Peer Community Journal

Section: Archaeology

Conference: CAA2025

Research article

Published
2026-07-24

Cite as
Steffen Strohm, Mattis thor Straten, Andrea Göhring, Hartwig Buenning, Peer Kröger, Oliver Nakoinz, Matthias Renz and Christoph Rinne (2026) *Bridging Interdisciplinary Research Data Management and Data Science through Modular Research Processes*, Peer Community Journal, 6: e69.

Correspondence
sts@cs.uni-kiel.de

Peer-review
Peer reviewed and recommended by PCI Archaeology, https://doi.org/10.24072/pci.archaeo.100686

CC BY

This article is licensed under the Creative Commons Attribution 4.0 License.

Bridging Interdisciplinary Research Data Management and Data Science through Modular Research Processes

Steffen Strohm [1,2], Mattis thor Straten [2], Andrea Göhring [1,3], Hartwig Buenning [1,2], Peer Kröger [1,2], Oliver Nakoinz [1,4], Matthias Renz [1,2], and Christoph Rinne [1,4]

Volume **6** (2026), article e69

https://doi.org/10.24072/pcjournal.755

Abstract

Transparent and reproducible research in large interdisciplinary projects, such as the Cluster of Excellence ROOTS, depends on combining two perspectives that are usually considered separately: the stewardship concerns of Research Data Management (RDM) and the analytical concerns of Data Science. This paper introduces Modular Research Processes (MRP), a framework representing a research endeavour as distinct, referable products and procedures described at several layers of granularity, expressing resolution rather than a hierarchy of authority. MRP provides a shared representation in which RDM and Data Science converge, focusing on the point where (materialised) products are FAIR-ready, while their procedural counterparts are not (yet) – a discrepancy that motivates the development of systematic descriptions and, where applicable, recommendations. We organise the framework's value into three application fields: illustration and communication; documentation and organisation; as well as comparison and recommendation. Our examinations and activities are based on the Data Management and Data Science Platform (DMDSP), which is an organisational component of the ROOTS project. We present the DMDSP as an example of how such a bridge can be shaped, aligned with an existing ecosystem, filled with collaborative method development, and sustained through training. Beyond the platform, it is MRP that is the transferable contribution, with a claim to cross-disciplinary and cross-project application. We report on two archaeological analyses that have been reconstructed as first proofs of concept, techniques for early empirical feedback on the conditions for adoption of an MRP-related tool, and a research roadmap anchored in this beginning.

[1]Cluster of Excellence ROOTS, Kiel University - Kiel, Germany, [2]Department of Computer Science, Kiel University - Kiel, Germany, [3]Leibniz Laboratory for Radiometric Dating and Stable Isotope Research, Kiel University - Kiel, Germany, [4]Institute of Prehistoric and Protohistoric Archaeology, Kiel University - Kiel, Germany

Peer Community Journal is a member of the Centre Mersenne for Open Scientific Publishing
http://www.centre-mersenne.org/

e-ISSN 2804-3871

Introduction

Large-scale interdisciplinary research projects produce heterogeneous data and, just as importantly, various methods, distributed across disciplines that rarely share standards, tools, or vocabularies. Making such research transparent, reproducible, and reusable requires bridging two perspectives that are usually pursued in isolation: the stewardship perspective of Research Data Management (RDM), which treats data as a sustainable and FAIR (Wilkinson et al., 2016) resource, and the analytical perspective of Data Science (DS), concerned with processes turning that data into insights. In an archaeology-focused setting such as the DFG-funded Cluster of Excellence ROOTS[1] at Kiel University, the gap between the two is especially visible: data and methods originate in archaeology and various other disciplins from natural sciences and humanities, and neither travel easily across disciplinary or paradigmatic boundaries.

In its effort to address this, ROOTS established the Data Management and Data Science Platform (DMDSP) as a dedicated organisational unit (see Figure 1). At its conceptual core lies the Modular Research Processes (MRP) framework, which represents an individual research endeavour as reusable products and procedures described at several layers of granularity. MRP provides the shared representation in which RDM and DS concerns meet. We distinguish the two deliberately: the DMDSP is a situated unit, specific to the ROOTS environment, whereas MRP is the transferable contribution, with a claim to cross-disciplinary and cross-project application.

