



Research Data Management Practices Across the Research Data Lifecycle and Their Potential for Collaboration in an International Higher Education Alliance

RESEARCH PAPER

SONJA N. KRALJ 

MATTHIAS LANDWEHR 

**Author affiliations can be found in the back matter of this article*

 ubiquity press

ABSTRACT

Proper research data management facilitates collaboration and promotes research progress through the synergetic process of publishing and reusing research data. Despite these advantages, extant literature reveals that researchers often struggle to manage their research data. In this article, we investigate the relation between research data management practices and collaboration in a context that was made to spark or intensify collaboration—a European Universities alliance. We conducted 39 semi-structured interviews addressing the whole research data lifecycle predominantly with researchers, but also support staff, from six universities that were or were shortly to become part of a European Universities alliance. Our findings show that researchers' practices affect their collaboration in various ways. For example, we find that sharing data within research teams constitutes an obstacle to collaboration due to international differences and technical impediments. Publishing data openly is also associated with certain reservations, such as a lack of knowledge. Furthermore, doubts toward existing data hinder researchers from reusing other researchers' data. International higher education alliances such as European Universities alliance should keep these impediments in mind and offer solutions which foster collaboration.

CORRESPONDING AUTHOR:

Sonja N. Kralj

University of Konstanz,
Germany

sonja.n.kralj@outlook.com

KEYWORDS:

research data management;
Open Science; European
Universities alliance; needs;
open data; collaboration

TO CITE THIS ARTICLE:

Kralj, S.N. and Landwehr, M. 2025 Research Data Management Practices Across the Research Data Lifecycle and Their Potential for Collaboration in an International Higher Education Alliance. *Data Science Journal* 24: 36, pp. 1–13. DOI: <https://doi.org/10.5334/dsj-2025-036>

Research data management (RDM) increases the success prospects of research projects (Briney, 2015; Kanza and Knight, 2022) and is essential for fruitful research collaborations (Borgman, Wallis and Mayernik, 2012; Zuiderwijk, Türk and Brazier, 2024). Recognizing these advantages, the European Commission advocates responsible RDM in line with the FAIR principles (being findable, accessible, interoperable, and reusable; Wilkinson *et al.*, 2016), data deposition in repositories, as well as measures that advance RDM skills as part of its Open Science policy (European Commission, n.d.). What is more, the European Commission and many other funding bodies have made good RDM a requirement for obtaining funding (forschungsdaten.info, 2023). Turning to the education sector, another priority of the European Commission is European Universities alliance established as part of the European Universities initiative, a key element of the European strategy for universities (European Commission Directorate-General for Education, Youth, Sport and Culture, 2024). As a result of its fifth call, 64 alliances involving higher education institutions in 35 European countries are, at the time of writing, forwarding international research collaborations, developing joint strategies, and creating seamless mobility offers for diverse status groups (*ibid.*).

Extant literature addresses challenges regarding RDM in consortia, collaborative projects, and partnerships (Bishop *et al.*, 2020; Jäckel and Lehmann, 2023; Mittal *et al.*, 2023) but has not shed light on European Universities alliance yet. In contrast, there is much literature on RDM needs along the research data lifecycle, which underlines the topic's relevance. Scholars have discovered needs in the planning phase (Bellgard *et al.*, 2023; Lefebvre, Bakhtiari and Spruit, 2020), the collection phase (Chen and Wu, 2017; Gownaris *et al.*, 2022), the processing phase (Anderson *et al.*, 2007; Bishop *et al.*, 2020), the analysis phase (Anderson *et al.*, 2007; Borghi and Van Gulick, 2021), the sharing and publishing phase (Tenopir *et al.*, 2020; Syn and Kim, 2022), the preserving phase (Anderson *et al.*, 2007; Donaldson and Koepke, 2022), as well as the reusing phase (Gownaris *et al.*, 2022; Mathiak *et al.*, 2023). To our knowledge, processing, publishing, and reusing are best researched, while the beginning of the research data lifecycle seems especially underrepresented.

Our study addresses the following research question: how do the RDM practices of researchers at universities, having just joined a European Universities alliance, affect these researchers' future research or RDM-related collaborations? We answer this question by examining researchers' needs across the research data lifecycle. Our study illuminates a wide range of RDM practices and differential starting points for collaboration across the universities, which provides an overview of the potential complexities for institutions considering joining an international higher education network such as a European Universities alliance. Furthermore, our study is beneficial for the European Commission, European University alliances, as well as members of the studied alliance, as they can build on our findings to improve current processes and infrastructure. Besides, our study extends and deepens existing knowledge on researchers' RDM practices and needs across the research data lifecycle. We heed Perrier *et al.*'s (2017) call to study the full research data lifecycle and thus shed light on its less-researched stages.

In the following sections, we will look at the importance of RDM for collaboration and give a brief overview of RDM needs across the research data lifecycle.

LITERATURE OVERVIEW

Collaboration is associated with certain RDM-related obstacles for partners in a consortium, which Schmitt and Burchinal (2011) summarize as a double challenge consisting of RDM coordination and ensuring data integrity. In addition to general obstacles to RDM, such as lack of time and resources, incentives, institutional support, or training (Borghi and Van Gulick, 2021), challenges can arise across the whole research data lifecycle.

