NORDDESIGN CONFERENCE
REYKJAVIK, ICELAND, 12-14 AUGUST 2024

Research Data Management Workflow - A Framework for Engineering Science

Max Leo Wawer, Johanna Wurst, Roland Lachmayer

Insitute of Product Development, Leibniz University Hannover

Abstract: Many research results in engineering science are based on the further processing of digital objects, which increases the need for reusable data. In this context, researchers are becoming increasingly familiar with implementing comprehensive data management as part of their research. Using the example of an engineering research result a generic research data management workflow was developed to ensure the traceability and reusability of research results and associated digital objects. This workflow represents a research process-orientated research data management driven by data provenance.

Keywords: Design Research, Data Driven Design, Design Process, Research Data Management, Workflow

1 Introduction

With the increasing amount of data generated in research and the associated requirement to make it reusable and comprehensible for further research, researchers are confronted with various new challenges in research data management (RDM). It includes all data practices, manipulations, enhancements, and processes that ensure that research data are of high quality, well organized, documented, preserved, sustainable, accessible, and reusable (Corti et al., 2020). The requirements for an RDM are often associated with the FAIR Data Principles according to Wilkinson et al., which aim to make research data findable, accessible, interoperable, and reusable (Wilkinson et al., 2016). The Researchers carrying out the research are usually responsible for implementing RDM during their research (Clements and McCutcheon, 2014). The implementation of holistic RDM is not only driven by external parties, such as funding bodies or the research community but also by the researchers' benefits from qualitatively assured RDM. On the one hand, the implementation of RDM allows the described research results to be reproduced and verified to support scientific findings; on the other hand, accessible data can be reused by the research community, which saves resources (time and money) and helps the scientific community to grow together, creating multiplier effects. The focus is on the reproducibility of the research results with a database, often referred to as data provenance, which describes the derivation of further processing by documenting the origin of the data and the procedures and methods used to collect it. Data provenance according to Gupta, *"refers to a record trail that accounts for the origin of a piece of data (in a document, database, or web repository) or a process together with an explanation of how, who, where, and why it got to the present place"* (Gupta, 2009). In a scientific context, the description of data provenance aims to be able to reconstruct the further processing of data up to the research result.

Due to the direct relationship between RDM and the actual research, these are directly related to each other and the data management should be aligned with different research areas in terms of context and thus be integrated as part of the actual research process. It is therefore essential that domain-relevant community standards are implemented for the context-related documentation and formats used for the digital objects, also named as digital artefacts, depending on the research method and associated research community. These need to be developed and established in and for the various communities. In Figure 1, a schematic data driven research process in engineering research with the associated RDM activities is shown. As part of the research process, data is collected, processed, and analyzed to produce a research result that is published in engineering research. In this context, RDM has become more relevant due to new methods and opportunities through the use of artificial intelligence like machine learning as the quality of the database is crucial for the reliable use of these methods (Kumar et al., 2017). To this end, data must be standardised and prepared through secure data management to be usable for further processing. In the research process, various data artefacts are generated and further processed in terms of content. These are then described in more detail in a publication, the research result. Based on the result, the data provenance refers to the reconstruction of the data origin contrary to the data processing in the research process, which can be ensured by a research-related RDM.

In engineering research, research data occurs in a wide variety of forms and types and is therefore referred to as digital objects that go beyond data in the narrower sense (Wittenburg, 2019). This includes all digital objects that arise digitally as artefacts in the research process and are primarily collected, reused, further processed, or further analyzed. In this context, the term FAIR digital object is also used when it comes to processing these artefacts as part of RDM and ensuring their reusability and reproducibility (Smedt et al., 2020; Schultes and Wittenburg, 2019). For this reason, in this paper, the term digital object is used to describe the digital artefacts generated in research, which is used as a synonym for the word data.