The contribution of this paper is a transferable account of how to bridge RDM and DS in interdisciplinary research. We introduce MRP at the level required to understand its role within the platform, reduced to modules and layers of granularity, alongside a three-field use-case typology, pairing illustration and communication, documentation and organisation, and comparison and recommendation, through which the framework aims to prove useful to the researchers involved. Both are grounded in the DMDSP, which serves as a case study showing how such a bridge can be organisationally supported through research, services, and training. The DMDSP itself is not presented as a globally deployable platform.

The remainder of this paper is structured as follows. After briefly introducing ROOTS and the DMDSP, *Background and Related Work* situates MRP between selected process reference models it builds on and the general-purpose notations from which it deliberately departs. The section *Core Concept: Modular Research Processes* introduces MRP and the use-case typology in the detail needed for the platform. Then *The DMDSP as a Situated Case* presents the platform in its project and contextual environments, describing how it fills the RDM-DS bridge through interdisciplinary data management, alignment with the existing ecosystem, scientific method development, and training. The *Research Roadmap* lays out the research that holds the platform together, and finally the paper concludes with a summary of our findings and reflections.

Interdisciplinary Research Project ROOTS

The Cluster of Excellence ROOTS, established in 2019, investigates the dynamic interactions between human societies and their environments over time. It brings together researchers from archaeology, the natural sciences, the humanities, and the life sciences to explore how cultural practices, environmental factors, and social structures have been connected in past contexts. The research framework is organised into several subclusters, each addressing a central theme (environmental hazards, dietary transformations, innovation and cognition, urbanisation processes, social inequality, as well as conflict and reconciliation) complemented by archaeological and historical laboratories that integrate empirical methods and theoretical perspectives. With about 25 Principal Investigators involved in the successful proposal for the project's second phase (2026–2032), ROOTS combines its members' multidisciplinary expertise to address transdisciplinary research questions, linking insights about the past to environmental

[1] https://gepris.dfg.de/gepris/projekt/390870439?language=en – retrieved 2026-07-13.

and societal challenges of the present and near future.[2] Besides the thematic subclusters, additional components shape ROOTS as a whole, including the Data Management and Data Science Platform (DMDSP).