Good planning advances successful collaboration (Mittal *et al.*, 2023). According to Lefebvre, Bakhtiari and Spruit (2020), however, RDM planning is still incipient. According to the authors, researchers invest little effort in the data management plan (DMP) and struggle to include consortial partners' information. Collaborators may need to formulate DMPs in more detail to capture each one's approach (Jäckel and Lehmann, 2023). What is more, some researchers do not seem to take DMPs seriously enough (Bellgard *et al.*, 2023), which may bear the risk of

conflicts, miscommunication, and a lack of accountability in collaborations. In terms of data collection, researchers often have institution-specific standard collection procedures, which may be difficult to adapt for different collaborations (Schmitt and Burchinal, 2011). Chen and Wu (2017), in turn, find that researchers lack methods and standards for collecting data, which may diminish their potential for successful contributions in collaborations. Furthermore, it is important to define a common vocabulary so that the partners generate comparable data (Urbano and Cagnacci, 2021).

In the processing phase, challenges arise based on integrating numerous data types collected in collaborative projects, data protection and security issues, as well as discrepancies between institutional and consortial processing workflows, policies, and regulations (Bishop et al., 2020; Mittal et al., 2023; Schmitt and Burchinal, 2011). Many studies imply improvable data processing preconditions for collaboration. For example, biomedical researchers and psychologists struggle with organizing their data, which is aggravated by a lack of standards (Anderson et al., 2007; Borghi and Van Gulick, 2021; Syn and Kim, 2022). Moreover, storage is a complex task in collaborative projects on which researchers lack clarity (Bishop et al., 2020). Researchers also use unfavorable storage options like personal computers (Chen and Wu, 2017; Tenopir et al., 2020), while Reichmann et al. (2021) advocate for more investments in improving institutional storage services. In the data analysis phase, literature identifies a lack of standardized analysis approaches and institutional support, as well as special software needs as critical (Anderson et al., 2007; Mittal et al., 2023).

Sharing data is a crucial element of collaboration, and publishing data may spark new ones. According to Mittal et al. (2023), sharing data with collaborators requires additional effort to fulfill data protection requirements, especially in the case of sensitive data. The authors postulate that a lack of quality control and time decrease sharing even in the case of a central sharing infrastructure. Insufficient metadata may decrease the data's usefulness for others (Mittal et al., 2023; Syn and Kim, 2022). Moreover, a lack of time or rights is another barrier to publishing data (Tenopir et al., 2020) and Mittal et al. (2023) acknowledge that publishing data is still not a standard procedure, as some researchers lack motivation and choosing the repository may be challenging.

Archiving data enhances integrity and reproducibility as a foundation for collaboration. However, researchers are concerned about the long-term availability and stability of their data (Donaldson and Koepke, 2022); they are dissatisfied with available solutions and lack training and assistance (Tenopir et al., 2020). Reusing data published by other researchers fosters a collaborative culture, may spark new partnerships, and saves resources, but findability issues, lack of metadata, and worries about data quality impede reuse (Gownaris et al., 2022; Mathiak et al., 2023). Kim and Yoon (2017) find significant effects of repository availability, perceived usefulness, perceived concern, as well as the availability of internal resources on researchers' data reuse intention.

METHOD

In this section, we describe our data collection and analysis process. Following extant studies using interviews (Bishop et al., 2020; Zuideerwijk, Türk and Brazier, 2024), we designed our study as qualitative to be able to get in-depth insights. Our sample consists of researchers and support staff working at six universities in six countries (see Appendix 1). Five universities were members of the European Reform University Alliance (<https://erua-eui.eu>) in its first funding phase (meaning that this alliance was new), one university was not a member of this alliance yet at the time of data collection but became a member in the alliance's second funding phase (starting shortly after data collection). Even though it was not an official member yet, the sixth university had expressed interest in Open Science and openness toward participation in activities in the time before it became a member, so we deemed it reasonable to include this university. None of the universities had been a member of another European Universities alliance before. According to the guidelines of the University of Konstanz, an ethics vote did not need to be obtained for the study.

We recruited our sample with the help of our board member colleagues at the partner universities. Our sampling was purposive, as we were looking for researchers with different academic backgrounds and experience levels, as well as support staff (e.g., research data managers, librarians). In total, we conducted 39 interviews with 40 informants (one interview was conducted with two informants; see Appendix 2). Among these informants, 31 were researchers and nine were support staff. The average length of an interview was 43 minutes

and 45% of our informants were female. Thirty-one interviews were conducted on-site at the respective universities, while seven were conducted online, and one was conducted in a written way. Two researchers, who were thoroughly briefed, assisted the principal investigator in collecting the data. All on-site and online interviews were recorded with a voice recorder and transcribed verbatim. Subsequently, all interviews were depersonalized by substituting informants' names with pseudonyms or removing project-specific paragraphs.

Our interviews were of a structured nature with two separate but similar guides for researchers and support staff (see Appendix 3). The interviews with researchers started with the latter's general understanding of RDM, followed by the main part, which comprised questions related to all research data lifecycle phases. A picture of the research data lifecycle served as visual support. In each phase, the first question related to actual practices, while the second question addressed needs. At the end, the interviewer asked about critical situations with regard to research data. The interviews with support staff included questions on their involvement and the support they give researchers in each lifecycle phase. We asked which opportunities and threats they see and which RDM elements they identify as critical. The interviews with support staff included questions on the communication with researchers and the infrastructure at their institution, too.

We analyzed the data using QDA Miner. We defined our codes during analysis and derived them from the data (Hsieh and Shannon, 2005), using the lifecycle phases as overarching categories. The process resembles Sendra, Late and Kumpulainen's (2025) work, but with more codes in total. The first author coded all interview transcripts. To ensure reliability and transparency, the coding was independently reviewed by the second author, and any discrepancies were discussed until consensus was reached. Besides the codes included in the findings, topics such as managing expectations and differences between generations and disciplines emerged. Next, we present our results regarding the question of how well-prepared researchers newly working in a European Universities alliance are for future collaborations.