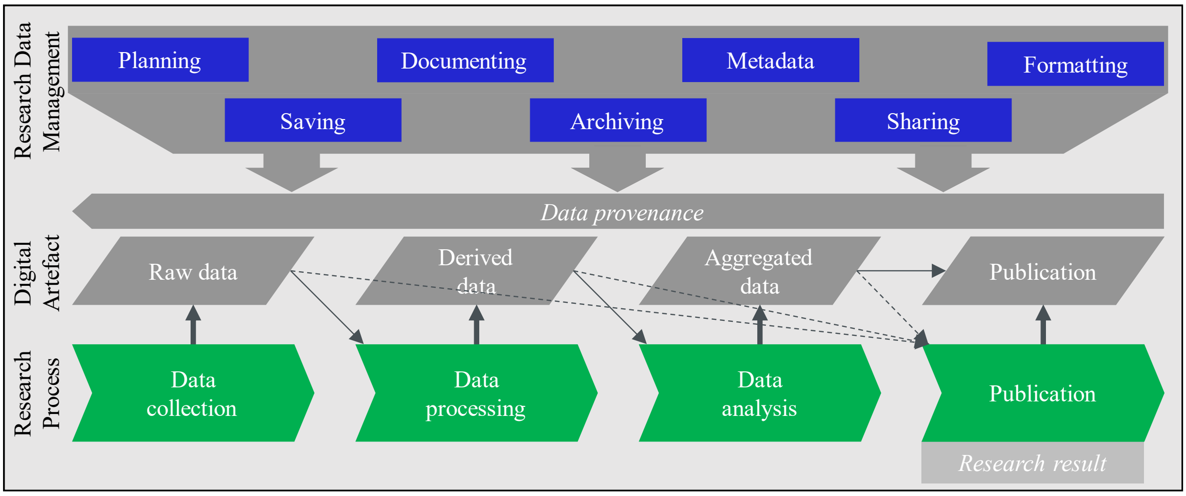

Figure 1. Data driven research process and fundamentals in RDM (according to (Iglezakis and Schembera, 2018))

In Germany, these topics are currently being addressed by the NFDI4Ing (National Research Data Infrastructure for the Engineering Sciences), a consortium of the NFDI e.V. that offers engineers services in the field of RDM. The perspectives of several engineering disciplines and research methods are included, and both are supported by overarching working groups (Schmitt et al., 2020).

In this paper, a generic workflow for ensuring data provenance and the research-oriented implementation of RDM is presented. The workflow model can support the implementation to illustrate relevant practices depending on the research process and the interaction of various digital objects that are collected in the course of research.

To this end, the following aspects and questions are discussed:
1. Which RDM tasks are to be realised by researchers within the research?
2. How can RDM be implemented in the research process?
3. How can the traceability of research results be ensured in the context of RDM?
4. How can a workflow reflect the heterogeneous research in engineering science?

Answering these questions enables the development of a generic workflow model that can be used to map the implementation of RDM throughout the research process and can be applied independently of the research method. Section 2 fundamentally summarizes existing work on the implementation of RDM and the representation of research processes and data lifecycles. The underlying method and the developed workflow for representing research process-accompanying RDM are described in Sections 3 and 4. In Section 5, the workflow development is exemplified using a use case from engineering research.

2 State of the Art

This chapter presents the fundamentals of RDM and corresponding models of data management-driven research processes. Data life cycle models are usually used to illustrate RDM and relate directly to the data (Weber and Kranzlmüller, 2019). In the models, phases are strung together sequentially in a cyclical representation. In these models, the individual phases basically cover "planning, collection, analysis, archiving, access, reuse". Reuse is then transferred back to planning to illustrate the subsequent use and further processing of data. In research, there are different representations of life cycle models, with which a structure for the consideration of operations is represented, which must be carried out in dealing with data throughout the entire life cycle (Sinaeepourfard et al., 2015; Ball, 2012; Wolf and Leppla, 2020). The application of lifecycle models to RDM support services has several advantages because from the creation to the use and finalisation of research data, it is necessary to prepare them in their format, application, use, and purpose to make them reusable and traceable. By identifying and naming the transformations that data goes through as phases in a larger lifecycle, organisations can better align their services (Carlson, 2014).