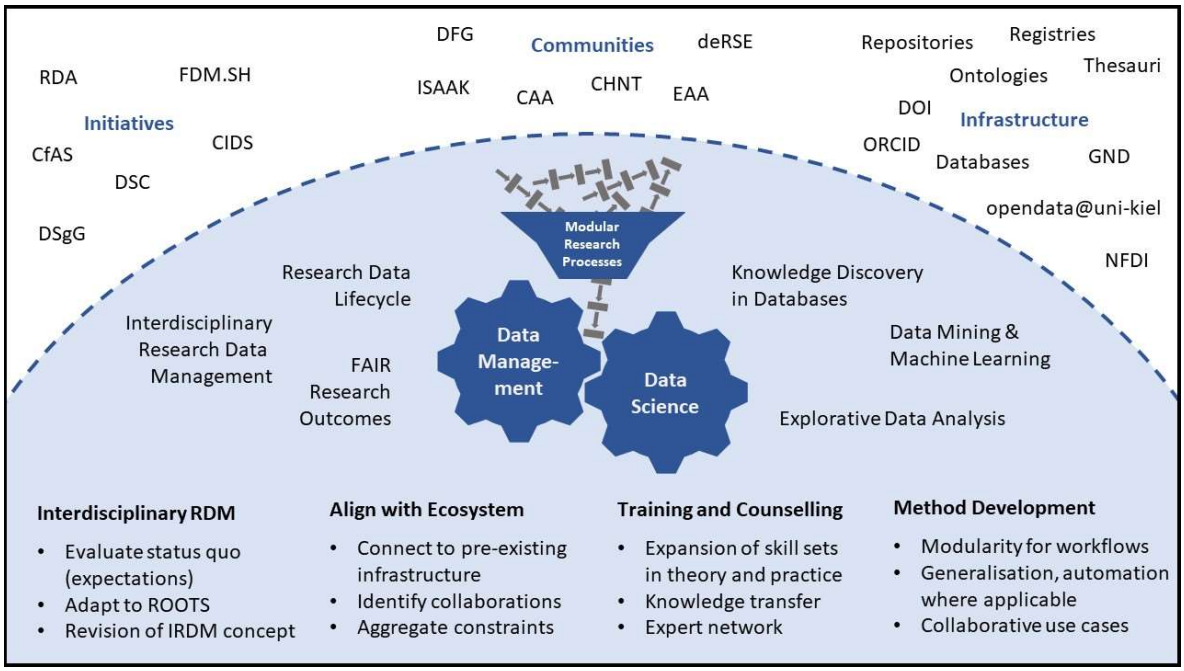

Figure 1 – Overview of the Data Management and Data Science Platform (DMDSP), situated within the surrounding landscape of initiatives, communities, and infrastructure, including the core concept of Modular Research Processes and the planned fields of activity.

Goals of the DMDSP

The goals of the DMDSP are aligned with the principles of sustainable, transparent, and interdisciplinary research within ROOTS. Its primary objective is the organisational and technical implementation of the FAIR principles, ensuring that research outcomes remain consistently retrievable and reusable over the long term. This implementation applies not only to research data in the narrow sense but also to workflows and, consequently, to research software (Barker et al., 2022). At the same time, the DMDSP is tailored to the specific needs of ROOTS, recognising the diverse formats, standards, and practices arising from its constituent disciplines. Beyond RDM, the platform fosters collaboration and provides concrete outcomes in the form of data management tools for the analysis, visualisation, and integration of heterogeneous datasets, promoting integrative workflows that combine machine learning techniques with established domain-specific methods and structured human input. Such input may come from disciplinary experts but also – where appropriate – from broader public engagement, aligning with open science values.

Background and Related Work

The MRP framework builds on two established perspectives on the research process rather than proposing one from scratch, and it deliberately positions itself as a lightweight alternative to general-purpose process notations. From RDM it inherits the Research Data Lifecycle, and from Data Science the Knowledge Discovery in Databases process model.

The Research Data Lifecycle (RDLC) has become a central model for structuring the dynamic processes of research in data management and curation, highlighting the stages through which

[2] https://www.uni-kiel.de/en/cluster-roots - retrieved 2026-07-13.

research data repeatedly pass: planning, collection, processing, analysis, preservation, sharing, and eventual reuse (Cox and Tam, 2018; UK Data Service, 2022). Its cyclic form underscores sustainability, transparency, and reproducibility, and emphasises the iterative nature of research, in which data and methods are revisited, revised, and repurposed.

The Knowledge Discovery in Databases (KDD) model, by contrast, is highly influential in Data Science and describes the systematic extraction of meaningful patterns from data through stages such as selection, pre-processing, transformation, mining, and interpretation (Fayyad et al., 1996). Focused on computational analysis, it pays comparatively little attention to longterm stewardship, versioning, and the contextual challenges that RDM foregrounds.

MRP bridges the two: it embeds lifecycle thinking into the analytically oriented KDD paradigm, so that a single research endeavour can be viewed as a modular composition of products and procedures that is at once computationally substantive and responsibly managed.

Representing a process as a sequence of modular steps is, of course, not new. General-purpose notations such as the Business Process Model and Notation (BPMN) and the Unified Modeling Language (UML) (Object Management Group, 2014, 2017) offer rich and precise constructs for control flow, routing, and execution. MRP is intentionally far less expressive. It is not meant to specify a process for automated enactment, but to make a methodical approach legible and reusable – and it trades the expressive power of these notations for what matters more in everyday interdisciplinary practice: low overhead and approachability for researchers who are not modelling specialists. What MRP adds in return is a dimension absent from such flat notations, describing the same process at several layers of granularity, from research intent down to technical details of implementation.

Core Concept: Modular Research Processes

Products, Procedures, and Layers of Granularity

Building on the two traditions introduced above, the Modular Research Processes (MRP) framework views a single, self-contained research endeavour as a modular composition of products and procedures. A product is an artefact of the research process (e.g., raw data, a cleaned dataset, a figure, a piece of code, a paragraph of interpretation, a publication) while a procedure is an activity that transforms input products into output products. Within the scope of this work, we assume that research endeavours typically involve the following stages: data acquisition, pre-processing, computational analysis, expert evaluation, and publication. MRP treats each step involved in these stages as a procedure together with the products it consumes and yields, and remains agnostic about their number, their repetition, and the cycles of refinement that connect them.