FINDINGS

PLANNING PHASE

In the planning phase, we identify three blocks of needs for collaborating successfully: more awareness, DMPs, and planning other lifecycle phases.

In order for RDM to initiate or support collaboration from the beginning, it needs to be planned thoroughly. Our data show that researchers are sometimes not aware of that. For example, for Victor (assistant professor, information sciences), RDM 'starts from the point that we gather data.' He admits later on, 'actually, yes, you are right. I tend to forget the planning phase.' William (postdoc, social sciences) reckons, 'it's such a counterintuitive field. Because you're looking at your own practice sort of from a side perspective. You don't think about this stuff when you're planning research.' Additionally, some researchers plan projects to avoid RDM by eschewing funding. Opening up a contrast between old and new European Union members, Elizabeth (professor, political science) states, 'Central Europe works slightly different than all the old EU. And we can manage research very frequently without funding, based, for example, on the relations with other institutions and colleagues at other universities.'

Another aspect reflecting a lack of awareness in collaborative projects is the missing role allocation. William (postdoc, social sciences) laments that RDM responsibility is assigned to the project coordinator who may not primarily identify with this role:

When you do have these collaborative projects and you are actually trying to coordinate them ..., the management capacity for this data dimension of the project, it's just integrated into the coordinator role completely or almost completely. We have this advisory function, but there's not a separate person who has that responsibility. ... So in that sense, I mean, there would be some good dynamics in separating those roles.

Another big issue in the planning phase is DMPs. Some researchers are unfamiliar with them, such as Caroline (doctoral candidate, humanities), who has heard of DMPs, but has no experience with them due to a lack of cooperation in project consortia:

We've never actually planned the data management per se. So we never planned how would we store the data, how would we share the data. ... I do believe that in Horizon 2020 such a plan is done. But it was probably made by the leaders and as we are not the leaders, we just follow what they say to us.

This raises the question of whether the answers given by the leaders really suit all partner institutions. A hierarchical approach is unlikely to contribute to sustainable RDM practices and may generate confusion, which may impede collaboration. Others see DMPs as a burden. Nicolas (professor, law) admits, 'we usually do copy and paste. ... We treat it like a tiresome necessity.' In the same vein, Charles (assistant professor, economics) remembers,

The first encounter was that it's a lot of bureaucracy. ... So, I still don't really know what it [DMP] is, to be honest. I know I have some boxes to tick and stuff to say. And I mean, I understand that it's basically making sure that the data is stored properly, but the properly is probably a bit vague.

Moreover, Oscar (associate professor, humanities) identifies the problem that researchers and support staff often misunderstand each other when writing DMPs:

What I find interesting with the law expert is that they can be very picky, but they are able to adapt or to interpret the law according to the situation. And what I encounter in the data advisor is that it's very difficult for them to adapt their protocol to the scientific needs or the researcher's needs. It's the same with engineers who develop tools for researchers. They don't communicate in the same words, they don't have the same language.

Planning other phases of the research data lifecycle emerged as important as well. For example, planning data analysis seems especially relevant in international project consortia. In Christina's (professor, engineering) opinion:

This has to be done from the proposal phase because after you get the project, it's impossible to start, you know, fighting. This is my data, no, this is your data or you do the analysis of ten cities, when you know, I have only the budget, you know, to do one or two cities. So data management is really critical for the success of the projects.

COLLECTION PHASE

Data protection issues are carried over from the planning phase into the data collection phase, where they must be resolved in practice to kick off a successful collaborative collection. Also, existing data sources represent a challenge.

Lucy (assistant professor, humanities) exemplifies a lack of concern with ethical and legal issues. When asked how she dealt with data protection in her interview study, she replied, 'you mean if I had to ask the people to sign a paper? No. ... Maybe because some of these people are people I know. And because they were willing to help me. And I never thought of something to ask.' All other informants are familiar with obtaining consent but often struggle with how to handle it. Oliver (professor, humanities) sees a conflict between respecting privacy and funders' demands:

If I interview a 19-year-old and, okay, this person is very young and I got her consent, but I'm not fully, you know, sure that she's aware of all possible consequences. ... We need to be very, you know, protective about people, but at the same time, I think that in the past, mostly the answer was, okay, we won't publish it because, you know, it's safer this way. But, you know, in terms of science or public funding, it's not the perfect answer.

Louise (support staff, cross-disciplinary) underlines this need:

There are a lot of questions with regard to consent forms. That's developing consent forms. It's more a GDPR [General Data Protection Regulation] issue. So it's not so much that I'm involved in that, but I can see that, their cases. And I discussed that yesterday as well when I gave a course at [university] on those consent forms. If the consent forms are five pages, even the researchers who get that to contribute to a specific project say 'no, I don't want to participate.'

Oscar (associate professor, humanities) laments that the process of obtaining consent is restrictive and may, in the case of doing research across borders, prolong the research process:

For simply interviewing people from organizations, we had to fill different forms that were validated through different commissions. It took months just to do a simple interview. So these ethical aspects are probably very necessary in certain aspects, but in certain projects, research, it's just, I would say, useless.

