

Practical Guide for Responsible Research Data Management in Academic-Industry Collaboration

Gerald Kremer^{1*}, Juliane Balder¹, Helena Ebel¹, Zhongtian Wang¹, Anvit Upadhyay¹, Eric Kempkens¹,
Sahar Ben Hassine¹, Rainer Stark¹

¹TU Berlin, Institut für Industrielle Informationstechnik, 10587 Berlin

*Corresponding author. E-Mail address: kremer@tu-berlin.de

Abstract: Responsible research data management (RDM) is crucial in industry-university collaborations, yet existing frameworks and data management plan (DMP) tools often remain abstract and detached from daily project work. This paper presents a role-based Business Process Model and Notation (BPMN) process map and an interactive prototype that depict the RDM lifecycle from project start to retention across key academic and industrial roles. RDM activities, phases, and roles were derived from lifecycle models, institutional policies, and project experience, then refined in a collaborative modelling workshop. Evaluation with nine researchers using think-aloud walkthroughs and a questionnaire shows that the artefact clarifies responsibilities and actions, while highlighting tensions between completeness and simplicity.

Keywords:

Research data management, RDM, Academic-industry collaboration, Data lifecycle, Process-oriented guidance

1 Motivation and Research Objective

Research Data Management (RDM) is a cornerstone of Data Science (DS) and Artificial Intelligence (AI) research, enabling reproducibility, reuse, and trustworthy insights from real world applications. In the domain of product creation, especially where academia and industry jointly innovate, data curation, a key aspect of RDM, that involves activities from e.g. data capture to storage and access, is crucial [1]. The work presented in the paper focusses on how to curate and manage research data responsibly in academic-industry collaborations, which brings unique challenges. Academics must balance open-science practices with proprietary and confidentiality constraints, ensure data quality, and comply with industry regulations. The relevance of the topic is underscored by the Open Science principles pushing for FAIR (Findable, Accessible, Interoperable, Reusable) data as a basis for know-ledge reuse [2]. At the same time industry partners must safeguard proprietary know-how and potential intellectual property (IP) embodied in datasets, which constrains sharing and reuse, unless protective measures (e.g., contractual restrictions, or technical de-risking such as anonymization/synthetic or algorithmically transformed data or encryptions) are in place.

Guidance through existing RDM frameworks is textual, abstract and often complicated. If researchers read the complete frameworks, they would learn what they should do, but not how these requirements translate into concrete actions for specific roles (e.g. project PI, research assistant, industry partner contact) at specific points in time. Translating high-level RDM principles into day-to-day practice remains difficult. Also prior literature identifies several challenges impeding RDM in practice, including misalignment between research workflows and data management planning, limited resources or incentives for RDM, and concerns about data governance [3].

This led to the research objective to develop a solution that provides easy-to-understand knowledge for and guides relevant stakeholders through the RDM life cycle of academic-industry research projects. Based on the objective, two research questions were derived:

RQ1: How can the RDM lifecycle academic-industry collaborations be formalized into an actionable model that integrates existing frameworks, standards and regulations and experiential knowledge of researchers to provide relevant and understandable knowledge to relevant stakeholders?

RQ2: How do researchers perceive the usefulness of such a developed solution?

2 State of Research

2.1 Conceptual frameworks and lifecycle models for RDM

RDM refers to the set of organizational and technical practices that are needed to handle research data and is typically framed through high-level principles and lifecycle models. The most influential are the FAIR Guiding Principles, which define that data should be *Findable, Accessible, Interoperable and Reusable* and emphasize machine-actionable metadata and persistent identifiers as requirements for reuse [2]. Building on these principles, different lifecycle models describe RDM as a sequence of phases, from planning and collecting to processing, analyzing, preserving and sharing data. The Digital Curation Centre (DCC) introduced a life cycle model, that distinguishes activities in “full life cycle actions” (e.g. representation, preservation planning, community watch), “sequential actions” (e.g. conceptualize, select, ingest) and “occasional actions” (e.g. dispose, migrate) emphasizes that curation is continuous rather than linear [4]. Other lifecycle frameworks, such as the Research data life cycle by the University College London define similar stages and activities but have a stronger university perspective [5]. Meanwhile there are a multitude of different models and frameworks. In a recently published review nearly 80 frameworks from different disciplines and from 20 different countries were analyzed, indicating that there is no need to develop yet another framework [6]. Instead, other institutions like the National Institute of Standards and Technology rather harmonize existing and established frameworks to work on a common ground instead of building niche solutions [7].