The framework's distinctive feature is that products and procedures may be described at four layers of granularity, which express resolution rather than authority: Vision (the research intent what is being done, and why), Approach (every task and problem class involved; by which means it is pursued), Algorithm (the specific method and its specification; how exactly it is done), and Implementation (the concrete realisation in code or another artefact). The same endeavour can thus be shown coarsely or finely, and – crucially in an interdisciplinary setting – a domain expert reading at the level of intent and a data scientist reading at the level of code engage the same structure at the resolution each one might need in a certain communicative situation.

This layering is also where MRP makes the relationship between RDM and Data Science precise. A product becomes materialised when its content exists as an actual artefact, which at the Implementation layer offers exactly the two anchors FAIR requires: a metadata record for findability and reuse, and a concrete referent to which a persistent identifier can attach. Products are therefore FAIR-ready once materialised. The MRP framework does not enforce FAIRness, but renders its preconditions structurally explicit. Procedures, their methodological counterpart, have no comparable anchor: a FAIR treatment of methods, with shared identifiers and description standards, remains an open problem. It is precisely this gap that motivates the platform's method-development activities and the sketched search space.

Because research is iterative, an MRP description is not static. Successive states of a product or procedure can be retained as versions, and the revisions and discarded alternatives that conventional reporting tends to hide can be preserved as part of the record rather than overwritten. This temporal expressivity (versioning and the documentation of a changing approach) is part of what makes the representation useful beyond a single snapshot, and we return to its consequences for persistence and tooling in the upcoming sections.

Three Fields of Application

An MRP representation is always put to use in a concrete situation, and the DMDSP organises its services around three such fields, distinguished by where a representation comes from and what it is for. The three are not mutually exclusive: a single endeavour may migrate between them as access to artefacts, intent, and collaborators changes. Also, they recur throughout this paper, both as platform activities and as strands of the research roadmap.

Illustration and Communication. An existing publication or codebase serves as the starting point, and a representation is reconstructed retrospectively from it to make a methodical approach and the choices behind it legible to others. Because the source artefacts already exist elsewhere, this field is deliberately storage-free: the representation links to and abstracts over artefacts it does not itself hold. This is the field in which MRP's communicative purpose is most direct and the one most readily demonstrated: a published analysis can be reconstructed and represented at a chosen resolution for cross-disciplinary exchange, teaching or review.

Documentation and Organisation. Here the representation grows prospectively, alongside a research process in action, as a companion while a question and its answer take shape. Its purpose is to organise the endeavour beyond the level of digital storage: products link to real artefacts (e.g., local folders, repositories, reference managers) and the structure of steps becomes the primary order imposed on the work. This is the field in which persistence becomes a genuine requirement. Whether researchers would actually adopt such a companion, and under what conditions, is an empirical question that the platform has already begun to investigate.

Comparison and Recommendation. The unit of interest is no longer a single endeavour but a population of MRP representations, read across projects, domains, and disciplinary cultures. On the analytical side, this supports the comparison of alternative methods and the epistemic cultures in which they are embedded; on the forward-looking side, an unspecified procedure between two specified products becomes a query for candidate methods. Realising this field requires a searchable space of methodological descriptions – the Methods and Algorithms Search Space (MASS) and an associated knowledge graph – which we name here only to bound the present scope and pursue as a separate line of future research.

The DMDSP as a Situated Case

The remainder of the paper turns from the framework to its organisational realisation. The DMDSP is a concrete unit within ROOTS, and its task is to fill the bridge between RDM and Data Science with research, service, and training for a specific research environment. Its hybrid nature requires a deliberate balance between new developments and the reuse of existing solutions: numerous concepts, recommendations, technologies, and schema- or data-level standards already exist and should inform decision-making rather than be reinvented. At this early stage, the platform's work is to identify where action is needed and to derive its fields of activity accordingly. We organise these activities along the three identified fields of application – illustration and communication, documentation and organisation, comparison and recommendation – which are linked to achieving synergetic effects with "natural" project demands.