When it comes to existing data sources, cooperation could help. For instance, for collecting media articles, David (professor, humanities) would like to have 'an overview of the types of databases that I could access. Because it seems we have a rather limited access to international databases from my university.' Monica's (support staff, cross-disciplinary) idea for improvement is collaborative and involves 'some sort of database that could tell new researchers who has already access to the databases they want to get access to', as 'it would be nice for them also to be able to contact other researchers using the same databases.'

PROCESSING AND ANALYZING PHASE

In the processing and analyzing phase, in addition to data protection, storage and organization emerged as the main areas in which researchers have needs to be able to collaborate successfully. The infrastructure for data storage represents one of the biggest needs.

Christina (professor, engineering) and her collaborators are worried about data security:

We are now having the data usually in the cloud of our university. But this is also a big discussion among, you know, different partners. Where should the data be stored? Is it safe to have it only in the cloud, or is it safe to have it in the servers of the university? What will happen if, you know, something, you know, an attack, for example, a cyberattack happened?

She also identifies a possible reason for researchers' preference for their own infrastructure: 'The problem that we have with the funding of the European projects as well as the national projects is that they do not pay for cloud services.' Besides, insufficient institutional support leads researchers in our sample to set up their own storage facilities, like Benjamin (professor, biology):

I don't know how IT understands itself but they are very busy with the general infrastructure. ... So we actually bought two servers for my lab. ... And we also have them in our own room basically (laughs) because putting them into their server room, which would make sure they are climate controlled, that they have, whatever, emergency electricity and so on, was also not without hick-ups.

David (professor, humanities) expresses the need for more funding of the humanities and social sciences. He advocates:

Some sort of capacity building to store data and to collect large numbers of data which is becoming, I mean which is becoming, at least in the digital humanities, more and more the way to go. ... At least in my country, the humanities and social sciences are less funded basically because we don't need equipment. I mean, we're supposed to basically do our research with this (points to laptop), this is even too sophisticated. This should be a pencil, right? And then some deep thought. ... Pencil, deep thought and perhaps chocolate.

Another critical element of the processing phase is depersonalization for safe collaboration. Oliver (professor, humanities) reckons that interdisciplinary experience exchange might help:

Anonymizing data is a challenge. I think that it's not always possible, but I can assume that people with some experience from other disciplines, maybe they have, you know, some new ideas how to do it, some clever tricks, and to do it transparently, ethically, but also, you know, to make those data usable.

Cross-country differences play a role in data protection efforts in international consortia as well, which may be, according to William (postdoc, social sciences), related to national differences in trust levels:

One of our participants is from Hungary and, you know, there's a lot of difference across Europe. ... We sign the things, but that's mainly because it's legal, right. Consent forms, you have to do it. And of course there is some kind of respect that you don't cite people in ways that could get them in trouble or whatever, but it's not a big problem to work with public sectors. But in Hungary it's something quite, quite different. I mean, they are expecting that they will probably have requests for anonymization and depersonalization at a level where the data really begins.

Furthermore, researchers working alone are struggling with organizing their data, especially in the case of large qualitative data. Hannah (postdoc, social sciences) is completely overwhelmed and would appreciate sharing experiences:

I'm realizing that I have so much data. So perhaps I could have had a better way of organizing already in the collect and capture phase because I was kind of relying on the methods that I was using before. And I can just see now, you know, my overview. I'm lost in data, basically. So yeah, well, what I've needed, I'm not sure. But perhaps, you know, just a collection of inspiration. How do other people work with ethnographic fieldwork?

Gabrielle (assistant professor, information science) is worried about the lack of control over how collaborators process the data. She desires transparent procedures, as the process currently entails asking 'how they will process the data ... because you don't know the systems they have, the policies they have, you don't know what kind of people are working, are they specialized or not? Are they going to harm this data or not?' Analyzing data collaboratively is another area researchers have needs in. Oliver (professor, humanities) had the following negative experience when using an analysis software: 'A few people used the same program, I think it was MaxQDA. And we had some problems with, you know, using by the whole team the same document. And it didn't always work the way we expected.'

Gabrielle (assistant professor, information sciences) also sees a need for an integrative tool regarding data protection issues, not just in the analysis phase, but also in earlier lifecycle phases. She laments differing national requirements and the corresponding coordination of partners:

In European projects, things are a bit complicated because there are a lot of partners from different countries with different backgrounds. I mean, now we have the General Data Protection Regulation, the GDPR, but this regulation somehow has been interpreted differently in various EU countries. So in some countries like Germany, there would be stricter requirements than any other EU country. So you have to orchestrate, let's say, the partners that are responsible for the corresponding tools, let's say, or for the processing, for the collection, for the storing, for the processing, for the analysis and so on.

SHARING PHASE

This paragraph will first illuminate practices regarding sharing research data with collaborators, followed by sharing data with the public. Sharing data with people from other (international) institutions constitutes a source of concern for researchers because of different national regulations, worries about data protection, and a need for convenient tools. Referring to a multi-country study, David (professor, humanities) reports, 'the data files are so big that we don't actually have very good ways of sending them. So, either we send our partners smaller files and then some data sharing is okay in my country, but then not okay in theirs.' Kevin (postdoc, health sciences) is not completely happy with his team's practice either:

We send each other download links, which is not very convenient. So yeah, having easier common cloud sharing possibilities would be much better. For internal people it works fine. But with external collaborators, it's really, I would say, very old school.

Frank (professor, chemistry) disapproves of the multitude of sharing platforms:

The main other frustration that I have with data is that in this university and many of the at least the universities in my country use different types. I mean, we have all

these, like, cloud platforms, but it's not easy to share outside of the university. So it would be really great if we would have like a, what's it called, an Eduroam, uh, an edu cloud or whatever.