While these models give a structural overview of RDM, they remain largely conceptual. They describe existing phases and their activities, but they say little about how individual roles act and interact, nor how concrete decisions about data are made in practice across a project. A data management plan (DMP) is a core RDM instrument that systematically outlines the procedures for collecting, storing, and managing data. Still, writing a DMP is often created when project proposals are handed in at funding agencies, but are rarely used during the projects themselves [8]. When researching on how to write DMPs, even the well-renowned DCC does not provide easy-to-understand knowledge on their website but instead provides a lengthy text (14 DINA4 pages) with 30 internal links to subpages and 10 references to be read. So, in summary there is no need for new frameworks, but better knowledge provision for researchers.

2.2 Policy and regulatory landscape in Germany

The conceptual frameworks are reinforced by a policy landscape at funder and institutional level. In Europe, Science Europe’s Practical Guide to the International Alignment of RDM provides core requirements for DMPs, criteria for selecting trustworthy repositories and a DMP evaluation rubric [2]. Many European funders and institutions have adopted these core requirements, which has led to a de facto standard for questions that should be addressed in a DMP (e.g. data types, documentation, storage, legal/ethical issues, long-term preservation).

In Germany, the DFG Code of Conduct “Guidelines for Safeguarding Good Research Practice” explicitly anchors research data as part of good scientific practice, requiring appropriate documentation, storage, and accessibility in line with legal and ethical constraints [9]. Similar national or institutional codes across Europe increasingly tie RDM explicitly to research integrity and to open science.

Additionally, data protection and privacy regulations such as the EU General Data Protection Regulation (GDPR) set strict conditions for processing and sharing personal and research data. In academic-industry collaborations, IP protection forms an additional governance layer beyond GDPR. Even when personal data is not involved, datasets can reveal product designs, process parameters, or competitive advantages. Practically, this requires early clarification of ownership and rights to use results and data, what can be published and under which license, and which de-risking strategies are acceptable (e.g., controlled access, synthetic data, or algorithmic transformations that preserve utility of data while at the same time reduce reverse-engineering risks).

2.3 Scientific tools and services for supporting RDM

There are different concepts that address how to support researchers during RDM. Much of it focuses on tools for planning and documenting RDM, particularly DMPs. These include templates and checklists (e.g. DCC DMP checklists) and machine-actionable DMPs (maDMPs) that can be interpreted and acted upon by information systems. At the same time RDM planning tools have been developed. For example the Data Stewardship Wizard (DSW) provides a knowledge-model-driven system where data stewards adapt questionnaires for specific domains and projects, generating DMPs [10]. Another example is the Research Data Management Organiser (RDMO), which enables institutions to model project-specific RDM requirements and generate structured plans and documentation [11]. All identified tools had in common that they focus on creating DMPs via templates or questionnaires and thereby primarily support documentation instead of providing day-to-day procedural guidance. In parallel, the German NFDI research groups (e.g., NFDI4Ing in engineering) provide complementary infrastructure services that operationalize FAIR practices in project work. The NFDI subproject RESODATE for example helps researchers find and reuse existing research artefacts such as software through semantic metadata [12]. However, no projects in the NFDI landscape could be found that focus on the procedural guidance on responsibilities and timing across roles.

2.4 Provision of RDM knowledge and training formats

Beyond tools, many universities librarians and data stewards have developed courses, workshops, and handbooks to support researchers in adopting practices [13]. Studies indicate, however, that researchers often experience RDM guidelines and policies as long, abstract and difficult to map to their day-to-day workflows. [14]. Training interventions can improve awareness and self-reported competence, but their effect is mostly short-term, as detailed procedural know-how is difficult to retain without continuous usage. [15].

Some universities also offer decision-support interfaces for RDM in the form of decision trees or visual checklists. In these tools, researchers answer a small set of yes/no questions and are then guided to more detailed information, services or templates. Such visual checklists help to break RDM down into a series of manageable tasks instead of presenting it as a single, large obligation [16].

More formal process-oriented approaches have also been proposed. In the humanities, BPMN-based models have been used to describe consultation and support workflows for RDM, including typical processes such as support with proposal preparation and data archiving, tailored to the needs of humanities projects rather than covering an entire cross-disciplinary lifecycle [17]. In addition, RDM maturity models provide generic process charts and classify activities into maturity levels to assess and

compare the state of RDM in organizations or projects. The specific needs of stakeholders are only weakly represented [18].

Overall, existing RDM resources provide strong conceptual and compliance-oriented guidance (lifecycle descriptions, policies, DMP templates), but they insufficiently operationalize this knowledge into role- and phase-specific actions that can be applied ad hoc in everyday academic–industry project work, including handovers, governance decisions, and the concrete artefacts to be produced.

3 Development Approach

Therefore, the research gap is the lack of a reusable, role-based and interactive representation of the RDM lifecycle for academic–industry collaborations that translates high-level principles and policies into concrete actions, decisions, and artefacts across roles.