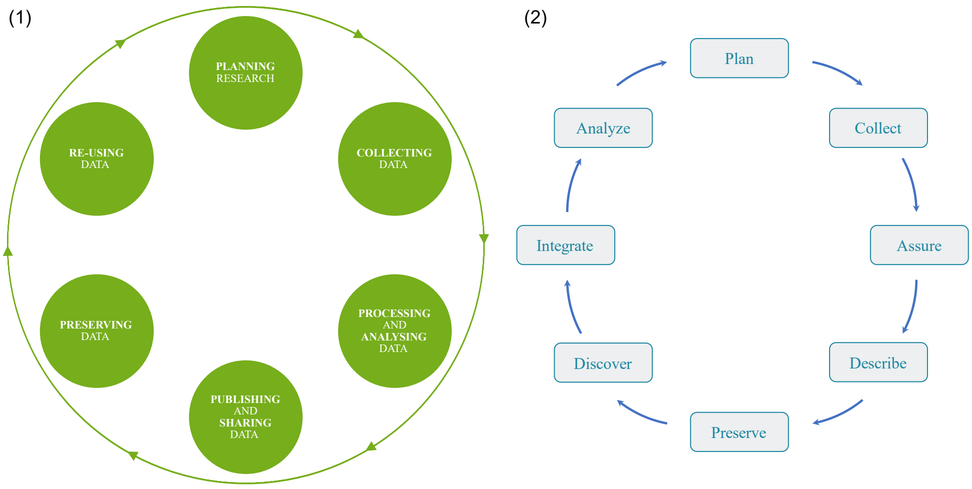

Figure 2. Research data lifecycle models by (1) UK Data (UK Data Service, 2019) and (2) DataOne (Michener et al., 2011)

The abstract life cycle models serve as a guide, but have their limitations, as life cycle models tend to represent an idealised version of the processes. These representations relate to RDM in general and contain relevant phases and content that need to be implemented for comprehensive RDM. However, an orientation towards the research process is only inadequately represented, as these models do not consider an end in the form of a research result or the end of a project; rather, the cyclical further processing of digital objects is at the center of these models. Research data management practices and services mainly cover tasks such as data storage and data sharing, as well as data handling and documentation; data management plans are also included to a lesser extent (Briney et al., 2020; Ashiq et al., 2022). For the planning of RDM, a data management plan is an integral part of the organisation of RDM in research projects and provides for project-specific planning aligned with the research method along the entire life cycle (Mozgova et al., 2021). To ensure that the data can be reused and retraced, the necessary contextual information and associated documentation must be collected in a research-specific manner. To describe data in more detail, further information must be defined in the form of metadata. Metadata can have different functions, they can describe the process of data collection in more detail, contain general administrative information, or also basic bibliographic information for the publication of data (Brase, 2009). To increase the findability and reusability of research data in data repositories, contextual domain-specific metadata must be defined and annotated to the data (Altun et al., 2023). Especially in collaborative centers, there is a great need for comprehensive and defined RDM solutions. The use of RDM systems in the form of knowledge management and data repositories can improve data exchange and knowledge acquisition at the project level. To this end, standardised vocabularies must be defined to ensure the traceability and interoperability of the data (Mozgova et al., 2020). The importance of tracing research data back to research results is demonstrated by the use of FDM systems with the linking of standardised documentation of data collection using the example of process chain development, in which various research projects investigate manufacturing processes along a process chain. Using a knowledge management system, sample production can be reconstructed and new process knowledge can be generated by semantically linking research data from previous process steps (Sheveleva et al., 2022).