Interdisciplinary Research Data Management

Interdisciplinary Research Data Management (IRDM) is the platform's response to the challenge of achieving project goals while dealing with the complexity, heterogeneity, and

interdisciplinarity of the researchers it supports (see Strohm et al., 2023a, 2023b) in four areas: alignment with the FAIR principles as a core funder requirement, reporting of results and project progression to funders and the wider community, promotion of collaboration and early data exchange before publication, and provision of an integrated search space for project data where applicable (see Strohm et al., 2024a, 2024b). Note that the latter is distinct from the methods-side search space discussed briefly discussed below, which concerns procedures rather than data, even though there is potential to view methods as data. This is, in practice, the documentation and organisation field: where products link to real artefacts and the FAIR-readiness of materialised products becomes an operational concern of roles, processes, and software adapted to the ROOTS environment. Identifying repeatable parts of a research process that can be supported, reflected or possibly automated, is a recurring task here, and it also provides the basis for the generalisation of methods toward cross-domain reuse.

Alignment with the Existing Ecosystem

The platform's activities are embedded in an ecosystem ranging from actors within the project to local, national, and global actors, each contributing constraints of differing intensity, from informational material and recommendations to fixed metadata requirements and standardised vocabularies. Internally, the platform collaborates with individual researchers, subclusters, and cross-cutting units such as the technical platform and the methods nucleus. Locally at Kiel University, the opendata@uni-kiel repository serves as the institutional home for FAIR research results, and the Centre for Interdisciplinary Data Science (CIDS)[3] advises researchers along the data life-cycle and provides an expert network for interdisciplinary method development; the ISAAK[4] community of practice in programming for archaeological analysis is a further cornerstone for practical exchange among researchers.

At the national level, the DMDSP follows the recommendations of the DFG and seeks active exchange with NFDI, the National Research Data Infrastructure, primarily with NFDI4Objects, regarding FAIR-compliant archiving, standardised vocabularies, and prospective integration into a domain knowledge graph, where applicable. The IANUS[5] data centre for archaeology and ancient studies serves as another potent exchange partner, while re3data serves as a registry through which domain-specific repositories can be located in the counselling process. Globally, the platform connects – being a part of ROOTS – with the Coalition for Archaeological Synthesis (CfAS)[6], the general-purpose repository Zenodo, the data publisher PANGAEA. It also considers infrastructures such as ARIADNE and the European Open Science Cloud (EOSC) as relevant to its environment.

Because the platform is archaeology-focused, particular attention is paid to domain-specific data and to interoperability with established archaeological repositories such as Open Context[7] and the Archaeology Data Service[8], as well as to community efforts on metadata standardisation for physical samples (IGSN-DataCite Archaeology and Cultural Heritage Community of Practice, 2025a, 2025b). FAIR remains the platform's core commitment; the CARE principles for data governance (Carroll et al., 2020) become directly relevant in the more sensitive contexts the cluster encounters – above all work involving human remains and ancient DNA, where questions of stewardship and community interest arise – and are engaged there on a case-by-case basis, rather than applied uniformly across all material.

[3] https://www.cids.uni-kiel.de – retrieved 2026-07-13.
[4] https://isaakiel.github.io/ – retrieved 2026-07-13.
[5] https://ianus-fdz.de/ – retrieved 2026-07-13.
[6] https://www.archsynth.org/ – retrieved 2026-07-13.
[7] https://opencontext.org/ – retrieved 2026-07-13.
[8] https://archaeologydataservice.ac.uk/ – retrieved 2026-07-13.

Scientific Method Development through Collaborative Use Cases

Method development addresses the side of the RDM-Data Science bridge for which fewer established anchor points are currently available. While materialised products are FAIR-ready, their procedural counterparts are not (yet): generalising and comparing methods across endeavours – in a comparison and recommendation scenario – requires representations that do not yet have settled standards, and developing them is an explicit objective of the platform. The vehicle for this work is concrete collaboration: the platform participates in ROOTS use cases that require adapting existing methods or developing new ones, which lets the framework be challenged and validated against real research data and analytical goals rather than in the abstract alone.