To address this gap, a modeling workshop was conducted with eight researchers with experience in industrial IT as well as grand funded research projects that jointly reconstructed typical collaboration and data flows and co-created a role-based BPMN model. Based on these results an interactive prototype was developed. It presents the RDM lifecycle and roles and reveals short textual guidance and example resources when process elements are selected. Afterwards the prototype was evaluated in individual think-aloud walkthroughs with nine other researchers at different career stages and with different experience levels with industry and research projects, combined with a brief questionnaire on perceived usefulness. The qualitative feedback was analyzed and used to revise both the BPMN model and the interactive prototype content.

4 Developed Artefact

4.1 Structure of the role-based BPMN model

The artefact consists of a BPMN 2.0 process model that represents the RDM lifecycle in industry–university collaborations from a role-based perspective. BPMN was chosen because it is a widely adopted standard for process modelling with a reasonably intuitive graphical notation and formal underpinnings, and is frequently used as a communication medium between heterogeneous stakeholders [19]. Empirical studies have shown that visual process models can support comprehension and task performance compared to purely textual descriptions. Although the effect depends on model quality and user characteristics [20], it is plausible that researchers and industry stakeholders can navigate BPMN-style maps with limited onboarding, especially when supported by legends and drill-down explanations.

Horizontally, the model is structured along a generic RDM lifecycle that integrates common phases from established frameworks like planning, data collection, transfer (which is especially important for industry collaboration and often not found in frameworks), processing, sharing and retention. This makes the model compatible with typical lifecycle descriptions while emphasizing the temporal order of decisions and tasks. Horizontally the chart has different swim lanes, but unlike in typical BPMN models the swim lanes here do not reflect different departments and their responsibilities. Instead, the swim lanes show the data lifecycle phase, the tasks, artefacts and possible resources. Instead of showing responsibilities in the swim lanes the model uses color-coding of activities for key roles in industry–university projects, for example doctoral researcher, principal investigator, IT / infrastructure, and industry partner representatives. Tasks and gateways indicate which actor is responsible for which RDM-relevant actions and decisions at a given phase, including handovers and points where consultation with support staff or the use of additional documents is recommended. Data-related artefacts (e.g. raw data, processed data, deletion contracts, Non-Disclosure Agreements, DMPs, Audit Documents) are represented as BPMN data objects linked to tasks. This makes the connection between

actions and RDM artefacts explicit and helps to disentangle where specific documents or datasets are created, updated or archived. The model intentionally avoids discipline-specific terminology so that it can be reused across different industrial projects, while still allowing local tailoring at the level of examples and annotations. A screenshot from the phase “data storage” can be found in fig. 1.

In this view, the tasks in the data storage phase are presented as a horizontally arranged process, including checks for data integrity, sensitivity reassessment, storage system selection, encryption, catalogue updates and periodic audits. Underneath the process flow, the artefact row lists the typical documents created or updated at each step (e.g. RDM catalogue entry, deletion contract, updated DMP, audit document). A color legend at the bottom maps tasks to roles and indicates whether tasks are mandatory or optional. Presenting RDM in such phase-level maps helps to decompose RDM into small, completable sub-tasks rather than a single, overwhelming obligation. The model is designed to be “high-level but concrete”: it abstracts from project-specific details yet is precise about who does what, when, and with which artefacts. This aligns with findings from process modelling research that emphasize the need for models that balance expressiveness with cognitive simplicity to remain understandable for non-expert users [20].