Kevin (postdoc, health sciences) emphasizes the importance of usability in light of data protection requirements:

I won't download a tool to implement encryption and decryption and then think about how I can share data with collaborators somewhere else who need to analyze part of the data and so on. And this is something, quite frankly, if this is not provided in an easy, accessible way, so one that has little extra work, then I won't do it. It's quite simple.

William (postdoc, social sciences) has witnessed a transformation over the past years, which requires adequate ways of sharing:

I suppose one thing that has changed a lot is the sharing of the raw data. Ten years ago, you would just put it up in some kind of shared cloud folder. Or it wasn't even a cloud then, you know. Maybe a Google Drive or something. You would just put it up there and people would just access it, all partners. I was working in EU projects back then also. So I have witnessed a little bit of that transformation. You would just share it anyway you could really.

Some researchers already benefit from collaborations, which raises hopes that new collaborations within the scope of the European Universities alliance will prove valuable. When talking about where she learned about ways of sharing data in an international project, Belinda (assistant professor, psychology) admits: 'My PI [principal investigator]. She is like a source of information, especially about these, like, more contemporary practices. Because here in my country, I don't have a lot of resources for that.'

In terms of publishing, the main issues are a lack of knowledge, trust, and incentives, as well as licensing, insufficient metadata, and a diversification of infrastructures. Caroline (doctoral candidate, humanities) admits she lacks clarity in different areas of Open Science, including publishing data: 'I'm really completely lost in these ways of making articles, data or other stuff open. I do believe that I need some time with someone who would be patient and ... in an easily to understand way describe how to do it.' Referring to the current societal and political climate, David (professor, humanities) is concerned that some data may, if published, entail negative consequences. In his opinion, 'all the ideas of sharing data exist in a more apolitical space than research actually exists in and I think we'll be existing for the foreseeable future and perhaps in an even more polarized environment.' He raises the point that publishing certain data might endanger researchers:

Some types of data are becoming more and more controversial, which means people might access data, misunderstand it, misuse it, misrepresent it, and then blame you. So, to some extent, making our data available to citizen science is a good idea. But what about people out there whose sole job would be to, sort of, make researchers look stupid and basically make fake research, just like fake news.

On a national level, Stella (professor, social sciences) refers to a lack of trust: 'We do not have a culture of sharing data with different researchers. And so I think this is the issue. And I think it's really related to trust also because maybe we are afraid of somebody.' Furthermore, William (postdoc, social sciences) alludes to the issue of different funding agencies having different requirements, as he views publishing data as an asset in obtaining funding by the European Commission. However, researchers can easily circumvent the requirements, reflecting a lack of incentives:

It's a selling point when you're doing the proposal. ... The Commission has a very sort of culture-shaping role in European research. And then I guess some countries need the money more than others. So that effect is probably stronger in some countries than others. In my country, where people have ample access to other funding, there's a little bit of a (snorts) dismissive attitude sometimes. But I do think, I mean, even if you compare it to national funding agencies, they are really raising the bar in terms of following contemporary standards of ethics and data management. ... It's

a little bit of a secret, but I'm a little bit of a fan of the European Commission in that sense. You shouldn't say that too loud in my country because it's really not a popular viewpoint (laughs).

In addition, two big areas of insecurities which hinder cooperation are licensing and metadata. Kim (support staff, cross-disciplinary) considers licensing 'a big topic. People usually don't know about it. And then sometimes they shy away because they don't know. Once they see the options and get explained in simple language what everything means, then they are ready to go forward with it.' Christina (professor, engineering) mentions interdepartmental misunderstandings which hinder the flow of information:

Data licensing, this would be extremely helpful. We do not have experts. We have IT experts who are supposed to know about data management. ... I think also that we speak a different language. ... We use the same notation, but it means totally (laughing) different things for them and for us.

In terms of metadata, Benjamin (professor, biology) laments some of his colleagues' practices: 'You can literally make it accessible in a way that it is completely useless to the majority of people.' He provides an example:

A GPS location, okay? It sounds trivial but it actually gives you access to satellite data and that gives you access to long-term climate, sunlight, whatever, temperature and so on. Alone from a precise location where the stuff was collected you get tons of data for free out of it associated with. ... Incentivize it. You should show people what you can do with it if they provide it. ... So that's a pretty low bar. Usually, you kind of know what months and year you are in, too. So it's two things that are really simple.

Beliefs about their data's uselessness may hinder researchers' collaboration, too. Working with social scientists, Louise (support staff, cross-disciplinary) has witnessed:

A narrative in the sense that other researchers will not be able to use my data because it's in a specific context, specific period with specific people, and they won't be able to transfer it just. And they can't be FAIR because they're sensitive or they're personal, so we have a lot of information, providing information to do and that's a long-term process, repeating and explaining.

Another issue hindering collaboration is the diversification of repositories instead of using existing ones. Referring to institutional repositories, Benjamin (professor, biology) reckons 'Nobody will know they exist.' He encourages to 'upload it to something that exists already and that allows you to cross-connect it. ... The more there is common repositories for stuff, I think, the easier it will get for people to access it, too.'

