian research organizations. Moreover, the professional backgrounds of these personnel vary, which highlights the need for standardized training and clearer career paths for data stewards. The authors suggest a stronger national coordination to develop trainings, more standardized professional profiles, and career paths for data stewards in Italian research organizations.”

The above findings are often clouded in too general or vague terms, for instance, the authors seem to use the term “organizations” and “Italian research organizations” interchangeably. The authors should use the term “organizations” whenever they wish to make generalizations about research performing organizations in general and “Italian research organizations” when they

generalize about Italian research organizations, and “participating organizations” or “participating Italian organizations” when they speak about the collected data (also “research landscape” vs. “Italian research landscape” and “fragmentation” vs. “no unified professional profile”?). The authors should carefully revise the entire manuscript to provide a clearer terminology as suggested above, esp. the Abstract and Discussion segments.

There are many one or two sentence paragraphs. In my opinion, article texts in high-quality scientific journals are grouped into larger paragraphs, facilitating understanding and making the reading experience better. The authors should review the entire text and see if they can group the many one or two sentence paragraphs into meaningful larger paragraphs, esp. Data cleaning and visualization, Analysis of survey results, Discussion segments.

The sentences can be long and involve several complex ideas. Try also to divide the most condensed broad sentences into perhaps two or three sentences for better readability.

Cross-check all numbers presented in text so that they correspond with the numbers in Tables and Figures. Then, carefully check that all numerical results in text, Tables and Figures correspond with the published data.

Title and abstract

There is a discrepancy between an answer given to reviewers and the papers execution.
A title “Data stewardship in Italy: Findings from a national survey” was previously suggested by a reviewer, and the authors claimed to accept this suggestion (“hence we accepted the suggested form “Data stewardship in Italy: Findings from a national survey”) However, the current execution of the title is “Data stewardship in Italy: Findings from the first national survey”. I still suggest the authors to consider wording the title as “Data stewardship in Italy: Findings from a national survey” as the use of “first national survey” can be problematic.

Analysis of survey results

I believe Table 2 and Figure 4 are misplaced. I believe Table 2 should be presented before going into the RQ2 segment in the findings segment. Also, I believe Figure 4 should be presented before going into the RQ3 segment.

When discussing the results to RQ3 the authors first list the response options within the survey “The list of the response options included: […]” in the text leading to Figure 5. I find this unnecessary. These options are presented in Figure 5. I suggest the authors to find a work-around for this.

Discussion

Discussion – 2nd paragraph
“We observed that only a minority of respondents participated in the survey on behalf of their organisation. Consistent with the European trends reporting a correlation between institutional maturity in RDM and the adoption of DS roles (Gruber *et al.*, 2021; Jetten *et al.*, 2021; Seidlmayer *et al.*, 2023; Wildgaard *et al.*, 2020) it is reasonable to infer that organisations with higher levels of awareness and engagement in RDM are more likely to be developing or implementing structured

RDM support services and to have designated a representative for survey completion." I do not understand the relation between these two ideas? Can the RDM maturity level of an organization really be inferred based on whether the respondent saw themselves as being able to respond on behalf of the organization? And if so, do the authors then mean that the Italian ROs are currently at low-maturity levels with regards to RDM? I think the latter claim would need more discussion and findings behind it – could the authors explain and expand?

I recommend presenting a concluding Conclusion paragraph or segment to end the Discussion segment, after the limitations. The current concluding content is in different places and is difficult to read and understand ("Overall, this study provides initial insights into the evolution and development of the DS role in Italian research performing organisations. The findings suggest a growing recognition of DS functions and their gradual, though heterogeneous, integration into institutional practices and service frameworks, which appear to be diversified. The DS role is intrinsically interdisciplinary, combining research, technical and support competencies, yet the results highlight the need for clearer role definition, more structured training and stronger national coordination."). Please divide the content into separate sentences, expand your concluding discussion, and perhaps provide a separate Conclusion segment based on these findings (see also general comments about presenting the main findings more clearly)

References

I appreciate that the authors broadened their literature review. However, there are still some prominent studies missing from their list of references (see examples below). It is on the authors' ethical responsibility to cite the prior relevant literature and therefore it would be important to still search for and expand the list of referenced literature.

Wendelborn, C., Anger, M., & Schickhardt, C. (2023). What is data stewardship? Towards a comprehensive understanding. *Journal of Biomedical Informatics*, *140*, 104337.

Fitsilis, P., Damasiotis, V., Dervenis, C., Kyriatzis, V., & Tsoutsa, P. (2024). Effective data stewardship in higher education: Skills, competences, and the emerging role of open data stewards. *arXiv preprint arXiv:2410.20361*.

Mons, B. (2018). *Data stewardship for open science: Implementing FAIR principles*. CRC Press.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Data stewardship, Information science, scientific communication

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Version 1

Reviewer Report 12 March 2026

https://doi.org/10.21956/openreseurope.24014.r69758

© 2026 Samsuddin S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Samsul Farid Samsuddin
Universiti Malaya, Kuala Lumpur, Kuala Lumpur, Malaysia

Overall, it is a good study that have been conducted. However, as the authors aware and concerns, the findings of the study cannot generalize the whole country findings. It was only small portion of institution involved. The sampling approach (snowball sampling and followed approach) are not suitable for this specific study. The sampling should be conducted in more concise so that all related institutions will be much involved and not left behind, and the authors could find a better finding represent the current objective of the study. Dealing with online responses, always across the common method bias issues. Providing how the authors overcoming the issues will much strengthen the responses gathered throughout the study. The findings of data analysis also a bit confusing. Inconsistent numbers of response being analyzed throughout the study. Also not involving the whole 77 responses without any concrete justifications. Which in the earlier part, the study only involved responses representing their institution and not involving the responses as an individual which are the bigger portion. Not providing any justifications on each decision made on the analyses. The discussion part was very comprehensive and in-depth in contributing to the body of knowledge. Much to be improved if the manuscript to be indexed. Current manuscript not ready to be indexed which it will caused misunderstandings and confused to the readers. Details of comments are as follows:

1. Objectives
    1. Map the state of data stewardship in Italian universities and research-performing institutions.
    2. A further objective of the survey was to assess the respondents' willingness to join a national community of practice on RDM.
    - The objective of the study not clear. Should detailed out what the authors meant on the state.
    - Providing clear several research questions will help readers to relate with all the analysis conducted and findings interpretation.

2. There were no theoretical frameworks or underpinning theories involved in this study seems a bit loose.

3. Materials and methods

3.1 The survey was officially launched in April–May 2023 (from 13/04/2023 to 31/05/2023) and reopened in November–December 2023 (from 07/11/2023 to 01/12/2023).

- No explanations or justification why it was reopened.

3.2 Institutional Characterization: Included 13 closed-ended questions to explore how the DS role is defined and how related services are structured in the respondent's institution.

- Not stated how the questions were developed or not based on any developed

questionnaire.

3.3 Abstract – Data cleaning was performed on the 77 valid entries collected (out of 88 responses) before conducting the descriptive analyses.
Data cleaning and visualization (in content) – A total of 84 anonymous responses were collected and
exported as a .csv file, on which data cleaning was performed.

- The number of responses is not consistent. 88 or 84 responses?

3.4 Dissemination followed a snowball sampling approach, beginning with targeted outreach, through two national mailing lists focused on Open Science, and expanding the audience through word-of-mouth, institutional social media channels, and an article in GARR Magazine.

- The dissemination approach of the survey was not compatible with the objective of the study. To map the state of data stewardship in Italian universities and research-performing institutions, a strategic and systematic approach should be conducted at the first place. List out all the Italian universities and research-performing institutions and contacted the related departments or units will benefit the study. The findings will be better and it will help future research to be conducted in clearer path with comprehensive findings.

4. Analysis of survey results
The majority of the respondents belong to the Technical/Administrative support staff (n=38), while another important portion of respondents identified themselves as a Researcher (n=27). Only a small number of Technologists (n=8) answered the survey,

- The total number according to the data reported is 73 responses. Not similar with earlier claim on 77 valid entries collected.
- The authors did not stated the Italian universities and research-performing institutions involved in the study. This might help future scholars to understand the current state.

5. Authors should consider to have a separated sub-heading of Conclusion section for the two last paragraphs.

6. Ethics and consent: Authors stated that “Ethical approval and consent were not required for this submission.” There was no further justification on this statement. What is the reason of such study did not require any ethical approval and consent when it involved human response. A lot of potential harms might affect when the response was collected, especially when it’s on behalf of the institution and the privacy of the respondents.

Is the work clearly and accurately presented and does it engage with the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
No

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Areas of expertise on quantitative research and information science fields (including information and research data management aspects).

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 22 Jul 2026

Giulia Caldoni

We thank the reviewer for the careful assessment of our manuscript and for the valuable comments and suggestions. Their feedback helped us improve the clarity of the text and provide further explanation on several aspects that required additional attention. Below, we provide a detailed point-by-point response to each comment. Our responses follow each of the reviewer's comments.
Overall, it is a good study that have been conducted. However, as the authors aware and concerns, the findings of the study cannot generalize the whole country findings. It was only small portion of institution involved. The sampling approach (snowball sampling and followed approach) are not suitable for this specific study. The sampling should be conducted in more concise so that all related institutions will be much involved and not left behind, and the authors could find a better finding represent the current objective of the study.
Dealing with online responses, always across the common method bias issues. Providing how the authors overcoming the issues will much strengthen the responses gathered throughout the study. The findings of data analysis also a bit confusing. Inconsistent numbers of response being analyzed throughout the study. Also not involving the whole 77 responses without any concrete justifications. Which in the earlier part, the study only involved responses representing their institution and not involving the responses as an individual which are the bigger portion. Not providing any justifications on each decision made on the analyses. The discussion part was very comprehensive and in-depth in contributing to the body of knowledge. Much to be improved if the manuscript to be indexed.
Current manuscript not ready to be indexed which it will caused misunderstandings and confused to the readers. Details of comments are as follows: 1. Objectives

1. Map the state of data stewardship in Italian universities and research-performing institutions.
2. A further objective of the survey was to assess the respondents' willingness to join a national community of practice on RDM.