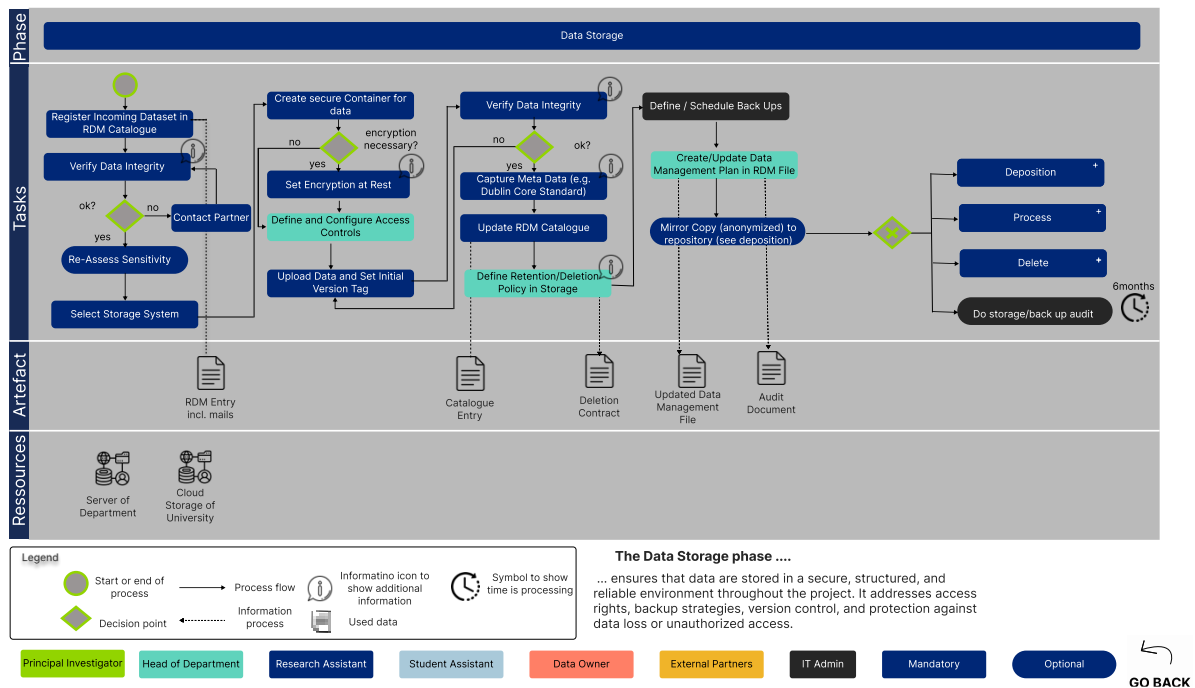


Figure 1: Overview of the role-based BPMN process map for the phase “data storage”

4.2 Interaction concept and demonstrator

Based on these BPMN diagrams, an interactive prototype in the form of a clickable demonstrator was developed that turns the process map into a navigable guideline for everyday project work. The primary intended users are early-career researchers who execute many RDM tasks but lack procedural knowledge, and industry contacts who require transparency about data handling. The demonstrator is not meant to automate RDM. Instead it supports coordination and decision-making by linking each task to role-specific actions, artefacts, and used methods. The full version can be found under [21]. Instead of working in a modelling tool, researchers navigate RDM via screens that show the phase-level process maps and allow them to click on individual activities or decision points. The intention is that users consult the prototype ad hoc at the point of need. For example, when they are about to store a new dataset and are unsure which steps are required and how these steps are fulfilled.

In such a scenario, a research assistant opens the Data storage view (Fig. 2) and follows the tasks from left to right until reaching “Set encryption at rest”. Here, “encryption at rest” means encrypting stored data on storage systems (including backups) so that unauthorized access to stored data does not reveal the content. In case the assistant does not know how to set the encryption at rest, they can click on the task. Clicking this task opens a drill-down screen (Fig. 3). The drill-down provides more detailed, structured guidance. A task panel lists concrete actions (e.g. determining the scope of encryption, configuring backups, establishing key management), a “Quality gates / Goals” panel summarizes what must be true for the step to be considered complete, and additional panels list helpful tools and relevant regulatory aspects (such as security-by-design requirements or data processing agreements for cloud storage). In this way, the prototype keeps the process context visible while offering focused, step-by-step instructions only when needed.

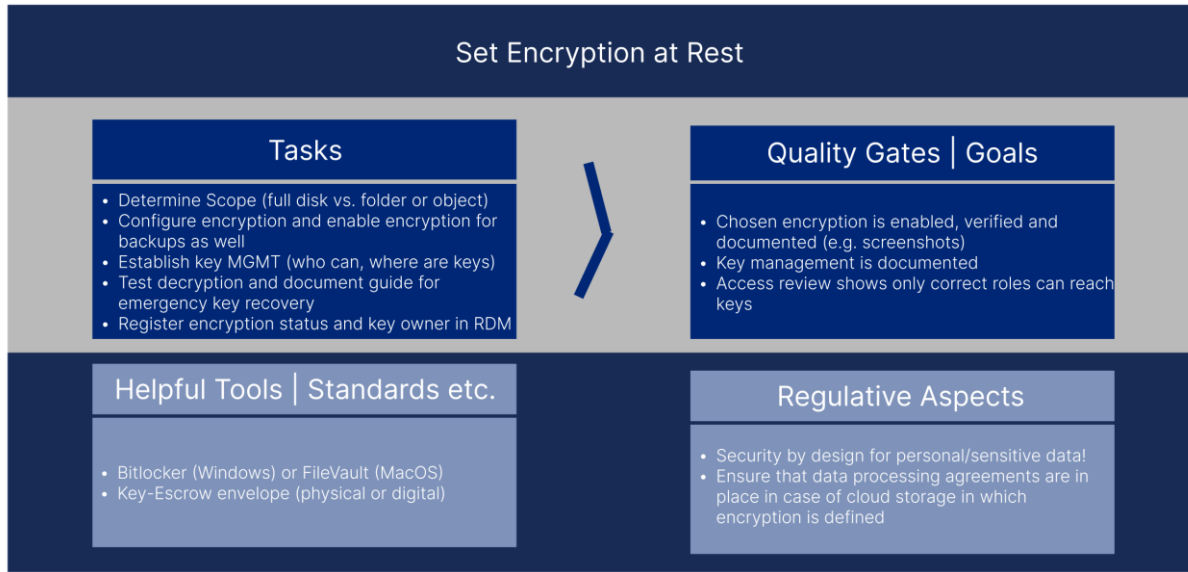


Figure 2: Detail page about the task “Set encryption at rest”