As early proofs of concept for the illustration and communication field, two published archaeological analyses have been reconstructed as MRP representations: a refinement of Nordic Bronze Age chronology built on a chain of correspondence analyses (Kneisel, 2013), and an ancient-DNA pipeline for geographic-origin inference are illustrated and made openly available.[9] Each recovers a methodical approach at a chosen resolution and was validated in consultation with the original authors, demonstrating that an existing analysis can be represented as legible, layered knowledge.

To probe the documentation and organisation field from the user's side, an early-feedback focus group was conducted with researchers from economics and the social sciences, as part of the SEA-EU collaboration between Kiel University and the University of Cádiz. The discussion confirmed the communicative and organisational motivation behind the concept and clearly identified several conditions for adoption. The persistence and trustworthiness of linked artefacts were dominant concerns, while low overhead and reusable templates emerged as the primary factors supporting adoption. A second focus group is planned at Kiel University, centred around ROOTS members of different disciplines, to test whether these conditions hold for the platform's local core audience; a fuller account of this empirical strand will be reported separately. These findings already feed back into the framework and the planned tooling, and exemplify the kind of empirical engagement on which the platform's development depends.

Training and Counselling

Training and counselling underpin all three fields by building the capacity researchers need to use them. A key partner is the young researchers' Academy within ROOTS, for whose members specific courses are being developed, in collaboration with the Kiel University Computing Centre and its Research Data Management division. Connecting experts in research data management, data science, and computer science with young researchers from archaeology and related disciplines helps the latter meet both their own research goals and funder requirements. The content and formats of these services are closely coordinated with the relevant schedules and with the hardware resources available through the CIDS Lab and the broader ROOTS environment.

Research Roadmap

The platform is held together less by its organisational chart than by a research direction that cuts across its activities. We outline this direction as four strands, each already anchored in work that has begun rather than being merely planned.

Concept development. The MRP framework briefly introduced above is the conceptual core, and its continued development is itself a research strand: refining the modules and their composition, and treating the development of a research process over time (versioning, revisions, discarded alternatives) as first-class structural elements rather than incidental history. This work

[9] Reconstructions available at https://doi.org/10.5281/zenodo.21335822 and https://doi.org/10.5281/zenodo.21335857 – retrieved 2026-07-13.

proceeds in dialogue with the platform's use cases, so that the framework is shaped by the research it is meant to represent.

A search space for methods. The comparison and recommendation field requires what materialised products already have and methods do not: a shared, searchable representation. The platform pursues the construction of a Methods and Algorithms Search Space (MASS) and an associated knowledge graph as a description space in which methodological segments can be compared, ranked by similarity, and recommended across endeavours and disciplines. Building a representation at once expressive enough for methodological nuance and disciplined enough for reliable comparison is a substantial undertaking, pursued as a separate line of research.

Tooling. For the documentation and organisation field to be usable in daily practice, a tool is needed that lets researchers build and maintain a representation alongside their work. Its requirements are not assumed but drawn from the empirical engagement reported above: persistence and trust in linked artefacts as the principal concern, and low overhead with reusable templates as the principal adoption lever. A design that links to rather than duplicates artefacts, with a backup or resolver strategy for broken links and a library of templates for recurring processual segments, follows directly from those findings.

Empirical validation and adoption. Whether the framework improves communication and organisation in practice cannot be settled by conceptual coherence alone. The focus group described above is a first step, to be followed by the planned Kiel session centred on archaeologists and further ROOTS disciplines, and ultimately by technology-adoption studies in which domain and computational researchers are brought together under different use case conditions.

Realising these strands requires an organisational effort that the platform treats as a pre-condition rather than as an outcome. It is people-centric: relevant roles must be staffed, and ROOTS members, Principal Investigators in particular, must be engaged early to articulate their data-management needs and help shape and prioritise the platform's activities. Most young researchers, joining in the second half of 2026, will be onboarded through counselling and courses, and the platform's goals will be revised continuously through sustained exchange with stake-holders within and close to the Cluster of Excellence. On the infrastructure side, this includes connecting with providers such as the Kiel University Computing Centre and the state initiative FDM.SH, and evaluating external services regarding knowledge graphs or commons developed within NFDI4Objects (von Rummel et al., 2024), NFDI4Culture (Bruns et al., 2024), and Base4NFDI offerings (Bernard et al., 2024) for reuse where possible.