- The objective of the study not clear. Should detailed out what the authors meant on the state.
- Providing clear several research questions will help readers to relate with all the

analysis conducted and findings interpretation.

Authors' response: We thank the reviewer for the observation. We defined 4 Research Questions (RQs) in the *Introduction* section. We have revised the *Analysis of Survey Results* section to organise more clearly the presentation of the findings around the four RQs, ensuring that each RQ is clearly addressed in a dedicated subsection. Furthermore, we have revised the *Discussion* section to summarise more explicitly the key findings in relation to each RQ.

2. There were no theoretical frameworks or underpinning theories involved in this study seems a bit loose.

Authors' response: We thank the reviewer for the comment. We made the rationale more explicit in the *Materials and Methods* section. The authors' direct professional experience was complemented by insights drawn from international reports, allowing to transfer international experiences to the Italian context.

3. Materials and methods 3.1 The survey was officially launched in April–May 2023 (from 13/04/2023 to 31/05/2023) and reopened in November–December 2023 (from 07/11/2023 to 01/12/2023).

- No explanations or justification why it was reopened.

Authors' response: We thank the reviewer for the comment. To improve clarity, we revised the *Materials and methods* section to provide more details on the reasons for reopening the survey. We acknowledged and discussed potential limitations of the dissemination approach, including the survey being reopened during two separate periods, in the *Discussion* section. 3.2 Institutional Characterization: Included 13 closed-ended questions to explore how the DS role is defined and how related services are structured in the respondent's institution.

- Not stated how the questions were developed or not based on any developed questionnaire.

Authors' response: We thank the reviewer for the comment. The development of the questionnaire has been explained in the final paragraph of the "Survey and question structure" section of the *Materials and Methods*. 3.3 Abstract – Data cleaning was performed on the 77 valid entries collected (out of 88 responses) before conducting the descriptive analyses. Data cleaning and visualization (in content) – A total of 84 anonymous responses were collected and exported as a .csv file, on which data cleaning was performed.

- The number of responses is not consistent. 88 or 84 responses?

Authors' response: We thank the reviewer for pointing out the issue. 88 is the correct number, we identified and corrected a typo.

3.4 Dissemination followed a snowball sampling approach, beginning with targeted outreach, through two national mailing lists focused on Open Science, and expanding the audience through word-of-mouth, institutional social media channels, and an article in GARR Magazine.

- The dissemination approach of the survey was not compatible with the objective of the study. To map the state of data stewardship in Italian universities and research-performing institutions, a strategic and systematic approach should be conducted at the first place. List out all the Italian universities and research-performing institutions

and contacted the related departments or units will benefit the study. The findings will be better and it will help future research to be conducted in clearer path with comprehensive findings.

Authors' response: We thank the reviewer for this important comment. We agree that a systematic census-based approach targeting all Italian universities and research-performing institutions could provide a more comprehensive institutional mapping. However, our study was conceived as a bottom-up community-driven initiative, aimed at obtaining an initial overview of the emerging data stewardship landscape in Italy while simultaneously engaging practitioners and stakeholders active in the field. To avoid misunderstandings, we have clarified this aspect in the manuscript, explicitly stating that the survey dissemination strategy was designed to reach individuals and institutions through existing professional networks and communities of practice, rather than through a formal institution-by-institution sampling process. We acknowledge this as a limitation of the study and note that future research could build on our findings by adopting a more systematic and targeted institutional approach.

4. Analysis of survey results The majority of the respondents belong to the Technical/Administrative support staff (n=38), while another important portion of respondents identified themselves as a Researcher (n=27). Only a small number of Technologists (n=8) answered the survey,

- The total number according to the data reported is 73 responses. Not similar with earlier claim on 77 valid entries collected.

Authors' response:

- We thank the reviewer for pointing out the issue. 77 is the correct number (Technologists n=12), we identified and corrected a typo.
- The authors did not stated the Italian universities and research-performing institutions involved in the study. This might help future scholars to understand the current state.

Authors' response: We thank the reviewer for this comment. Details about the Italian universities and research-performing institutions involved in the study, are available in the dataset *Data_survey.csv* which has been deposited on Zenodo (https://doi.org/10.5281/zenodo.17907703).

5. Authors should consider to have a separated sub-heading of Conclusion section for the two last paragraphs.

Authors' response: We thank the reviewer for the comment. We thoroughly revised the Discussion section to make it clearer. We decided not to divide it into sections as we believe that it would have hindered the readability of the paragraph.

6. Ethics and consent: Authors stated that "Ethical approval and consent were not required for this submission." There was no further justification on this statement. What is the reason of such study did not require any ethical approval and consent when it involved human response. A lot of potential harms might affect when the response was collected, especially when it's on behalf of the institution and the privacy of the respondents.

Authors' response: We thank the reviewer for highlighting this important issue. We have

revised the Ethics Statement to clarify the rationale for not requiring formal ethical approval and individual written consent. First, the survey was anonymous. Participants were informed in the survey introduction that: "The survey is anonymous and therefore none of the answers will be related to the participant's personal data." No personal identifiers were collected, and no responses could be linked to individual respondents. Second, the questionnaire did not collect sensitive personal information, nor did it include questions regarding health status, political or religious beliefs, personal behaviors, psychological traits, or other topics that could expose participants to personal, social, professional, or legal risks. The survey focused exclusively on institutional aspects related to the presence, role, training, and activities of data stewards and similar research-support professionals within Italian universities and research institutions. Participation was voluntary, and respondents were informed about the purpose of the study before completing the questionnaire. By choosing to complete and submit the survey, participants provided their informed consent to participate.

Competing Interests: No competing interests were disclosed.

Reviewer Report 10 March 2026

https://doi.org/10.21956/openreseurope.24014.r69762

© **2026 Slifka J.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Jan Slifka
Czech Technical University in Prague, Prague, Prague, Czech Republic

This article presents a valuable first national survey of data stewardship in Italian universities and public research institutions. The topic is timely and relevant, the manuscript is well situated within the broader European discussion on research data stewardship, and the exploratory survey design is appropriate for mapping an emerging professional landscape. The paper also has important strengths in transparency: the authors make the underlying dataset, survey templates, and code available. The descriptive analysis is suitable for the size and nature of the sample, and the main message of the paper is useful: many institutions appear to have staff performing data-stewardship-related functions, but these roles are rarely formally recognised under that title.

My recommendation is **Approved with Reservations**, because the main revisions I would request concern clarity, internal consistency, and the framing of some interpretations, rather than the overall validity of the study. First, the manuscript would benefit from a single, transparent accounting of denominators throughout the Results. Some changing sample sizes are understandable and appear to follow the survey logic, since later analyses are restricted to respondents reporting DS or equivalent roles and several figures explicitly use subset denominators. However, the reporting is not always easy to follow, and one inconsistency seems to require clarification: the paper states that 77 valid responses remained after cleaning; the

institutional affiliation counts sum to 77 (33 from universities and 44 from public research institutions); yet the role counts reported in the same section (38 technical/administrative staff, 27 researchers, and 8 technologists) sum to 73. Because the methods state that roles were clustered into three macro-categories, the authors should clarify whether four responses were missing or unclassifiable, whether another category was omitted from the narrative, or whether the relevant denominator for that statement is smaller than 77. More generally, the manuscript would benefit from stating the applicable denominator more explicitly whenever a new subset of respondents is being analysed. Relatedly, there also appears to be some inconsistency in the description of the survey responses across sections: the Abstract mentions 88 responses, while the "Data cleaning and visualisation" section reports 84 collected responses before cleaning. Clarifying this progression of numbers (initial responses, cleaned dataset, and analysed subset) in a consistent way throughout the manuscript would improve transparency and readability.

Second, the Zenodo citations should be corrected and made internally consistent. In the Methods, the survey templates and dataset are cited as *Caldoni et al. (2025a)*, but in the reference list *Caldoni et al. (2025a)* corresponds to the *Manifesto of the "Italian data steward community"*. The data and software availability section instead points to *Caldoni et al. (2025b)* for the data package and lists the survey templates there, but the DOI in that section does not match the DOI given in the reference list. This appears to be a citation and findability issue rather than a lack of data sharing, but it should be corrected so that readers can unambiguously locate the survey templates, dataset, and code.

Third, the methods and limitations would benefit from a little more explanation in places. It would be helpful to explain why the survey was reopened, to clarify more explicitly how the activity list used in Figure 7 was compiled, and to discuss the likely effects of dissemination through Open Science-oriented mailing lists and related channels on the composition of the respondent pool. The paper already acknowledges that the sample covers only a limited number of institutions, and a slightly fuller discussion of representativeness and self-selection would strengthen the manuscript further.

Finally, I encourage the authors to soften a few interpretive statements in the Discussion and Conclusions so that they remain closely aligned with what a descriptive, self-selected survey can support. In particular, some wording could be made more cautious where the manuscript infers broader institutional visibility or communication patterns, suggests a "growing emphasis" on the technical dimension, or describes the proposed activity list as exhaustive. These points do not undermine the usefulness of the study, but a more restrained phrasing would make the conclusions stronger and more precise. With these revisions, I think the article would make a useful contribution to the literature on data stewardship and research support in national contexts.