Taken together, the BPMN model and demonstrator form a didactic artefact that combines an overview of the RDM lifecycle with role-specific, on-demand explanations. In contrast to static diagrams or purely textual guidelines, the artefact is intended to be used repeatedly during a project as a navigable “map” of RDM activities.

5 Evaluation

Nine researchers participated in the evaluation (7 doctoral researchers, 1 postdoc, 1 professor; 6 male, 3 female). Their research experience ranged from 0.5 to 22 years (mean 6.5, median 5), and they had completed 1 to 50 industry-collaboration projects (mean 11, median 5). Self-rated RDM competence (1–7 scale) ranged from 2 to 6 (mean 3.6, median 3). Overall, the sample combines early-career researchers with low RDM confidence and highly experienced researchers, enabling assessment across different seniority levels.

5.1 Perceived usefulness

Fig. 4 presents the perceived usefulness. The items were measured on 7-point Likert scales, following the Technology Acceptance Model [22], where perceived usefulness is a core determinant of technology acceptance. Across items, medians cluster between *rather agree* (5) and *agree* (6) for most positive statements. Participants tended to agree that the guideline helps them clarify roles and responsibilities, derive concrete actions, and that overall, the guideline is useful for their work (medians around 5, with

upper quartiles at 6–7). Ratings for time-related benefits (e.g. creating an RDM plan or preparing audits more quickly) are more mixed, with central tendencies around the neutral to “rather agree” range, indicating that participants primarily see the artefact as a conceptual and coordinative aid rather than as a strong time-saving tool. The negatively framed item (“the guide has no added value for me”) received low ratings with a median clearly below neutral, confirming that participants generally perceived the process map as beneficial. The raw data can be found under [23].

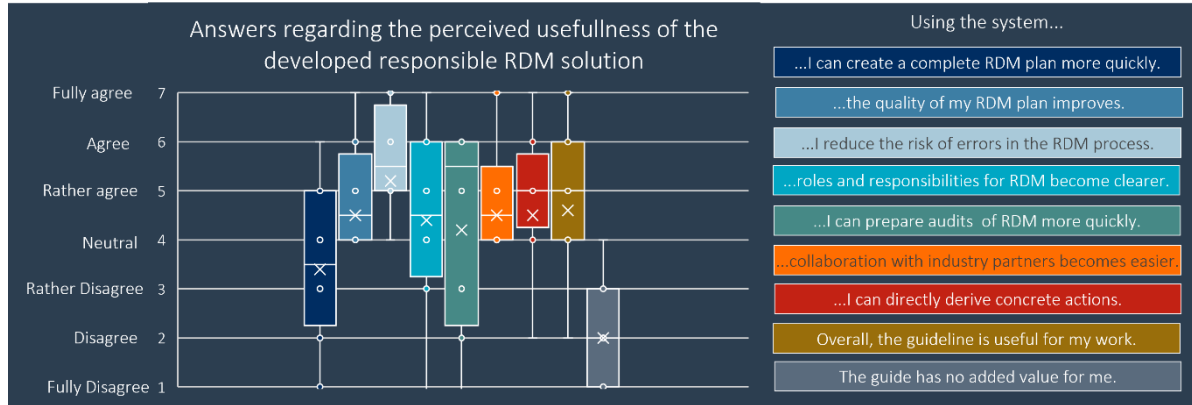


Figure 3: Boxplots of ratings for perceived usefulness and related statements

5.2 Qualitative feedback from think-aloud walkthroughs

During the think-aloud walkthroughs, participants applied the prototype to given industrial collaboration project scenarios in groups of two researchers while verbalizing their thoughts. The recordings were transcribed, anonymized and analyzed. The transcribed recordings can be found under [24].