Conclusion

Transparent, reproducible, and reusable research in an interdisciplinary setting depends on bridging two perspectives that are usually pursued apart: the stewardship concerns of Research Data Management and the analytical concerns of Data Science. This paper has argued that the Modular Research Processes (MRP) framework provides a shared representation in which the two perspectives can meet, describing a research endeavour as products and procedures across layers of granularity, and making the relationship between them precise at the point where materialised products are FAIR-ready while their procedural counterparts are not yet. The three application fields that guided the development of this representation – illustration and communication, documentation and organisation, comparison and recommendation – provide the framework with concrete purpose, and they recur as both the activities a supporting platform performs and the strands along which its research advances.

We have grounded this account in the Data Management and Data Science Platform (DMDSP), presented deliberately as a situated case: one cluster's organisational answer to the question of how such a bridge can be staffed, aligned with an existing ecosystem, filled with collaborative method development, and sustained through training. MRP, not the platform, is the transferable contribution, with a claim to cross-disciplinary and cross-project application that the DMDSP illustrates rather than exhausts.

The work reported here is at an early stage, and we have been explicit about what that means: two published analyses reconstructed as available proofs of concept[9], a focus group probing the conditions for adoption (first results shared later in 2026), and a research roadmap whose strands are grounded in this initial work rather than merely projected from it.

What remains is to carry these strands forward: a searchable space of methodological descriptions, a tool that meets the practitioners' requirements concerning persistence and low overhead, and an adoption study in which the framework's central promise, that a shared structured representation eases the transfer of methodology across disciplinary boundaries, can be put to the test.

Acknowledgements

The authors would like to thank Jutta Kneisel, Ben Krause-Kyora and Nicolas A. da Silva for the fruitful collaboration on MRP representations of their respective analytical research cases. Preprint version 2 of this article has been peer-reviewed and recommended by Peer Community in Archaeology (https://doi.org/10.24072/pci.archaeo.100686; Pavlidis, 2026).

Funding

This work has been supported by the German Research Foundation (DFG) under Germany's Excellence Strategy through the ROOTS Cluster of Excellence (EXC 2150–390870439) and under the National Research Data Infrastructure through the Research Data Infrastructure for the Material Remains of Human History (NFDI4Objects, PNr. 501836407).

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article. The authors declare the following non-financial conflict of interest: Oliver Nakoinz, Steffen Strohm (PCI Reviewers).

Data, script, code, and supplementary information availability

Early examples from the Illustration and Communication field were semi-automatically reconstructed and validated by the original authors.

The paper-based example (Kneisel, 2013) is available at: https://doi.org/10.5281/zenodo.21335822 (Strohm, 2026a).

The code-based example (not yet published in its domain) is available at: https://doi.org/10.5281/zenodo.21335857 (Strohm, 2026b).

References

Barker M, Chue Hong NP, Katz DS, Lamprecht AL, Martinez-Ortiz C, Psomopoulos F, Harrow J, Castro LJ, Gruenpeter M, Martinez PA, Honeyman T (2022). Introducing the FAIR Principles for research software. *Sci. Data* **9**, 622. https://doi.org/10.1038/s41597-022-01710-x.

Bernard L, Jander M, Manske A, Miller B, Rettberg N, Ritter X, Rißler-Pipka N, Schäfer-Neth C, Stein R, Zänkert S (2024). Base4NFDI Policy Paper: "Base4NFDI Services and EOSC: Guidance for Interoperability". Version 1. https://doi.org/10.5281/zenodo.13946300.

Bruns O, Söhn LC, Tietz T, Steller JJ, Ondraszek SR, Posthumus E, Schrade T, Sack H (2024). What's cooking in the NFDI4Culture kitchen? A KG-based research data integration work-flow. In: *Proceedings of the 4th Workshop on Metadata and Research (objects) Management for Linked Open Science (DaMaLOS 2024), co-located with the Extended Semantic Web Conference (ESWC 2024)*. Hersonissos, Crete, Greece: PUBLISSO. https://doi.org/10.4126/FRL01-006474028.

Carroll SR, Garba I, Figueroa-Rodríguez OL, Holbrook J, Lovett R, Materechera S, Parsons M, Raseroka K, Rodriguez-Lonebear D, Rowe R, Sara R, Walker JD, Anderson J, Hudson M (2020). The CARE principles for indigenous data governance. *Data Sci. J.* **19**. Article 43, 1–12. https://doi.org/10.5334/dsj-2020-043.