ARCHIVING PHASE

Researchers' needs in terms of archiving data are related to tools, data format, and maintenance. With regard to tools, Natalie (doctoral candidate, biology) has a plan for the long-term preservation of her physical samples, but less so for her digital data: 'For all my physical material I have a good plan, but for the virtual, the digital, I am a bit more uncertain because I basically just can't remember where I'm supposed to store it at the moment.' Fiona (support staff, cross-disciplinary) considers sustainable solutions most critical:

Finding the long-term storage solutions. Especially we haven't had that case yet, but if someone really collects big data, we would have to think how to do that. It's not very efficient to have each time the same discussion and, you know, finding solutions for one particular researcher. We would have to come up with one idea.

Stella (professor, social sciences) is concerned about a lack of responsibility after the end of collaborative projects:

I feel it is very sad because I do not feel that we archived correctly our research data. We had some cloud and we shared some information on a cloud, but I am not sure what happens with this information after finishing the research project. And I think that the principal investigators do not feel responsibility for the issue of archiving data. If they didn't have a direct obligation to do it from the institution of funding.

Regarding the data format, updating the data continuously represents a need. Kim (support staff, cross-disciplinary) delineates:

making sure that in ten years' time, you can still open them and use them. And not even the professors from the computer science department have an answer to that. So even they have a problem with 'I tried opening this program that my PhD student did eight years ago but I can't.'

REUSING PHASE

One reason which hinders collaboration in our alliance is that researchers are not reusing others' research data because of worries about their quality. Kevin (postdoc, health sciences) has quite a determined standpoint against reusing data about which or about whose collectors he lacks knowledge:

If I don't know the lab, if I don't know the people, if I can't ask anyone, I don't think I would really use anonymous or data from other labs. ... It's not just the quality. It's also, just to give a small example, often in a dataset, you know, in our field, it's not that the instructions to the subject, for example, are that clean. So if you ask them, what did you tell them, and that can change the behavior of the subjects. Then there are small things. I mean, either they had breaks in between or they didn't have breaks in between recordings that can cause fatigue and all these little details.

Stella (professor, social sciences) is 'wondering about qualitative data. And probably it is more challenging to understand correctly existing qualitative data than quantitative. They are more complex.' Oliver (professor, humanities) reports incentives for collecting instead of reusing data: 'If you prepare your grant proposal, you can get money for new interviews.' Elizabeth (professor, political science), in turn, would like a virtual research environment which supports preceding lifecycle phases, too, but she is not familiar with the term:

The most needed help from my perspective is to have a ... space where I can store data, but at the same time I can get access to the data established by somebody else. ... What I mean, I think that I would like to have a general system, you know like we have academia.edu or ResearchGate, for data. ... So I have a secure place where I can upload my data. I can reuse them, being 100% sure that everything's okay, and I can reuse somebody else's data and provide my data to other people.

DISCUSSION AND RECOMMENDATIONS

In this section, we give recommendations on how RDM may encourage collaboration. Our paper addresses the research question of how the RDM practices and needs of researchers who work at new European Universities alliance member universities affect their future collaborations by providing insights on all research data lifecycle phases. Most researchers in our sample who already do collaborative research perceive it as fruitful in general. While many already engage in beneficial practices, they also have various needs which pertain to the individual (e.g., lack of knowledge), institutional (e.g., inadequate infrastructural support), national (e.g., perceived collective lack of trust), and consortial (e.g., lack of role allocation) levels. It is important to sharpen and define the kinds of collaboration our study is conducive to, as in our case, the first funding phase of the European Reform University Alliance primarily served to explore general possibilities of cooperation.

Many researchers in our sample consider RDM only partially an integral part of the research process (Bellgard *et al.*, 2023). We confirm a disregard of planning and unequal contributions to DMPs reported by extant literature (Lefebvre, Bakhtiari and Spruit 2020). Structures like European Universities alliance should set up DMP meetings for collaborating researchers and encourage project-specific DMPs (Jäckel and Lehmann, 2023), as well as the usage of tools like ARGOS (<https://argos.openaire.eu/home>) or Data Stewardship Wizard (Pergl *et al.*, 2019). This would enhance and simplify contribution by all partners. Our findings also convey the risk that researchers avoid DMPs by applying for specific funding and escape the radar if they do not have international partners. Collaborations, which are, in our case, sparked by the new membership in the European Universities alliance, may thus entail that more researchers

apply RDM. Clarity in DMPs may also solve the problem of dealing with inconsistencies in data protection regulations in many lifecycle stages by setting clear guidelines.

Another important topic which emerged from our data and which goes hand in hand with collaboration is exchange. Our data show that peer exchange can be powerful and possibly of higher value to researchers than help which comes solely from support staff. We encourage formats which foster experience exchange among collaborators, but also beyond—for instance, through an ambassador program (Kralj, Heber and Helmecke, 2025). Besides, not just researchers but also support staff of collaborating universities should exchange best practices. Furthermore, rather than possibly discrepant vocabulary between collaboration partners (Urbano and Cagnacci, 2021), our informants report intrainstitutional misunderstandings (e.g., between researchers and the IT department). Institutions should find ways to overcome this issue—for instance, by encouraging regular exchange and needs assessment.

Collaboration can hardly be enforced, but diligently, ethically, and soundly designed collaborative tools could, after clarifying potential licensing requirements, address the reported needs in integrating, storing, sharing, and jointly analyzing data while at the same time saving resources. Examples are research infrastructures or platforms like CLARIAH (<https://clariah.nl/>), Coscine (<https://about.coscine.de/en/>), EOSC (<https://open-science-cloud.ec.europa.eu/>), or virtual research environments like openBIS (Plass et al., 2023). Collaboration should point researchers to adequate existing infrastructure or tools. From an institutional perspective, while university and/or government policies are a valuable framework, our study finds that they are yet to become an important influencing factor. Policies may regulate important parameters and thus provide a common basis on which collaboration can take place more effectively, which is why further research is needed on their specific effect on collaboration.