5.2.1 Positive perceptions

Across groups, the BPMN guideline was described as a systematic way of approaching RDM. One group summarized “I think this is a systematic system, a systematic way of proceeding ...” and emphasized its value for communication with partners: “I find it useful that we can demonstrate how we deal with the data to our industry partners.” The walkthroughs made the complexity of RDM visible in a way participants considered helpful: “You notice how complex the process actually is. You could see that while going through it.” Another participant concluded: “Overall this process is very good so that you don’t forget anything when you carry out research data management.” Another group confirmed this impression and highlighted that the governance and task descriptions make aspects like metadata and anonymization more explicit than their usual practice: “It looks like what we have discussed but more precise and from the side of data ... we did not discuss metadata and anonymization in so much detail. With such a task description it is more visible what steps are behind our steps.” They also noted that the governance view essentially covers the whole project duration: “This is exactly a guideline what we can do from the beginning of a project till three months afterwards.” For projects where partners explicitly value structured RDM, the guideline was described as “very valuable”.

5.2.2 Critical feedback

At the same time, participants pointed to several weaknesses. First, the purpose and scope were not fully clear: “It was not 100 percent clear what the goal was, whether the aim was the complete execution of RDM ... or whether you are only creating the RDM file, and RDM plan.” Second, several terms and decision points were criticized as unclear. Examples include “use case” (“the term ... is always rather irritating for me”), “data deposition” (“Data Retention is something different. We both apparently have not understood ‘Data Deposition’ yet.”), and labels such as “check requirements of using technological solutions for needed data”, which one participant said they did not understand. Gateways in particular

were perceived as confusing: “No matter what I do, I always end up at ‘Derive needed data’. Then this is not really a gateway. That is confusing.” “In general, the gateways are described wrongly.” Third, some participants felt the process was too detailed and prescriptive: “I think you are making it too complicated ... It should not be too complicated, not too many possibilities that could be, but not everything prescriptive.”. They warned that “too much determinism” can create mistrust and make people feel like “slaves of the process”. Fourth, several usability issues of the prototype were mentioned. These include font size (“The font size is too small.”), unclear starting point and missing back navigation (“I would now think that you have to click somewhere on the left first, but that doesn’t work.”; “There is no back button somehow.”), and confusing color coding for roles (“For me this is all the same color ... as someone with red–green weakness.”). The nature of the “RDM file” artefact was also unclear: “What is RDM file? Is it a text file, an XML or JSON or what?”. Another group additionally noted confusion about how mandatory vs. optional tasks are represented: “Because it is color coded, but here it is not color coded anymore ... ‘mandatory, optional’ – and here the research assistant seems to be mandatory and optional at the same time.”. Together, these statements show that while the overall idea is appreciated, the semantics of elements and the current interaction design are not self-explanatory.

5.2.3 Immediate design changes

Based on the feedback, a set of three low-effort improvements were implemented, which are already reflected in the version presented in this paper. First, a short introductory panel was added that explicitly states the goal of the tool (provision of knowledge about specific tasks of RDM plan during all phases of project lifecycle, not automating all RDM tasks), to address the confusion about scope. Secondly, unclear labels (including “use case”, “check requirements of using tech solutions...”) were rephrased and several gateways whose paths previously all led to the same activity were simplified, to avoid pseudo-decisions. Thirdly, a concise legend was added explaining the meaning of colors, shapes and the distinction between mandatory and optional tasks, and reworked the color scheme to be more easily distinguishable, also for users with red–green weakness. Additionally, explanation boxes for a better understanding for the meaning of each phase were added.

6 Discussion and Resulting Improvements

The evaluation indicates that the role-based BPMN demonstrator is generally perceived as useful across a heterogeneous group of researchers with varying experience in RDM and industry collaboration. Most participants reported that the prototype helps them gain an overview of the RDM life cycle, clarify responsibilities, and derive concrete actions. Time savings and direct support for writing funder-specific DMPs were rated slightly lower, which is consistent with our design focus on understanding procedural knowledge rather than automating documentation. These findings fit well with evidence that visual process models can effectively convey procedural knowledge when they are well-designed. A systematic review on process model comprehension emphasizes that graphical process models support users in organizing and integrating procedural information, provided that cognitive load is controlled through appropriate notation and layout [25]. Eye-tracking results show that readers of BPMN models use stable visual strategies and that well-structured models support efficient navigation, especially for non-experts [26]. The participants’ comments that the demonstrator “gives a good overview” and makes “roles and responsibilities clearer” are in line with these observations. In other words, the prototype appears to exploit the strengths of BPMN as a medium for externalizing and sharing process knowledge rather than merely prescribing compliance steps.

At the same time, the results highlight a tension that is also well known in the process-modeling literature: some participants asked for more detail and additional gateways or decision points, whereas others perceived parts of the model as too detailed or complex [25]. The partly contradictory feedback

therefore seems less like a flaw of the prototype and more like evidence that different user groups need different levels of abstraction. This supports the design decision to use drill-down views and suggests that further personalization by role, experience level, or project type would be beneficial.