Is the work clearly and accurately presented and does it engage with the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: My research and professional activities focus on research data management, data stewardship, and data management planning. I contribute to the development of the Data Stewardship Wizard tool and participate in ELIXIR and national initiatives related to data stewardship and RDM support in research institutions.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 22 Jul 2026

Giulia Caldoni

We thank the referee for the encouraging feedback and for the thoughtful suggestions. They allowed us to improve the text significantly, and to clarify some aspects that deserved more space and attention. Below, we provide a detailed point-by-point response to each comment. Our responses follow each of the reviewer's comments. This article presents a valuable first national survey of data stewardship in Italian universities and public research institutions. The topic is timely and relevant, the manuscript is well situated within the broader European discussion on research data stewardship, and the exploratory survey design is appropriate for mapping an emerging professional landscape.
The paper also has important strengths in transparency: the authors make the underlying dataset, survey templates, and code available. The descriptive analysis is suitable for the size and nature of the sample, and the main message of the paper is useful: many institutions appear to have staff performing data-stewardship-related functions, but these roles are rarely formally recognised under that title.
My recommendation is Approved with Reservations, because the main revisions I would request concern clarity, internal consistency, and the framing of some interpretations, rather than the overall validity of the study. First, the manuscript would benefit from a single, transparent accounting of denominators throughout the Results. Some changing sample sizes are understandable and appear to follow the survey logic, since later analyses are restricted to respondents reporting DS or equivalent roles and several figures explicitly use subset denominators.
However, the reporting is not always easy to follow, and one inconsistency seems to require

clarification: the paper states that 77 valid responses remained after cleaning; the institutional affiliation counts sum to 77 (33 from universities and 44 from public research institutions); yet the role counts reported in the same section (38 technical/administrative staff, 27 researchers, and 8 technologists) sum to 73. Because the methods state that roles were clustered into three macro-categories, the authors should clarify whether four responses were missing or unclassifiable, whether another category was omitted from the narrative, or whether the relevant denominator for that statement is smaller than 77. More generally, the manuscript would benefit from stating the applicable denominator more explicitly whenever a new subset of respondents is being analysed. Relatedly, there also appears to be some inconsistency in the description of the survey responses across sections: the Abstract mentions 88 responses, while the "Data cleaning and visualisation" section reports 84 collected responses before cleaning. Clarifying this progression of numbers (initial responses, cleaned dataset, and analysed subset) in a consistent way throughout the manuscript would improve transparency and readability.

Authors' response: We thank the reviewer for the very thorough comment, which helped us revise the consistency of the results that we considered to respond to each research question. Particularly, we identified and corrected some typos which hindered the understanding:

- Role sum: we reported 8 technologists while 12 was the correct number, hence 77 is the correct total. The typo error has now been corrected.
- Number of total responses: 88 is the correct number, now reported in the text.

We also specified, when answering each research question, the pool of responses we are basing our analysis on in order to improve clarity. Second, the Zenodo citations should be corrected and made internally consistent. In the Methods, the survey templates and dataset are cited as *Caldoni et al. (2025a)*, but in the reference list *Caldoni et al. (2025a)* corresponds to the *Manifesto of the "Italian data steward community"*. The data and software availability section instead points to *Caldoni et al. (2025b)* for the data package and lists the survey templates there, but the DOI in that section does not match the DOI given in the reference list. This appears to be a citation and findability issue rather than a lack of data sharing, but it should be corrected so that readers can unambiguously locate the survey templates, dataset, and code.

Authors' response: We thank the reviewer for the observation. The references have been corrected in the revised version of the article. Third, the methods and limitations would benefit from a little more explanation in places. It would be helpful to explain why the survey was reopened, to clarify more explicitly how the activity list used in Figure 7 was compiled, and to discuss the likely effects of dissemination through Open Science-oriented mailing lists and related channels on the composition of the respondent pool. The paper already acknowledges that the sample covers only a limited number of institutions, and a slightly fuller discussion of representativeness and self-selection would strengthen the manuscript further.

Authors' response: We thank the reviewer for the observation. To improve clarity, we revised the *Materials and methods* section to provide more details on the reasons for reopening the survey. We acknowledged and discussed potential limitations of the

dissemination approach, including the survey being reopened, in the *Discussion* section. Finally, I encourage the authors to soften a few interpretive statements in the Discussion and Conclusions so that they remain closely aligned with what a descriptive, self-selected survey can support. In particular, some wording could be made more cautious where the manuscript infers broader institutional visibility or communication patterns, suggests a "growing emphasis" on the technical dimension, or describes the proposed activity list as exhaustive. These points do not undermine the usefulness of the study, but a more restrained phrasing would make the conclusions stronger and more precise. With these revisions, I think the article would make a useful contribution to the literature on data stewardship and research support in national contexts.

Authors' response: We thank the reviewer for the suggestion. To improve the paper readability and the interpretation of the results, we carefully revised the manuscript to ensure the use of more cautious wording and to tone-down expressions that might have overemphasized the results.

Competing Interests: No competing interests were disclosed.

Reviewer Report 09 March 2026

https://doi.org/10.21956/openreseurope.24014.r69757

© 2026 Adamaki A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Angeliki Adamaki
SciLifeLab Data Centre, Uppsala University (Ringgold ID: 463758), Uppsala, Sweden

Summary

The article presents the results of a national survey exploring the presence, roles, and activities of professionals involved in research data management (RDM) in Italian universities and research institutions. The survey collected 77 valid responses (from 49 distinct institutions based on the CSV file and the ROR IDs provided) and provides an initial overview of how data stewardship-related tasks are currently distributed across institutional roles. The findings show that several institutions (25 based on the ROR IDs provided in the CSV data file and filtered for those knowing that they have DSs in their institutions) rely on staff performing data stewardship activities, such as support/guidance for FAIR data practices, data management plans, training, and policy guidance, although these responsibilities are not always formally recognised under the title of data steward. The study highlights a fragmented professional landscape and emphasises the need for clearer role definitions and professional pathways. At the same time, the work has potential impact beyond the Italian context, as it provides a landscape analysis at national level, contributes to community-building efforts through the creation of the Italian Data Steward Community, and proposes a methodological approach that could support similar mapping exercises in other countries.

General Comments

The article has the potential to generate significant impact, as the survey could contribute to:

- landscape analyses of data stewardship practices following a reusable approach,
- community-building efforts,
- the development of a roadmap for the professionalisation of data stewardship.

Some statements in the article appear to reflect community experience rather than being directly supported by the presented data. Where possible, the authors should clarify this. Given the relatively small sample size (for statistical analysis), some statistical interpretations in the manuscript should be treated with caution. Qualitative expressions such as “a few” or “substantial” or “significant” can be difficult to interpret without corresponding numerical values (e.g. providing absolute numbers alongside percentages) which help readers better assess the significance of the results. Some recommendations for improvements are given below, considering the sample size and the visualisation options.

Throughout the manuscript, the activities of data stewards are frequently described as providing “support” to researchers. While this term is common in RDM contexts, it may underestimate the broader role that data stewards often play (as also discussed in the article). In many cases, data stewards act not only as support staff but also as enablers of research, contributing to the development of infrastructures and relevant policies, and the design of research environments, frameworks and workflows, that make FAIR and sustainable data practices possible.

The authors also mention “policy” in some parts of the text (e.g. “policy consultation”, “policymakers”), it will be good to clarify if they mean data policy, institutional policy, Open Science policy etc. GARR (pg. 5) is mentioned without explanation of the acronym and/or link.

The article builds on the assumption that the importance of RDM is recognised within institutions (e.g. beginning of section “Data stewardship in Italy”, and in “Discussion”). It would be interesting to see (perhaps in a linked, follow up analysis) how this recognition emerged in the Italian context, and which factors contributed to it. For example:

- the role of research infrastructures (RIs),
- the influence of European policies or funding programmes,
- the identification of researchers’ needs perhaps via institutional initiatives.

Future developments could also consider connections with initiatives such as research assessment reform and recognition frameworks, which may influence how such roles are acknowledged within academic institutions.

Questions (Q)

Q1: (related to respondents) Although it becomes clear as the analysis is presented, the authors can clarify early in the survey description that the participants were not necessarily data stewards themselves, and they replied either representing their institutions or based on their personal experiences.

Q2: (related to EUSurvey) Are there technical or ethical considerations that others should know if they want to replicate the methodology?