As to limitations, there may be a light bias in our sample, as members of support staff (mostly our board members) recruited the interviewees at our partner universities. There might possibly be cases in which the researchers who were contacted had already been in touch with our board members. We also have a slight majority of humanities and social sciences scholars, since those are the core fields of many European Reform University Alliance universities. As we only provide a temporal snapshot, longitudinal research is needed to study changes in practices and needs over time. Future research could also accompany and evaluate formats such as ambassador or mentoring programs.

CONCLUSION

To sum up, our study provides rich insights into how the RDM practices and needs of researchers from different disciplines and career levels who work at six universities in a European Universities alliance prepare them for future collaborations. We show that researchers have various needs and that many needs appear across the participating universities. Often, such RDM needs come in the way of successful research collaborations. Therefore, it is crucial to address these needs at the university, country, and European levels. Overall, RDM is a topic to which European Universities alliance can make a valuable contribution—for instance, by fostering formats of experience exchange on an alliance or even cross-alliance level.

DATA ACCESSIBILITY STATEMENT

The dataset is available via KonDATA, the repository of the University of Konstanz, at <https://doi.org/10.48606/tfsvgw4qvt5b97f9>. The published dataset contains 35 interviews with 36 out of a total of 40 informants. Three informants refused consent to their transcripts being published on KonDATA, while one person did not respond to the email which included the depersonalized transcript, asking for permission.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendices.** Appendix 1 to 3. DOI: <https://doi.org/10.5334/dsj-2025-036.s1>

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 101035808. This paper is based on the deliverable 'Research Data Management Community Study Report' available on the CORDIS platform.


COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

S. Kralj collected most of the data, processed and analyzed the data and wrote the article, while M. Landwehr provided the topic, gave feedback throughout the project, and edited the paper.

AUTHOR AFFILIATIONS

Sonja N. Kralj  orcid.org/0000-0002-8516-8034
University of Konstanz, Germany

Matthias Landwehr  orcid.org/0000-0001-9274-2578
University of Konstanz, Germany

REFERENCES

- Anderson, N.R., Lee, E.S., Brockenbrough, J.S., Minie, M.E., Fuller, S., Brinkley, J. and Tarczy-Hornoch, P.** (2007) 'Issues in biomedical research data management and analysis: Needs and barriers', *Journal of the American Medical Informatics Association*, 14(4), pp. 478–488. Available at: <https://doi.org/10.1197/jamia.M2114>
- Bellgard, M.I., Bennett, R., Wyborn, Y., Williams, C., Barner, L. and Zeps, N.** (2023) 'RDM+PM checklist: Towards a measure of your institution's preparedness for the effective planning of research data management', *Data Science Journal*, 22(1), p. 36. Available at: <https://doi.org/10.5334/dsj-2023-036>
- Bishop, B., Gunderman, H., Davis, R., Lee, T., Howard, R., Samors, R., Murphy, F. and Ungvari, J.** (2020) 'Data curation profiling to assess data management training needs and practices to inform a toolkit', *Data Science Journal*, 19(4), pp. 1–8. Available at: <https://doi.org/10.5334/dsj-2020-004>
- Borghi, J.A. and Van Gulick, A.E.** (2021) 'Data management and sharing: Practices and perceptions of psychology researchers', *PLoS ONE*, 16(5), p. e0252047. Available at: <https://doi.org/10.1371/journal.pone.0252047>
- Borgman, C.L., Wallis, J.C. and Mayernik, M.S.** (2012) 'Who's got the data? Interdependencies in science and technology collaborations', *Computer Supported Cooperative Work*, 21(6), pp. 485–523. Available at: <https://doi.org/10.1007/s10606-012-9169-z>
- Briney, K.** (2015) *Data management for researchers: Organize, maintain and share your data for research success*. Exeter: Pelagic Publishing. Available at: <https://doi.org/10.2307/jj.28833749>
- Chen, X. and Wu, M.** (2017) 'Survey on the needs for chemistry research data management and sharing', *Journal of Academic Librarianship*, 43(4), pp. 346–353. Available at: <http://doi.org/10.1016/j.acalib.2017.06.006>
- Donaldson, D.R. and Koepke, J.W.** (2022) 'A focus groups study on data sharing and research data management', *Scientific Data*, 9, 345. Available at: <https://doi.org/10.1038/s41597-022-01428-w>
- European Commission.** (n.d.) *Open science*. Available at: https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science_en (Accessed: 6 March 2025).
- European Commission Directorate-General for Education, Youth, Sport and Culture.** (2024) *European universities: A key pillar of the European education area*. Available at: <https://data.europa.eu/doi/10.2766/521265> (Accessed: 6 March 2025).
- Forschungsdaten.info.** (2023) *Funder guidelines*. Available at: <https://forschungsdaten.info/praxis-kompakt/english-pages/funder-guidelines/> (Accessed: 6 March 2025).
- Gownaris, N.J., Vermeir, K., Bittner, M.I., Gunawardena, L., Kaur-Ghumaan, S., Lepenies, R., Ntsefong, G.N. and Zakari, I.S.** (2022) 'Barriers to full participation in the open science life cycle among early career researchers', *Data Science Journal*, 21(1), p. 2. Available at: <https://doi.org/10.5334/dsj-2022-002>
- Hsieh, H. and Shannon, S.E.** (2005) 'Three approaches to qualitative health research', *Qualitative Health Research*, 15(9), pp. 1277–1288. Available at: <https://doi.org/10.1177/1049732305276687>