The broader literature on DMPs also helps to contextualize the findings. Smale et al. argue that while DMPs are widely mandated, evidence for their effectiveness in changing data practices is limited, and researchers often experience them as an administrative burden [27]. A recent survey similarly reports inconsistent integration of DMPs into research workflows and calls for additional mechanisms other than through DMPs to teach good data management practices [28]. Against this backdrop, it is notable that the participants valued the demonstrator not primarily as a way to produce DMP text, but as a means to understand what should happen in each phase and who is responsible. This aligns with the idea that effective RDM support must go beyond plan documents and instead embed RDM into day-to-day project activities.

At the same time, the study is more exploratory than many of the controlled evaluations in the RDM training literature. The perceived usefulness was evaluated and a qualitative feedback was collected, but it was not tested whether the demonstrator leads to measurable improvements in knowledge, skills, or behavior, as done in studies by Xu et al. on online RDM instruction [15]. The small sample ($n = 9$) limits the generalizability of the findings. Future work should therefore include controlled comparisons with existing RDM support formats (e.g., PDF guidelines, DMP tools, or standard workshops), larger and more diverse participant groups, and behavioral outcome measures such as the quality of actual RDM plans or adherence to lifecycle steps during a project.

7 Conclusion and Future work

This paper introduced a role-based BPMN process map and an interactive demonstrator for research data management in industry–university collaborations. It addressed the two research questions:

RQ1: We developed a role-based BPMN process map and an interactive demonstrator for research data management in industry-university collaborations, focusing on procedural understanding (who does what, when, and with which artefacts) rather than on DMP text generation.

RQ2: In an evaluation with nine researchers, the artefact was perceived as useful mainly for clarifying roles and responsibilities, deriving concrete actions, and communicating the RDM “big picture”. It was less associated with time savings for writing DMPs, consistent with its purpose as a complement to existing planning and automation tools.

Still, the current model and demonstrator should be seen as a starting point for a trans-university shared knowledge provision tool. They are meant to be adapted and extended by other institutions, especially those with limited resources to develop their own RDM process models.

References

- [1] K. Briney, H. Coates, and A. Goben, “Foundational Practices of Research Data Management,” *Res. Ideas Outcomes*, vol. 6, p. e56508, Jul. 2020, doi: 10.3897/rio.6.e56508
- [2] Science Europe, “Practical Guide to the International Alignment of Research Data Management - Extended Edition,” Jan. 2021, doi: 10.5281/ZENODO.4915861.
- [3] G. Birkbeck, T. Nagle, and D. Sammon, “Challenges in research data management practices: a literature analysis,” *J. Decis. Syst.*, vol. 31, no. sup1, pp. 153–167, Dec. 2022, doi: 10.1080/12460125.2022.2074653.
- [4] S. Higgins, “The DCC Curation Lifecycle Model,” *Int. J. Digit. Curation*, vol. 3, no. 1, pp. 134–140, Dec. 2008, doi: 10.2218/ijdc.v3i1.48.
- [5] J. A. J. Wilson, *UCL Research Data Lifecycle 2020/2023*. University College London, 2024, p. 661695 Bytes. doi: 10.5522/04/25471591.