Q3: (related to Figure 1) Can the authors comment on survey coverage relative to the number of universities and research infrastructures in each region? This would help assess how representative the survey is. The figure helps the reader understand the sample, but not the data coverage. Of course each institution might have diverse departments with different needs in data

stewardship, and only one response from a specific organisation might not be representative, so this should be clarified as well.
Q4: Instead of excluding data from part of the analysis, would it be possible to present all responses, while clustering for individuals and institutional representatives? If subset of the dataset eventually is used (as it is now), the authors should explain this choice.
Q5: (pg 7) "The majority of the respondents belong to.", the numbers in parentheses add up to 73, but total number is 77?
Q6: (Figures 2-3) Would the authors consider combining the figures in e.g. clustered or stacked bar chart? While pie charts are commonly used to visualize percentage, including counts alongside percentages will improve interpretability.
Q7: (Figure 4a) When looking at the free-text responses there is at least one response where the participant writes that there is one officially identified DS and then other cases in their organisation, although they earlier answer "No" to whether there is a formally identified DS. Since the sample size is small, such details change the results presented. Besides this, the data shows that the respondents come from different institutions, which shows good data coverage and the authors could consider mentioning this.
Q8: (General recommendation on figures) The authors may consider adapting the type of visualisation to the number of categories presented. For statistics with only a few categories (e.g., 2–4), pie charts can be effective, provided that absolute counts are shown alongside percentages. For analyses involving larger numbers of categories, horizontal bar charts may be more appropriate (absolute counts shown alongside percentages), as they allow category labels to be placed next to the bars, improving readability.
Q9: (pg 8) The text describing Figure 6 can be improved if rephrased. The numbers in parentheses add up to 60, but these cannot be the respondents (n=38). The authors probably list the answers they collected as each respondent could give multiple answers.
Q10: (Figure 5) Similar comments as in Q8 about the visualisation. If there are 13 respondents, then 1/13 is the minimum we expect. How did the 5% shown on the pie chart result?
Q11: (Figure 7 and corresponding text) The authors mention percentages in the text, which are not shown on the Figure or in a table, and it's not very easy to follow. On the figure it's probably not necessary to mention the numbers at each bar, as they don't add up to the number of respondents (multiple answers). These could be changed to match the text. If the aim of this part of the analysis is to show how the tasks of a DS are distributed, then the authors might consider describing this as e.g. RDM tasks are reported in x% of cases together with DMP, y% also involves training, etc.
Q12: (Figure 8) Would it be interesting to distinguish between Unis and RIs?
Q13: (Figure 9b) Similar comments as in Q8 regarding the visualisation in pie chart.
Q14: (pg 12) "Conversely, it is notable that about 35%..", although it's indeed 7/20 responses, those do not correspond to 35% of the survey responses as the observation only refers to those representing their institutions. The authors might consider rephrasing this part, and perhaps also mentioning that these 20 responses seem to be from 20 different organisations.
Q15: (pg 12) "..we can hypothesise..", is this an assumption based on the authors' experience from the community activities and fora?
Q16: (pg 12) "..who answered in a personal capacity may reflect..", could this be also because of weaknesses in the internal structure of each organisation, no mandate, etc.?
Q17: (pg 12) "..difficulties in recognising who qualifies as a DS..", can the authors elaborate?
Q18: (pg 13) "..Additionally the notable representation of respondents..", how is the role of the respondents relevant to the survey?
Q19: (pg 13) There are some statements that are given as facts, while they could be better

expressed as possible explanations based on the authors' experience from their work, the DS community, etc. The survey presented here is a snapshot of time, and it would be interesting to see how these observations change and hopefully evolve.

Is the work clearly and accurately presented and does it engage with the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research Data Management and Open Science, with a focus on Research Infrastructures and community/service development.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 22 Jul 2026

Giulia Caldoni

We thank the referee for the encouraging feedback and for the thoughtful suggestions. They allowed us to improve the text significantly, and to clarify some aspects that deserved more space and attention. Below, we provide a detailed point-by-point response to each comment. Our responses follow each of the reviewer's comments.

Summary The article presents the results of a national survey exploring the presence, roles, and activities of professionals involved in research data management (RDM) in Italian universities and research institutions. The survey collected 77 valid responses (from 49 distinct institutions based on the CSV file and the ROR IDs provided) and provides an initial overview of how data stewardship-related tasks are currently distributed across institutional roles. The findings show that several institutions (25 based on the ROR IDs provided in the CSV data file and filtered for those knowing that they have DSs in their institutions) rely on staff performing data stewardship activities, such as support/guidance for FAIR data

practices, data management plans, training, and policy guidance, although these responsibilities are not always formally recognised under the title of data steward.
The study highlights a fragmented professional landscape and emphasises the need for clearer role definitions and professional pathways. At the same time, the work has potential impact beyond the Italian context, as it provides a landscape analysis at national level, contributes to community-building efforts through the creation of the Italian Data Steward Community, and proposes a methodological approach that could support similar mapping exercises in other countries.

General Comments The article has the potential to generate significant impact, as the survey could contribute to:

- landscape analyses of data stewardship practices following a reusable approach,
- community-building efforts,
- the development of a roadmap for the professionalisation of data stewardship.

Some statements in the article appear to reflect community experience rather than being directly supported by the presented data. Where possible, the authors should clarify this. Given the relatively small sample size (for statistical analysis), some statistical interpretations in the manuscript should be treated with caution. Qualitative expressions such as “a few” or “substantial” or “significant” can be difficult to interpret without corresponding numerical values (e.g. providing absolute numbers alongside percentages) which help readers better assess the significance of the results. Some recommendations for improvements are given below, considering the sample size and the visualisation options.
Authors’ response: We thank the reviewer for the suggestions. Accordingly, we have revised the wording throughout the manuscript to avoid any implication of inferential statistical significance and to more accurately reflect the descriptive and exploratory nature of the study. We have deleted qualitative expressions where appropriate and consistently reported absolute numbers alongside percentages throughout the manuscript. We also acknowledge the reviewer's concern regarding the relatively small sample size and have more explicitly addressed this limitation in the *Discussion* section. Throughout the manuscript, the activities of data stewards are frequently described as providing “support” to researchers. While this term is common in RDM contexts, it may underestimate the broader role that data stewards often play (as also discussed in the article). In many cases, data stewards act not only as support staff but also as enablers of research, contributing to the development of infrastructures and relevant policies, and the design of research environments, frameworks and workflows, that make FAIR and sustainable data practices possible. The authors also mention “policy” in some parts of the text (e.g. “policy consultation”, “policymakers”), it will be good to clarify if they mean data policy, institutional policy, Open Science policy etc. GARR (pg. 5) is mentioned without explanation of the acronym and/or link.
Authors’ response: We thank the reviewer for the observation, which helped us improve the precision of the manuscript. To better reflect the scope and specificity of the profession, we replaced several occurrences of “support to data management” with the terms “data stewardship” and “DS role(s)” throughout the text. Moreover, we clarified the meaning of “policies” in the *Introduction* section. explained acronyms. We revised the text to explain acronyms, except “GARR” which can’t be explained as it is not an acronym. We provided a link to the Consortium GARR webpage for clarity. The article builds on the assumption that

the importance of RDM is recognised within institutions (e.g. beginning of section “Data stewardship in Italy”, and in “Discussion”). It would be interesting to see (perhaps in a linked, follow up analysis) how this recognition emerged in the Italian context, and which factors contributed to it. For example:

- the role of research infrastructures (RIs),
- the influence of European policies or funding programmes,
- the identification of researchers’ needs perhaps via institutional initiatives.

Future developments could also consider connections with initiatives such as research assessment reform and recognition frameworks, which may influence how such roles are acknowledged within academic institutions.
Authors’ response: We thank the reviewer for this suggestion. Although we decided to remove from the manuscript the reference to a future, expanded version of the survey, this remains one of the objectives of the Italian Data Stewards Community, whose establishment was facilitated by this initial survey. We will take the reviewer’s suggestions into account when planning future survey activities.

Questions (Q) Q1: (related to respondents) Although it becomes clear as the analysis is presented, the authors can clarify early in the survey description that the participants were not necessarily data stewards themselves, and they replied either representing their institutions or based on their personal experiences.
Authors’ response: We thank the reviewer for the observation. We clarified this aspect in the *Materials and Methods* section when describing the structure of the questionnaire.

Q2: (related to EUSurvey) Are there technical or ethical considerations that others should know if they want to replicate the methodology?
Authors’ response: We thank the reviewer for the question. EUSurvey is an open-source platform provided by the European Commission, which ensures compliance with European data protection standards. It also allows anonymous responses, with no personal or sensitive data being collected or requested. This has been added in the ethical statement of the article.

Q3: (related to Figure 1) Can the authors comment on survey coverage relative to the number of universities and research infrastructures in each region? This would help assess how representative the survey is. The figure helps the reader understand the sample, but not the data coverage. Of course each institution might have diverse departments with different needs in data stewardship, and only one response from a specific organisation might not be representative, so this should be clarified as well.
Authors’ response: We thank the reviewer for this interesting observation. We agree that assessing the geographical coverage of the survey at the regional level would provide additional insight, particularly considering that responses were obtained from at least one university or research organisation in all but three Italian regions. However, conducting such an analysis would require a more detailed assessment of institutional coverage, as the Italian research landscape is highly fragmented and organisationally complex. For example, the National Research Council of Italy (CNR), which counts as a single research organisation, comprises more than 200 seats distributed across the country. To address the reviewer(s) question, we have revised the manuscript to specify that responses were obtained from 17

of the 104 universities and research-performing organisations operating in Italy. We have discussed that this underrepresentation may affect the generalisation of the findings. We also clarified that the primary aim of the study is to provide a snapshot of current practices rather than a fully comprehensive or statistically representative national mapping. These additions can be found in the *Discussion* section.

Q4: Instead of excluding data from part of the analysis, would it be possible to present all responses, while clustering for individuals and institutional representatives? If subset of the dataset eventually is used (as it is now), the authors should explain this choice.
Authors' response: We thank the reviewer for the suggestion. We deeply revised the Analysis of Survey Results section to clearly specify which subset of the dataset was used to answer every Research Question (RQ). We decided to consider only responses provided on behalf of organisations (n=20) to answer RQ1 ("To what extent are DSs formally integrated in research performing organizations in Italy?") as our interest was focused on capturing the formal and institutional definition of the DS role. Further RQs aimed, on the other hand, to characterize the professional identity of DSs, hence we based our analysis on the total of answers received.

Q5: (pg 7) "The majority of the respondents belong to.", the numbers in parentheses add up to 73, but total number is 77?
Authors' response: We thank the reviewer for the comment. There was a typo concerning the number of Technologists: 8 instead of 12. The typo error has now been corrected.