- Jäckel, D. and Lehmann, A. (2023) 'Benefits and challenges: Data management plans in two collaborative projects', *Data Science Journal*, 22, p. 25. Available at: <https://doi.org/10.5334/dsj-2023-025>
- Kanza, S. and Knight, N.J. (2022) 'Behind every great research project is great data management', *BMC Research Notes*, 15, p. 20. Available at: <https://doi.org/10.1186/s13104-022-05908-5>
- Kim, Y. and Yoon, A. (2017) 'Scientists' data reuse behaviors: A multilevel analysis', *Journal of the Association for Information Science and Technology*, 68(12), pp. 2709–2719. Available at: <https://doi.org/10.1002/asi.23892>
- Kralj, S.N., Heber, M. and Helmecke, P. (2025) 'Extending awareness of open science via peers: The ERUA open science ambassador programme', *LIBER Quarterly*, 35(1), pp. 1–26. Available at: <https://doi.org/10.53377/lq.22303>
- Lefebvre, A., Bakhtiari, B. and Spruit, M. (2020) 'Exploring research data management planning challenges in practice', *it - Information Technology*, 62(1), pp. 29–37. Available at: <https://doi.org/10.1515/itit-2019-0029>
- Mathiak, B., Juty N., Bardi, A., Colomb, J. and Kraker, P. (2023) 'What are researchers' needs in data discovery? Analysis and ranking of a large-scale collection of crowdsourced use cases', *Data Science Journal*, 22, p. 3. Available at: <https://doi.org/10.5334/dsj2023-003>
- Mittal, D., Mease, R., Kuner, T., Flor, H., Kuner, R. and Andoh, J. (2023) 'Data management strategy for a collaborative research center', *GigaScience*, 12, p. gjad049. Available at: <https://doi.org/10.1093/gigascience/gjad049>
- Pergl, R., Hooft, R., Suchánek, M., Knaisl, V. and Slifka, J. (2019), "'Data Stewardship Wizard": A tool bringing together researchers, data stewards, and data experts around data management planning', *Data Science Journal*, 18, p. 59. Available at: <https://doi.org/10.5334/dsj-2019-059>
- Perrier, L., Blondal, E., Ayala, A.P., Dearborn, D., Kenny, T., Lightfoot, D., Reka, R., Thuna, M., Trimble, L. and MacDonald, H. (2017) 'Research data management in academic institutions: A scoping review', *PLoS ONE*, 12(5), p. e0178261. Available at: <https://doi.org/10.1371/journal.pone.0178261>
- Plass, F., Englisch, S., Apele Zubiri, B., Pflug, L., Spiecker, E. and Stingl, M. (2023) 'Using OpenBIS as virtual research environment: An ELN-LIMS open-source database tool as a framework within the CRC 1411 Design of Particulate Products', *Data Science Journal*, 22, p. 44. Available at: <https://doi.org/10.5334/dsj-2023-044>
- Reichmann, S., Klebel, T., Hasani-Mavriqi, I. and Ross-Hellauer, T. (2021) 'Between administration and research: Understanding data management practices in an institutional context', *Journal of the Association for Information Science and Technology*, 72(11), pp. 1415–1431. Available at: <https://doi.org/10.1002/asi.24492>
- Schmitt, C.P. and Burchinal, M. (2011) 'Data management practices for collaborative research', *Frontiers in Psychiatry*, 2, p. 47. Available at: <https://doi.org/10.3389/fpsy.2011.00047>
- Sendra, A., Late, E. and Kumpulainen, S. (2025) 'From data lifecycle to research activity model: Research data management in data-intensive social sciences and humanities research', *Aslib Journal of Information Management*, ahead of print. Available at: <https://doi.org/10.1108/AJIM-12-2024-0959>
- Syn, S.Y. and Kim, S. (2022) 'Characterizing the research data management practices of NIH biomedical researchers indicates the need for better support at laboratory level', *Health Information and Libraries Journal*, 39(4), pp. 347–356. Available at: <https://doi.org/10.1111/hir.12433>
- Tenopir, C., Rice, N.M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R. and Sandusky, R.J. (2020) 'Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide', *PLoS ONE*, 15(3), p. e0229003. Available at: <https://doi.org/10.1371/journal.pone.0229003>
- Urbano, F. and Cagnacci, F. (2021) 'Data management and sharing for collaborative science: Lessons learnt from the Euromammals initiative', *Frontiers in Ecology and Evolution*, 9, p. 727023. Available at: <https://doi.org/10.3389/fevo.2021.727023>
- Wilkinson, M.D. et al. (2016) 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, 3, p. 160018. Available at: <https://doi.org/10.1038/sdata.2016.18>
- Zuiderwijk, A., Türk, B.O. and Brazier, F. (2024) 'Identifying the most important facilitators of open research data sharing and reuse in epidemiology: A mixed-methods study', *PLoS ONE*, 19(2), p. e0297969. Available at: <https://doi.org/10.1371/journal.pone.0297969>

TO CITE THIS ARTICLE:

Kralj, S.N. and Landwehr, M. 2025 Research Data Management Practices Across the Research Data Lifecycle and Their Potential for Collaboration in an International Higher Education Alliance. *Data Science Journal* 24: 36, pp. 1–13. DOI: <https://doi.org/10.5334/dsj-2025-036>

Submitted: 02 May 2025

Accepted: 20 November 2025

Published: 04 December 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.