- [6] J. Jiang, D. Maurici-Pollock, and R. Tang, "Research Data Lifecycle (RDLC): An Investigation into the Disciplinary Focus, Use Cases, Creator Backgrounds, Stages and Shapes of RDLC Models," *Int. J. Digit. Curation*, vol. 19, no. 1, p. 30, Feb. 2025, doi: 10.2218/ijdc.v19i1.860.
- [7] R. Hanisch, D. Kaiser, A. Yuan, A. Medina-Smith, B. Carroll, and E. Campo, "NIST Research Data Framework (RDaF): Version 2.0," National Institute of Standards and Technology, Gaithersburg, MD, NIST SP 1500-18r2, Feb. 2024. doi: 10.6028/NIST.SP.1500-18r2.
- [8] T. Miksa, J. Cardoso, and J. Borbinha, "Framing the scope of the common data model for machine-actionable Data Management Plans," in *2018 IEEE International Conference on Big Data (Big Data)*, Seattle, WA, USA: IEEE, Dec. 2018, pp. 2733–2742. doi: 10.1109/BigData.2018.8622618.
- [9] Deutsche Forschungsgemeinschaft, "Guidelines for Safeguarding Good Research Practice. Code of Conduct," Apr. 2022, doi: 10.5281/ZENODO.6472827.
- [10] R. Pergl, R. Hooft, M. Suchánek, V. Knaisl, and J. Slifka, "'Data Stewardship Wizard': A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning," *Data Sci. J.*, vol. 18, p. 59, Dec. 2019, doi: 10.5334/dsj-2019-059.
- [11] I. Anders et al., "The Research Data Management Organiser (RDMO) – a Strong Community Behind an Established Software for DMPs and Much More," *Data Sci. J.*, vol. 23, p. 28, May 2024, doi: 10.5334/dsj-2024-028.
- [12] R. Pueblas, P. Kuckertz, J. M. Weinand, L. Kotzur, and D. Stolten, "ETHOS.PASSION: An open-source workflow for rooftop photovoltaic potential assessments from satellite imagery," *Sol. Energy*, vol. 265, p. 112094, Nov. 2023, doi: 10.1016/j.solener.2023.112094.
- [13] Z. Xu, "Research Data Management Training in Academic Libraries: A Scoping Review," *J. Librariansh. Sch. Commun.*, vol. 10, no. 1, Jul. 2022, doi: 10.31274/jlsc.13700.
- [14] T. Austin, K. Bei, T. Efthymiadis, and E. Koumoulos, "Lessons Learnt from Engineering Science Projects Participating in the Horizon 2020 Open Research Data Pilot," *Data*, vol. 6, no. 9, p. 96, Sep. 2021, doi: 10.3390/data6090096.
- [15] Z. Xu, X. Zhou, A. Kogut, and M. Clough, "Effect of online research data management instruction on social science graduate students' RDM skills," *Libr. Inf. Sci. Res.*, vol. 44, no. 4, p. 101190, Oct. 2022, doi: 10.1016/j.lisr.2022.101190.
- [16] University of Edinburgh Library, "Research Data Service. (o.J.). Quick Guide 10 – Research Data Management Flowchart [PDF]. University of Edinburgh. Available at: <https://library.ed.ac.uk/research-support/research-data-service/guidance> (accessed 1 December 2025)."
- [17] P. Helling, J. Blumtritt, and B. Mathiak, "Der Beratungsworkflow des Data Center for the Humanities (DCH) an der Universität zu Köln," *O-Bib Offene Bibl. Herausgeber VDB*, 2018, doi: 10.5282/O-BIB/2018H4S248-261.
- [18] M. L. Wawer et al., "A maturity based data management integration in engineering research projects," *Proc. Des. Soc.*, vol. 5, pp. 169–178, Aug. 2025, doi: 10.1017/pds.2025.10031.
- [19] H. Ritchi, M. Jans, J. Mendling, and H. A. Reijers, "The Influence of Business Process Representation on Performance of Different Task Types," *J. Inf. Syst.*, vol. 34, no. 1, pp. 167–194, Mar. 2020, doi: 10.2308/isis-52385.
- [20] S. A. Kassim et al., "Benefits and limitations of business process model notation in modelling patient healthcare trajectory: a scoping review protocol," *BMJ Open*, vol. 12, no. 5, p. e060357, May 2022, doi: 10.1136/bmjopen-2021-060357.
- [21] G. Kremer, "Practical Guide: Responsible Research Data Management (RDM) - BPMN-based top-level process and phase protocols." Technische Universität Berlin, Jan. 22, 2026. doi: 10.14279/DEPOSITONCE-25164.
- [22] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Q.*, vol. 13, no. 3, pp. 319–340, 1989, doi: 10.2307/249008.
- [23] G. Kremer, "Perceived usefulness ratings (7-point Likert) for an interactive BPMN-based Research Data Management guideline," Jan. 21, 2026, *Technische Universität Berlin*. doi: 10.14279/DEPOSITONCE-25162.
- [24] G. Kremer, "Think-aloud walkthrough transcripts: interactive BPMN-based Research Data Management guideline (academic-industry collaboration)," Jan. 21, 2026, *Technische Universität Berlin*. doi: 10.14279/DEPOSITONCE-25161.
- [25] K. Figl, "Comprehension of Procedural Visual Business Process Models: A Literature Review," *Bus. Inf. Syst. Eng.*, vol. 59, no. 1, pp. 41–67, Feb. 2017, doi: 10.1007/s12599-016-0460-2.
- [26] M. Tallon et al., "Comprehension of business process models: Insight into cognitive strategies via eye tracking," *Expert Syst. Appl.*, vol. 136, pp. 145–158, Dec. 2019, doi: 10.1016/j.eswa.2019.06.032.
- [27] N. A. Smale, K. Unsworth, G. Denyer, E. Magatova, and D. Barr, "A Review of the History, Advocacy and Efficacy of Data Management Plans," *Int. J. Digit. Curation*, vol. 15, no. 1, p. 30, 2020, doi: 10.2218/ijdc.v15i1.525.
- [28] M. Langseth, E. Sellers, G. Donovan, and A. Liford, "Assessing the value and usage of data management planning and data management plans within the U.S. Geological Survey: U.S. Geological Survey Open-File Report 2023–1069, 44 p., <https://doi.org/10.3133/ofr20231069>," Open-File Report, 2023.