Q6: (Figures 2-3) Would the authors consider combining the figures in e.g. clustered or stacked bar chart? While pie charts are commonly used to visualize percentage, including counts alongside percentages will improve interpretability.
Authors' response: We thank the reviewer for the observation, shared also by referee 2. We removed figure 2 and reported the numbers only in the text. We maintained figure 3 as we think it's relevant to give a visual snapshot of the number of formally recognized data stewards in the national landscape. We also removed pie charts and used tables instead.

Q7: (Figure 4a) When looking at the free-text responses there is at least one response where the participant writes that there is one officially identified DS and then other cases in their organisation, although they earlier answer "No" to whether there is a formally identified DS. Since the sample size is small, such details change the results presented. Besides this, the data shows that the respondents come from different institutions, which shows good data coverage and the authors could consider mentioning this.
Authors' response: We thank the reviewer for the very attentive comment. When answering RQ1, we based our quantitative analyses on responses to the structured survey questions. In the specific case that is pointed out, the respondent selected "No" when asked whether the individuals performing data stewardship functions were formally identified as data stewards. Furthermore, the accompanying comment suggests that, although at least one individual may hold the title of Data Steward, data stewardship activities within the organisation are predominantly carried out by personnel who are not formally designated as such. Therefore, the response was considered consistent with the intended interpretation of the survey question and was classified in the "No" category. According to the reviewer's suggestion, we commented on the good institutional data coverage at the

beginning of the *Analysis of Survey Results* section.

Q8: (General recommendation on figures) The authors may consider adapting the type of visualisation to the number of categories presented. For statistics with only a few categories (e.g., 2–4), pie charts can be effective, provided that absolute counts are shown alongside percentages. For analyses involving larger numbers of categories, horizontal bar charts may be more appropriate (absolute counts shown alongside percentages), as they allow category labels to be placed next to the bars, improving readability.
Authors' response: We thank the reviewer for this comment. Figure 4b has been converted into a table to clearly present the different categories and the number of responses for each.

Q9: (pg 8) The text describing Figure 6 can be improved if rephrased. The numbers in parentheses add up to 60, but these cannot be the respondents (n=38). The authors probably list the answers they collected as each respondent could give multiple answers.
Authors' response: We thank the reviewer for the comment. For the analysis now reported in RQ3 and RQ4 we considered all respondents who reported the presence of Data Stewards (or equivalent roles) within their organisation (n = 38), regardless of whether the questionnaire had been completed on behalf of the institution or in an individual capacity. Multiple (non–mutually exclusive) responses were allowed. To avoid misunderstanding, we have clarified this selection criterion in the manuscript and in the corresponding figure captions.

Q10: (Figure 5) Similar comments as in Q8 about the visualisation. If there are 13 respondents, then 1/13 is the minimum we expect. How did the 5% shown on the pie chart result?
Authors' response: We thank the reviewer for the comment. Figure 5 has been converted into a table to better evaluate the number of responses for each role. Multiple (non-mutually exclusive) responses were allowed.

Q11: (Figure 7 and corresponding text) The authors mention percentages in the text, which are not shown on the Figure or in a table, and it's not very easy to follow. On the figure it's probably not necessary to mention the numbers at each bar, as they don't add up to the number of respondents (multiple answers). These could be changed to match the text. If the aim of this part of the analysis is to show how the tasks of a DS are distributed, then the authors might consider describing this as e.g. RDM tasks are reported in x% of cases together with DMP, y% also involves training, etc.
Authors' response: We thank the reviewer for the suggestion. Accordingly, we have reported absolute numbers alongside percentages throughout the manuscript.

Q12: (Figure 8) Would it be interesting to distinguish between Unis and RIs?
Authors' response: We thank the reviewer for the suggestion. We did not report results stratified by institution type, as this distinction was not central to the study objectives and did not contribute substantially to the overall interpretation of data stewardship distribution in Italy.

Q13: (Figure 9b) Similar comments as in Q8 regarding the visualisation in pie chart.

Authors' response: We thank the reviewer for the comment. Figure 9b has been converted into a table to better present results as absolute numbers. Multiple (non-mutually exclusive) responses were allowed.

Q14: (pg 12) "Conversely, it is notable that about 35%..", although it's indeed 7/20 responses, those do not correspond to 35% of the survey responses as the observation only refers to those representing their institutions. The authors might consider rephrasing this part, and perhaps also mentioning that these 20 responses seem to be from 20 different organisations. **Authors' response:** We thank the reviewer for this comment. The sentence has been rephrased as follows: "*We also note that 35% of the respondents (7 out of 20, Figure 3) reported the absence of DSs in their institution, suggesting uneven levels of awareness on this topic.* " We did not further comment on the responding organisations.

Q15: (pg 12) "..we can hypothesise..", is this an assumption based on the authors' experience from the community activities and fora?
Authors' response: We thank the reviewer for the observation. We carefully revised the text to avoid over-interpretations of the results, hence the specific sentence mentioned has been deleted.

Q16: (pg 12) "..who answered in a personal capacity may reflect..", could this be also because of weaknesses in the internal structure of each organisation, no mandate, etc.?
Authors' response: We have included the reviewer's suggestion in the updated version of the article. Now the sentence has been rephrased as follows: "*We observed that only a minority of respondents participated in the survey on behalf of their institution. Consistent with the European trends reporting a correlation between institutional maturity in RDM and the adoption of DS roles (Gruber, 2021; Jetten et al., 2021; Seidlmayer et al., 2023; Wildgaard et al., 2020), it is reasonable to infer that institutions with higher levels of awareness and engagement in RDM are more likely to be developing or implementing structured RDM support services and to have designated a representative for survey completion.*"

Q17: (pg 12) "..difficulties in recognising who qualifies as a DS..", can the authors elaborate?
Authors' response: We thank the reviewer for this comment. The sentence has been removed, and the key concepts have been revised and rephrased throughout the manuscript.

Q18: (pg 13) "..Additionally the notable representation of respondents..", how is the role of the respondents relevant to the survey?
Authors' response: We thank the reviewer for the comment. Respondents' roles were collected to contextualize the survey findings. Since participation was open to professionals beyond formal data stewards, the results reflect a broader data stewardship community. The survey did not explicitly assess whether responses were based on personal experience. We modified the text to address this point both in the *Analysis of Survey Results* and in the *Discussion* sections.

Q19: (pg 13) There are some statements that are given as facts, while they could be better expressed as possible explanations based on the authors' experience from their work, the DS community, etc. The survey presented here is a snapshot of time, and it would be

interesting to see how these observations change and hopefully evolve.
Authors' response: We thank the reviewer for this comment. We revised the manuscript and rephrased some sentences to avoid over-interpretation of the results. Although we decided to remove from the manuscript the reference to a future, expanded version of the survey, this remains one of the objectives of the Italian Data Stewards Community, whose establishment was facilitated by this initial survey. We will take the reviewer's suggestions into account when planning future survey activities.

Competing Interests: No competing interests were disclosed.

Reviewer Report 09 March 2026

https://doi.org/10.21956/openreseurope.24014.r69765

© 2026 Saul L. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Louise Saul
University of Southampton, Southampton, England, UK

This article describes a survey that was circulated to research institutions in Italy, via a multifaceted snowball approach, within 2 two-month time periods in the year 2023.

It builds upon previous work by other groups to define and professionalise the role of Data Steward in the research setting, particularly looking at work in the Netherlands (1), Germany (2), and Denmark (3) and seeks to undertake the same project in an Italian setting. The introduction of this manuscript frames the narrative and need for this work.

This is an extremely timely piece of work as various EU projects came to a close there was a realisation of the requirement for a continued effort to define the Data Steward role across Europe. However, whilst there needs to be a united effort to ensure that a ubiquitous language is used to describe this role, each nation's research ecosystem will be subtly unique. Understanding this variability is key to ensuring a cohesive approach to embedding Data Stewardship, which necessitates an understanding of the current landscape.

This work seeks to describe the current status of both Data Stewards and people who undertake work that would be completed by Data Stewards at research institutions in Italy, capturing information about their activities, responsibilities, geographical location, official roles, training requirements, recognition for their work, and area in which they are located.

This is a very important piece of work; the survey is comprehensive without being a burden for the user to complete. I was particularly interested in the approach taken to include a description of a 'Data Steward' at the beginning of the survey; it may have been interesting to allow respondent to comment upon this definition, however I appreciate that may complicate the analysis. It would be

interesting to understand the rationale for developing the survey and inclusion of questions, however in emerging fields there is often not sufficient information to do so and as stated above, often national landscapes can be unique enough to prevent a direct transposition. I assume that the survey design was based upon previous work (4).

Promoting reproducibility and data re-use for systematic review and meta-analyses, the authors have made the data, survey, and analysis code fully available with persistent identifiers, also including information about the methods used to clean the data. It would have been interesting to include more information about the institution 'types' that respondent worked at (i.e. Independent research institute, higher education, government sponsored institute) in order to fully understand the context in which the answers were given.

The way the responses have been summarised in this manuscript presents an interesting and useful picture that will be useful to compare to other countries' summaries of the same role. I have some questions over the breakdown of responses for figure 4b and figure 5, where descriptive statistic may not be helpful for so few respondents and a table or similar may be more appropriate. I am also unclear as to why only 38 respondents (I presume identified from the 'Technical/Administrative staff' cluster) were included in the analysis for figure 6/7/8/9, but this may be as a result of my own understanding. The results were illuminating and are very valuable for understanding the role of Data Steward in Italian research organisations.

I found the discussion to be very insightful and appreciated the summary that the authors provided. I look forward to seeing outputs from the Italian Data Stewards Community.