Cox AM, Tam WWT (2018). A critical analysis of lifecycle models of the research process and research data management. *Aslib J. Inf. Manag.* **70**, 142–157. https://doi.org/10.1108/AJIM-11-2017-0251.

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996). The KDD process for extracting useful knowledge from volumes of data. *Commun. ACM* **39**, 27–34. https://doi.org/10.1145/240455.240464.

IGSN–DataCite Archaeology and Cultural Heritage Community of Practice (2025a). Archaeological Sample Metadata Profile. Version 1.0. Zenodo. https://doi.org/10.5281/zenodo.17254974.

IGSN–DataCite Archaeology and Cultural Heritage Community of Practice (2025b). *IGSN–DataCite Archaeology CoP Crosswalk Recommendation*. Zenodo. https://doi.org/10.5281/zenodo.17446557.

Kneisel J (2013). New Chronological Research of the Late Bronze Age in Scandinavia. *Danish Journal of Archaeology* **2**, 95–111. https://doi.org/10.1080/21662282.2013.904129.

Object Management Group (2014). *Business Process Model and Notation (BPMN), Version 2.0.2*. OMG Specification. Accessed: 2026-07-22. Object Management Group. https://www.omg.org/spec/BPMN/2.0.2.

Object Management Group (2017). *OMG Unified Modeling Language (OMG UML)*. Specification formal/2017-12-05. Version 2.5.1. Accessed: 2026-07-22. Object Management Group. https://www.omg.org/spec/UML/2.5.1/.

Pavlidis G (2026). Bridging research data management and data science through modular representations of research workflows. *Peer Community in Archaeology* **100686**. https://doi.org/10.24072/pci.archaeo.100686.

Strohm S (2026a). *Supplementary Material (HTML): Git Repository Release on MRP Example 1 (Kneisel 2013, Correspondence Analyses)*. Version 1.0. Zenodo. https://doi.org/10.5281/zenodo.21335822.

Strohm S (2026b). *Supplementary Material (HTML): Git Repository Release on MRP Example 2 (da Silva, aDNA analysis)*. Version 1.0. Zenodo. https://doi.org/10.5281/zenodo.21335857.

Strohm S, Buenning H, Renz M (2023a). (Poster) Hit by Reality – How Socio-Technical Challenges Shift the Design of an Information System in an Interdisciplinary Research Environment. Poster presented at the 19th IEEE International Conference on e-Science (eScience 2023), Limassol, Cyprus, October 9–13, 2023. Zenodo. https://doi.org/10.5281/zenodo.10047720.

Strohm S, Buenning H, Renz M (2023b). Implementing a FAIR information system for archaeology-related interdisciplinary research. In: *2023 IEEE 19th International Conference on e-Science (e-Science)*. Limassol, Cyprus: IEEE, pp. 1–2. https://doi.org/10.1109/e-Science58273.2023.10254840.

Strohm S, Wölker Y, Renz M (2024a). (Poster) Project-Specific Research Data Management Beyond Repositories for FAIR Research Practice. Poster presented at the 20th IEEE International Conference on e-Science (eScience 2024), Osaka, Japan, September 16–20, 2024. Zenodo. https://doi.org/10.5281/zenodo.13897370.

Strohm S, Wölker Y, Renz M (2024b). Project-Specific Research Data Management Beyond Repositories for FAIR Research Practice. In: *2024 IEEE 20th International Conference on e-Science (e-Science)*, pp. 1–2. https://doi.org/10.1109/e-Science62913.2024.10678716.

UK Data Service (2022). *Research Data Lifecycle*. Accessed: 2026-07-22. https://ukdataservice.ac.uk/learning-hub/research-data-management/.

von Rummel P, Keller C, Fricke F (2024). NFDI4Objects: Warum brauchen wir eine gemeinsame Forschungsdateninfrastruktur für die materiellen Hinterlassenschaften der Menschheitsgeschichte? *Archäologische Informationen* **47** (1), 63–72. https://doi.org/10.11588/ai.2024.1.110390.

Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018. https://doi.org/10.1038/sdata.2016.18.