References

1. Jetten, M., Grootveld, M., Mordant, A., Jansen, M., et al.: Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship (1.1). *https://doi.org/10.5281/zenodo.4623713*. 2021.
2. Newbold, E., Wang, Y., Lehtsalu, L., Zhou, B., et al.: RDA Professionalising Data Stewardship IG - What does a career track for data stewards look like?. *https://doi.org/10.15497/RDA00102*. 2024.
3. Wildgaard, L., Vlachos, E., Nondal, L., Larsen, A. V., et al.: National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management (DM Forum) (Version 1). *https://doi.org/10.5281/zenodo.3609516*. 2020.
4. Förstner, K. U., Seidlmayer, E., ZB MED-Informationszentrum Lebenswissenschaften, Bundesministerium für Bildung und Forschung, et al.: Forschung unterstützen - Empfehlungen für Data Stewardship an akademischen Forschungsinstitutionen. *https://doi.org/10.4126/FRL01-006441397*. 2023.

Is the work clearly and accurately presented and does it engage with the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Data Stewardship skills and careers

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 22 Jul 2026

Giulia Caldoni

We sincerely thank the reviewer for the positive feedback and insightful comments, which enabled us to refine the manuscript. Below, we provide a detailed point-by-point response to each comment. Our responses follow each of the reviewer's comments.
This article describes a survey that was circulated to research institutions in Italy, via a multifaceted snowball approach, within 2 two-month time periods in the year 2023. It builds upon previous work by other groups to define and professionalise the role of Data Steward in the research setting, particularly looking at work in the Netherlands (1), Germany (2), and Denmark (3) and seeks to undertake the same project in an Italian setting. The introduction of this manuscript frames the narrative and need for this work. This is an extremely timely piece of work as various EU projects came to a close there was a realisation of the requirement for a continued effort to define the Data Steward role across Europe. However, whilst there needs to be a united effort to ensure that a ubiquitous language is used to describe this role, each nation's research ecosystem will be subtly unique. Understanding this variability is key to ensuring a cohesive approach to embedding Data Stewardship, which necessitates an understanding of the current landscape. This work seeks to describe the current status of both Data Stewards and people who undertake work that would be completed by Data Stewards at research institutions in Italy, capturing information about their activities, responsibilities, geographical location, official roles, training requirements, recognition for their work, and area in which they are located.
This is a very important piece of work; the survey is comprehensive without being a burden for the user to complete. I was particularly interested in the approach taken to include a description of a 'Data Steward' at the beginning of the survey; it may have been interesting to allow respondent to comment upon this definition, however I appreciate that may complicate the analysis. It would be interesting to understand the rationale for developing the survey and inclusion of questions, however in emerging fields there is often not sufficient information to do so and as stated above, often national landscapes can be unique enough to prevent a direct transposition. I assume that the survey design was based upon previous work (4).

Authors' response: We thank the reviewer for the observation. We made the rationale more explicit in the *Materials and Methods* section. The authors' direct professional experience was complemented by insights drawn from international reports, allowing to transfer international experiences to the Italian context. Promoting reproducibility and data re-use for systematic review and meta-analyses, the authors have made the data, survey, and analysis code fully available with persistent identifiers, also including information about the methods used to clean the data. It would have been interesting to include more information about the institution 'types' that respondent worked at (i.e. Independent research institute, higher education, government sponsored institute) in order to fully understand the context in which the answers were given. The way the responses have been summarised in this manuscript presents an interesting and useful picture that will be useful to compare to other countries' summaries of the same role. I have some questions over the breakdown of responses for figure 4b and figure 5, where descriptive statistic may not be helpful for so few respondents and a table or similar may be more appropriate.

Authors' response: We thank the reviewer for the observation. We decided to use tables to report the breakdown data that were previously presented in the pie charts of figures 4b and 5b. I am also unclear as to why only 38 respondents (I presume identified from the 'Technical/Administrative staff' cluster) were included in the analysis for figure 6/7/8/9, but this may be as a result of my own understanding. The results were illuminating and are very valuable for understanding the role of Data Steward in Italian research organisations.

Authors' response: We thank the reviewer for the observation, which helped us clarify the reporting of results. For the analysis now reported in RQ3 and RQ4 we considered all respondents who reported the presence of Data Stewards (or equivalent roles) within their organisation (n = 38), regardless of whether the questionnaire had been completed on behalf of the institution or in an individual capacity. Multiple (non–mutually exclusive) responses were allowed. To avoid misunderstanding, we have clarified this selection criterion in the manuscript and in the corresponding figure captions. I found the discussion to be very insightful and appreciated the summary that the authors provided. I look forward to seeing outputs from the Italian Data Stewards Community.
Authors' response: We thank the reviewer for the positive assessment of the value of these results. We are happy to share that all the outputs from the Italian Data Steward Community are available in our Zenodo community:
https://zenodo.org/communities/italiads/records?q=&l=list&p=1&s=10&sort=newest

Competing Interests: No competing interests were disclosed.

Reviewer Report 09 March 2026

https://doi.org/10.21956/openreseurope.24014.r69761

© 2026 Rousi A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Antti M. Rousi
Aalto University, Espoo, Espoo, Finland

Congratulations to the authors for this interesting research. This work was a pleasure to review. I agree that the present study -- along with its valuable open data -- could serve as a benchmark for similar national questionnaires mapping the state of data stewardship in other countries. Although the number responses may seem small, these respondents represent a highly specialized and rare group of professionals. I believe that the final version of this work could be indexed in ORE.

However, I also believe this work must be improved before it can be accepted for indexing. Although my general sentiment for this research is positive and I wish to support the authors' research, I recommend a major revision followed by a new round of reviews. I hope the authors find my suggestions helpful.

Work title

The current manuscript is titled "Mapping data stewardship in Italy: Findings from the first national survey". I understand the use of term "mapping" within this context, but I wonder if the title would work better with search engines without it. Also, the term "first" seems unnecessary in the title. I invite the authors to contemplate whether the title could be shortened into "Data stewardship in Italy: Findings from a national survey".

Terminology and RQs

The understandability of the article suffers from terminological discrepancies.

- The abstract states the research aim as to "map the presence, roles, and activities of data stewards or equivalent professionals across universities and research institutions, to provide an initial and systematic overview of the national landscape"
- The introduction, however, seems to be built upon the three following core concepts of data stewardship, data stewardship profiles, and data stewardship service models
- The end of the Introduction states that "[...] the idea emerged to launch a public survey to map the state of data stewardship in Italian universities and research-performing institutions. A further objective of the survey was to assess the respondents' willingness to join a national community of practice on RDM."
- Furthermore, within the methodology section, the structure of the survey questionnaire is based on "Respondent Profiles" and "Institutional Characterisation".

This terminological looseness is exacerbated by the fact that the research aim and research questions are not addressed in the article text.

Looking at the survey questionnaire, the research aim and questions could perhaps be formulated, for instance, as follows: "The research aim of the present work is to investigate how data stewardship services are organised within Italian universities and research institutions. The research questions are posed as follows.

- RQ1 What kinds of professionals undertake data stewardship in Italian universities and

research institutions?
- RQ2 Which service organizations employ the personnel undertaking data stewardship activities?
- RQ3 Which kinds of data stewardship support actions are these professionals currently undertaking?"

I believe the authors must:
- Provide a general introduction to their research aim and define their core terminology in the Introduction before the literature review segment
- Link the literature review with the core terminology defined in the Introduction (e.g. "The question of what kinds of professionals undertake data stewardship is addressed in the research branch focussing on DS profiles"; note! I am not sure if the latter is factually correct; it is more of a style example)
- Define their research aim and research questions in detail after the literature review segment (see the example RQs above)
- Make sure the methods section is terminologically aligned with the research aim and RQs
- Base the reporting of their survey results around the RQs, e.g., one Findings sub segment per RQ
- Briefly summarize the most important findings per RQ in the Discussion segment

Literature review

The authors must improve their literature review. The draft only cites 17 works, which is a small amount for a research article. Furthermore, the draft mostly cites EOSC, RDA and national working group reports when of course the literature on data stewardship is more broad. A simple Google Scholar search with the term "data stewardship" produces several relevant works currently missing from cited references. Besides research articles focussing on data stewardship, also literature on how academic libraries have arranged RDM services are in my view relevant to this work. As a rule of thumb, the authors should strive to double the amount of cited references when adding research articles into the list of references.

Misc. notes and improvement suggestions
- For a better flow of the current Introduction text, I suggest the authors experiment with removal of the subtitles (e.g. "International efforts to professionalise research data stewardship"). Furthermore, the authors could experiment with moving the current 2nd paragraph of introduction as the new 4th paragraph of introduction. The latter solution could improve the text flow introducing the concepts of DS profiles and DS service models.
- I don't believe one needs figures to illustrate yes/no questions or other survey questions with only two answer options. Figures 2 and 3 could thus be removed and this information presented in article text only
- I would avoid pie charts in research articles (see e.g. https://zenodo.org/records/17550233). I suggest you would use column charts instead throughout the paper
- I suggest you state the limitations of your study as the second last paragraph of your discussion section, followed by a concluding paragraph. It is important to acknowledge the limitations of your work, but stating them at the beginning of the discussion can divert the attention from your main results

- After thorough revision of the article, I recommend sending it for professional English language editing

Is the work clearly and accurately presented and does it engage with the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Data stewardship, Information science, scientific communication

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 22 Jul 2026

Giulia Caldoni

We thank the referee for the encouraging feedback and for the thoughtful suggestions. They allowed us to improve the text significantly, and to clarify some aspects that deserved more space and attention. Below, we provide a detailed point-by-point response to each comment. Our responses follow each of the reviewer's comments.

Work title The current manuscript is titled "Mapping data stewardship in Italy: Findings from the first national survey". I understand the use of term "mapping" within this context, but I wonder if the title would work better with search engines without it. Also, the term "first" seems unnecessary in the title. I invite the authors to contemplate whether the title could be shortened into "Data stewardship in Italy: Findings from a national survey".
Authors' response: We thank the reviewer for the observation. We agree that the shorter version of the title is more appropriate, hence we accepted the suggested form "Data stewardship in Italy: Findings from a national survey".
Terminology and RQs The understandability of the article suffers from terminological discrepancies.

- The abstract states the research aim as to “map the presence, roles, and activities of data stewards or equivalent professionals across universities and research institutions, to provide an initial and systematic overview of the national landscape”
- The introduction, however, seems to be built upon the three following core concepts of data stewardship, data stewardship profiles, and data stewardship service models
- The end of the Introduction states that “[...] the idea emerged to launch a public survey to map the state of data stewardship in Italian universities and research-performing institutions. A further objective of the survey was to assess the respondents’ willingness to join a national community of practice on RDM.”
- Furthermore, within the methodology section, the structure of the survey questionnaire is based on “Respondent Profiles” and “Institutional Characterisation”.

Authors’ response: We thank the reviewer for their observations. We agree that the manuscript contained some terminological discrepancies that could affect its readability. To address this, we revised the text to consistently distinguish between data steward role(s) and data stewardship services/service models throughout the manuscript. We also revised the Introduction to improve conceptual coherence and ensure consistency across the Abstract, Introduction, and Methodology sections. This terminological looseness is exacerbated by the fact that the research aim and research questions are not addressed in the article text. Looking at the survey questionnaire, the research aim and questions could perhaps be formulated, for instance, as follows: “The research aim of the present work is to investigate how data stewardship services are organised within Italian universities and research institutions. The research questions are posed as follows.

- RQ1 What kinds of professionals undertake data stewardship in Italian universities and research institutions?
- RQ2 Which service organizations employ the personnel undertaking data stewardship activities?
- RQ3 Which kinds of data stewardship support actions are these professionals currently undertaking?”

Authors’ response: We thank the reviewer for the suggestion. We made RQs explicit in the *Introduction* section. I believe the authors must:

- Provide a general introduction to their research aim and define their core terminology in the Introduction before the literature review segment
- Link the literature review with the core terminology defined in the Introduction (e.g. “The question of what kinds of professionals undertake data stewardship is addressed in the research branch focussing on DS profiles”; note! I am not sure if the latter is factually correct; it is more of a style example)
- Define their research aim and research questions in detail after the literature review segment (see the example RQs above)
- Make sure the methods section is terminologically aligned with the research aim and RQs
- Base the reporting of their survey results around the RQs, e.g., one Findings sub segment per RQ
- Briefly summarize the most important findings per RQ in the Discussion segment

Authors’ response: We thank the reviewer for their observations. We have revised the *Analysis of Survey Results* section to organise more clearly the presentation of the findings around the four RQs, ensuring that each RQ is clearly addressed in a dedicated subsection.

Furthermore, we have revised the *Discussion* section to summarise more explicitly the key findings in relation to each RQ.

Literature review The authors must improve their literature review. The draft only cites 17 works, which is a small amount for a research article. Furthermore, the draft mostly cites EOSC, RDA and national working group reports when of course the literature on data stewardship is more broad. A simple Google Scholar search with the term "data stewardship" produces several relevant works currently missing from cited references. Besides research articles focussing on data stewardship, also literature on how academic libraries have arranged RDM services are in my view relevant to this work. As a rule of thumb, the authors should strive to double the amount of cited references when adding research articles into the list of references.

Authors' response: We thank the reviewer for the suggestion. The literature review has been revised accordingly, and 10 more references were added. Misc. notes and improvement suggestions

- For a better flow of the current Introduction text, I suggest the authors experiment with removal of the subtitles (e.g. "International efforts to professionalise research data stewardship"). Furthermore, the authors could experiment with moving the current 2nd paragraph of introduction as the new 4th paragraph of introduction. The latter solution could improve the text flow introducing the concepts of DS profiles and DS service models.

Authors' response: We thank the reviewer for the observation. We changed the order of the paragraphs as suggested to improve text flow. We decided to maintain the subtitles to identify paragraphs, as we believe it helps readability.

- I don't believe one needs figures to illustrate yes/no questions or other survey questions with only two answer options. Figures 2 and 3 could thus be removed and this information presented in article text only
- I would avoid pie charts in research articles (see e.g. https://zenodo.org/records/17550233). I suggest you would use column charts instead throughout the paper

Authors' response: We thank the reviewer for the observation, shared also by referee 4. We removed figure 2 as suggested, while we maintained figure 3 as we think it's relevant to give a visual snapshot of the number of formally recognized data stewards in the national landscape. We also removed pie charts and used tables instead.

- I suggest you state the limitations of your study as the second last paragraph of your discussion section, followed by a concluding paragraph. It is important to acknowledge the limitations of your work, but stating them at the beginning of the discussion can divert the attention from your main results

Authors' response: We thank the reviewer for the suggestion. We revised the *Discussion* section and reported the limitations of the study at the end of the section.

- After thorough revision of the article, I recommend sending it for professional English language editing

Authors' response: We thank the reviewer for the suggestion. The manuscript went through copyediting to enhance clarity and conciseness.

Competing Interests: No competing interests were disclosed.

Reviewer Report 03 March 2026

https://doi.org/10.21956/openreseurope.24014.r69766

© 2026 Coffey A et al. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aoife Coffey
University College Cork Library, University College Cork National University of Ireland University College Cork Boole Library (Ringgold ID: 494323), Cork, County Cork, Ireland

Pádraig Mac Aodhgáin
University College Cork Library, University College Cork National University of Ireland University College Cork Boole Library (Ringgold ID: 494323), Cork, County Cork, Ireland

This manuscript presents the findings of the first national survey on data stewardship in Italian universities and research institutions. The survey responds to the growing need for research data management support in line with Open Research and the FAIR Principles. The key findings suggest that data stewardship in Italy is not formalised, there is a spectrum of job titles and responsibilities and the backgrounds of those in data stewardship roles varies widely. It concludes that there is a growing presence of data stewardship in Italy but that limited formal recognition of the role are preventing clear career pathways and recognition. Highlighting the need for national coordination to strengthen professional identities and recognition and work towards standardize practices.

Is the work clearly and accurately presented, and does it engage with the current literature?
The work is well-grounded in relevant literature and contextualised within the European research landscape. We encourage the authors revisit the statement *"only 33% of the European respondents actually had the job title "data steward"* The figure mentioned does not seem to be supported by the data underlying the cited report and potentially this needs to be rephrased for accuracy. Data available here https://zenodo.org/records/10117910

Is the study design appropriate and is the work technically sound?
The design is appropriate and socialised in relevant and available channels.

- The survey was reopened for a period could the reason and justification for this be included in the methods.
- Some additional details of the limitations of the survey would be useful such as an acknowledgement of the limitation of using mailing list which are open science focused so could be seen as self-selecting.
- We also suggest including details of how well the 77 responses represent the full Italian research ecosystem – number of eligible institutions vs the number of responses and an estimated national research population.

- Could the authors also clarify how the list of data stewardship activities outline in Figure 7 was complied.

Are sufficient details of methods and analysis provided to allow replication by others?
The methodology is described in detail and is sound and appropriate.
However, there could be some improvements.

- In figure 1 two of the data points beside the number 17 are the same for consistency should 2 be 1.5 times the size of 1. Pie charts for cities with 2-5 respondents were visualised as 1.5 the size of the units.
- Names of major cities could be included in figure 1 for those who are not familiar with the geography of Italy.
- Figure 8 – could the axis be explained perhaps include a legend to clarify what each point on the bar chart means.

Are all the source data and materials underlying the results available?
Yes all data and material have been shared

- The underlying data provides the survey questions in English but could the column headers in the dataset also be translated to English.

If applicable, is the statistical analysis and its interpretation appropriate?
The analysis is descriptive which is appropriate given the profile of the data and sample size.

- It may be helpful to refine some of the wording used in the descriptions of the findings. Softening the phrasing will ensure that the interpretations come across as balanced and precise. An example would be the phrase 'striking majority' which seems a bit of an overstatement.

Are the conclusions drawn adequately supported by the results?
The authors acknowledge the limitations of the study in the conclusion, however there are several instances where the discussion appears to over-interpret or extend the findings beyond what can be supported by the data.

- The conclusions section would benefit from being more concise and from reframing the content to highlight a small number of clear, well-defined take-home messages.
- The conclusion section would also benefit from clarification of the authors definition of DS-like staff, DS- activities, DS-like responsibilities, DS-like professionals as they are mentioned in the discussion but not elsewhere. It would be good to under the meaning of these phases in the context of the survey.
- *"In the future, we plan to perform another survey including as many universities and research institutions as possible."* We suggest removing this as it is should be an aim from the outset.
- We suggest that this statement is revisited *"On the other hand, the large proportion of respondents who answered in a personal capacity may reflect limited institutional awareness/policies related to RDM needs or a lack of internal communication about existing practices."* It is an over interpretation of the findings but could be rephrased as question.
- For the following line it is not clear how this assertion is supported by figure 6 and 7. " *According to the responses that we gathered here, a DS is a professionals with (at least one or more) cross-cutting skills in RDM (disciplinary, IT and technical, legal) and who very often acts as a bridge between researchers (i.e. producers and users of research data), infrastructures and research organisations (Figure 6 & Figure 7)."*
- The suggestion that the list of data stewardship task was *'exhaustive'* seems to be an overreach considering the sample size and the early stage of data stewardship socialization and adoption in Italy in 2023 as described by the author themselves. Also in light of earlier statement such as *"DS-like activities are performed informally or distributed across multiple roles, making them difficult for respondents to identify."*

- While the broad range of backgrounds among data stewards is often viewed as a strength, enabling them to contribute effectively across a diverse research landscape the phrase *"fragmented professional identity*" and the following paragraph seems to imply a negative interpretation of the diverse educational backgrounds. You may want to consider rephrasing.
- It is also not clear that the data indicates that there is a *'growing emphasis'* on the technical dimension of the role given the sample size.
- We suggest the last paragraph in the conclusion is giving the heading/title **Next Steps**

Notes:
There minor typos which need to be corrected and acronyms which should be clarified . Such as
" Consortium GARR"
"a DS is a professionals with"

Is the work clearly and accurately presented and does it engage with the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research data stewardship

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 22 Jul 2026

Giulia Caldoni

We thank the referees for their positive feedbacks and for the thoughtful comments on our work. They allowed us to improve the text significantly, and to clarify some aspects that deserved more space and attention. Below, we provide a detailed point-by-point response to each comment. Our responses follow each of the reviewers' comments.

Is the work clearly and accurately presented, and does it engage with the current literature? The work is well-grounded in relevant literature and contextualised within the European research landscape. We encourage the authors revisit the statement *"only 33% of the European respondents actually had the job title "data steward"* The figure mentioned does not seem to be supported by the data underlying the cited report and potentially this needs to be rephrased for accuracy. Data available here https://zenodo.org/records/10117910
Authors' response: We thank the reviewers for raising this point. We carefully re-examined the underlying data from the *Research Data Alliance Interest Group Professionalising Data Stewardship Career Tracks Survey Dataset*. Among the 205 European respondents who answered this question, 68 reported the job title *Data Steward*, whereas 137 reported other job titles. Thus, individuals with the job title *Data Steward* accounted for approximately 33% of the European respondents (68/205 = 33.2%). Based on this verification, we consider the statement in the Introduction to be accurate and supported by the underlying data. Therefore, no modification to the manuscript was deemed necessary.

Is the study design appropriate and is the work technically sound? The design is appropriate and socialised in relevant and available channels.

- The survey was reopened for a period could the reason and justification for this be included in the methods.
- Some additional details of the limitations of the survey would be useful such as an acknowledgement of the limitation of using mailing list which are open science focused so could be seen as self-selecting.

Authors' response: We thank the reviewers for the helpful comments. To improve clarity, we revised the *Materials and methods* section to provide more details on the reasons for reopening the survey. We acknowledged and discussed potential limitations of the dissemination approach, including the survey being reopened, in the *Discussion* section.

- We also suggest including details of how well the 77 responses represent the full Italian research ecosystem – number of eligible institutions vs the number of responses and an estimated national research population.

Authors' response: We thank the referees for raising this point. To address this point, we revised the text to indicate that responses were obtained from 17 out of the 104 Italian universities and research centers. We explicitly discussed that this underrepresentation may affect the generalisation of the findings. We also clarified that the primary aim of the study is to provide a snapshot of current practices rather than a fully comprehensive or statistically representative national mapping. These additions can be found in the *Discussion* section.

- Could the authors also clarify how the list of data stewardship activities outline in Figure 7 was complied.

Authors' response: We thank the reviewers for the suggestion. We have clarified the manuscript by specifying that the list of data stewardship activities presented in Figure 7 was developed through an iterative design process based on the authors' professional experience and refined using insights from international reports, particularly those from the Netherlands and Denmark. This information has now been explicitly included in the Methods section.

Are sufficient details of methods and analysis provided to allow replication by others? The methodology is described in detail and is sound and appropriate. However, there could

be some improvements.

- In figure 1 two of the data points beside the number 17 are the same for consistency should 2 be 1.5 times the size of 1. Pie charts for cities with 2-5 respondents were visualised as 1.5 the size of the units.
- Names of major cities could be included in figure 1 for those who are not familiar with the geography of Italy.
- Figure 8 – could the axis be explained perhaps include a legend to clarify what each point on the bar chart means.

Authors' response: We thank the reviewers for the suggestions.

- We fine-tuned the pieces of information shown in figure 1 following referees' suggestions. First, we corrected the consistencies of the cities mentioned in the review, according to the logic mentioned in the Materials and Methods section of the manuscript.
- We decided to include labels for all cities to make it easier to identify the respondents' places of origin, even for those unfamiliar with the geography of Italy.
- We converted figure 8 into a table to improve the clarity of data interpretation.

Are all the source data and materials underlying the results available? Yes all data and material have been shared

- The underlying data provides the survey questions in English but could the column headers in the dataset also be translated to English.

Authors' response: We thank the referees for the suggestion. Since the survey was conducted in Italian, the dataset was deposited in its original language to ensure fidelity to the source data. We believe that translating only the column headers, while leaving the response values in Italian, would reduce dataset readability and create ambiguity for data reuse. To support accessibility for non-Italian-speaking users, we have provided an English version of the survey questions in the accompanying documentation. For these reasons, we have retained the original Italian column headers in the archived dataset.

If applicable, is the statistical analysis and its interpretation appropriate? The analysis is descriptive which is appropriate given the profile of the data and sample size.

- It may be helpful to refine some of the wording used in the descriptions of the findings. Softening the phrasing will ensure that the interpretations come across as balanced and precise. An example would be the phrase 'striking majority' which seems a bit of an overstatement.

Authors' response: We thank the reviewers for the observation. To improve the paper readability and the interpretation of the results, we carefully revised the manuscript to ensure the use of more cautious wording and to tone-down expressions that might have overemphasized the results. As an example, we removed the phrase 'striking majority'.

Are the conclusions drawn adequately supported by the results? The authors acknowledge the limitations of the study in the conclusion, however there are several instances where the discussion appears to over-interpret or extend the findings beyond what can be supported by the data.

- The conclusions section would benefit from being more concise and from reframing the content to highlight a small number of clear, well-defined take-home messages.
- The conclusion section would also benefit from clarification of the authors definition

of DS-like staff, DS- activities, DS-like responsibilities, DS-like professionals as they are mentioned in the discussion but not elsewhere. It would be good to under the meaning of these phases in the context of the survey.

Authors' response: We thank the reviewers for the comments. We have carefully revised the *Discussion* section to make it more concise and ensure that interpretations remain consistent with the results. We have softened statements that could be perceived as overinterpretations and aligned the wording more closely with the exploratory scope of the study. We uniformed the terminology, using "DS functions" to what was referred as "DS-like task", "DS-like responsibilities" and "DS-like activities" in the previous version. In the *Introduction section of the* revised text, we have clarified the terminology, where we explicitly state that the term "DS" "DS functions" and "data stewardship" are used in a broad sense to include professionals who perform data stewardship functions. We have also ensured consistent wording throughout the manuscript to avoid ambiguity.

- "In the future, we plan to perform another survey including as many universities and research institutions as possible." We suggest removing this as it is should be an aim from the outset.

Authors' response: We thank the reviewers for the suggestion. We removed this specific sentence and revised the entire Discussions section highlighting the impact that the survey already had in the Italian and international landscape, rather than outlining future aims.

- We suggest that this statement is revisited "On the other hand, the large proportion of respondents who answered in a personal capacity may reflect limited institutional awareness/policies related to RDM needs or a lack of internal communication about existing practices." It is an over interpretation of the findings but could be rephrased as question.

Authors' response: We thank the reviewers for the observation. We revised the entire text to make it more adherent to the results we are presenting, hence we removed this sentence.

- For the following line it is not clear how this assertion is supported by figure 6 and 7. "According to the responses that we gathered here, a DS is a professionals with (at least one or more) cross-cutting skills in RDM (disciplinary, IT and technical, legal) and who very often acts as a bridge between researchers (i.e. producers and users of research data), infrastructures and research organisations (Figure 6 & Figure 7)."

Authors' response: We thank the reviewers for the remark, we removed this sentence as the same concept is better explained in a following paragraph of the *Discussion* section.

- The suggestion that the list of data stewardship task was 'exhaustive' seems to be an overreach considering the sample size and the early stage of data stewardship socialization and adoption in Italy in 2023 as described by the author themselves. Also in light of earlier statement such as "DS-like activities are performed informally or distributed across multiple roles, making them difficult for respondents to identify."

Authors' response: We thank the reviewers for the observation. We revised the sentence, using a more cautious wording, and specified how the list of activities was compiled.

- While the broad range of backgrounds among data stewards is often viewed as a strength, enabling them to contribute effectively across a diverse research landscape the phrase "fragmented professional identity" and the following paragraph seems to imply a negative interpretation of the diverse educational backgrounds. You may want to consider rephrasing.

- It is also not clear that the data indicates that there is a ‘growing emphasis’ on the technical dimension of the role given the sample size.

Author’s response: We thank the reviewers for the observation. We revised the sentences to make our statements clearer.

- We suggest the last paragraph in the conclusion is giving the heading/title Next Steps

Authors’ response: Thank you for the suggestion. We revised the entire Discussions section to clarify our observations, hence we decided to highlight the impact that the survey already had in the Italian and international landscape, rather than to outline possible next steps.
Notes: There minor typos which need to be corrected and acronyms which should be clarified . Such as “ Consortium GARR” "a DS s a professionals with"

Authors’ response: We thank the reviewers for these notes, we carefully checked the text, corrected typos and explained acronyms. “GARR” can’t be explained as it is not an acronym. We provided a link to the Consortium GARR webpage for clarity.

Competing Interests: No competing interests were disclosed.