With the integration of research-related content into data management workflows, Kowlazky's work presents a digital research data lifecycle in which research methodological content is integrated. This proceeds from data collection to quality control to a content phase ends in data collection. All four phases are accompanied by the necessary data management, which is defined in the phases. In addition, the first three phases are complemented by formatting and contextualisation of the data in each phase, which is finally brought together in the content phase. Research methods and scientific quality control are also embedded in the model (Kowalczyk, 2017). In the work by Wissik and Ďurčo, a research data workflow is presented that relates to institutional research data management. This involves various stakeholders and relates to the development of data management within an institution. The workflow contains six phases, from the initiation of a project with a project idea to data processing with the publication and storage of data and continuous quality control to enable the reuse of the data. Except for the quality assurance phase, the phases are similar to those of the data life cycle models (see Figure 2). Process steps such as "data acquisition" and "analysing" and continuous "documentation" are defined and linked to associated FDM systems (Wissik and Ďurčo, 2015).

3 Methodology

To enable a research-oriented integration of RDM into the research process and to present it in a workflow model, a top-down approach was pursued in this work, as the research processes in the engineering sciences do not yet include a full implementation of RDM. The core processes of the RDM are identified and integrated into the research process based on the target-oriented process identification, which, according to Schmelzer and Sesselmann, pursues a top-down approach to business process management (Schmelzer and Sesselmann, 2020). Starting from the literature, RDM practices, and data lifecycle models, the integration into the research process is carried out and then illustrated using a use case from engineering research. The requirements, which up to now have mostly been holistic, are thus to be mapped in the research process in an implementation-orientated manner. Due to the different research methods and research fields in engineering research, a workflow model is developed that can be generically adapted to different research methods. To this end, the digital object that is generated in research and must be handled as part of the RDM is used directly as the reference point for the workflow model. The existing data lifecycle models and RDM models are a general representation of RDM processes at a higher level in the form of phases, with the lifecycle models having no defined beginning and end, which results in the need for a research-orientated representation. To demonstrate a research process-orientated implementation, the research process in this model is presented in a result-orientated manner with the achievement of a research result to show a limitation in the research process. Any type of research-relevant content from the researcher's perspective that provides an answer to one or more research questions is regarded as a research result. In the engineering sciences, this usually involves the publication of research results in text form. This results in a workflow model for the research process-orientated implementation of RDM, which aims to embed the measures and activities of RDM in the research process of engineering researchers in a process-orientated manner. So that it is not considered as an additional burden at the end of the research process.

4 RDM Provenance Workflow

The following presents a workflow model for implementing RDM in a research process. The model reconstructs research results based on the research process, aligning data management tasks with digital objects. This helps researchers implement research-oriented RDM.

For this purpose, the research process is delimited by (sub-)research questions and an associated research result. Furthermore, different types of digital objects that arise in the course of research are shown. This is followed by the integration of RDM practices aligned to the research process concerning the digital object, which is transferred into the generic workflow model.

4.1 Engineering research process

In the heterogeneous engineering community, the great diversity of problems, which are concentrated in certain (sub-) research areas, has led to highly specialised, individualised approaches to solutions. In data-driven research, digital objects are the basis for the scientific cognition process. The approaches to collecting digital objects in engineering range from simulations, modeling, experimental measurements, and theoretical approaches. A certain behavior of a system is observed in data collection. Based on an input in the form of a start configuration, the behavior of the system is observed by measuring or recording properties, thereby generating an output. The output then represents digital objects, which are subsequently processed and prepared concerning the properties to be analysed. This involves filtering and summarising the data in a visualization and format suitable for analysis. The processed digital objects are then used to analyse the content to present the research results. A final digital object usually reflects the research result, which is published. In engineering research, for example, a research result is usually associated with the initiation of a publication, which in this work represents the boundary and conclusion of a research process. The beginning is the underlying (sub-)research question that initiates the creation of the first digital object. Research results in engineering science are usually published in text-based form. In addition to the text-based form, the results are presented in the publication in the form of plots, methods, or source code, depending on the research method. This presentation of results usually only includes the final processing step of a digital object (see Figure 1). Within a research project, many such research processes take place with the achievement of several research results and associated publications

Data collection in the research process is therefore not a onetime event, but an ongoing process involving the collection in the form of generation and gathering of digital objects. There are different approaches to collecting digital objects based on different research methods. On the one hand, digital objects can be primarily generated in the research process through experiments, observations, or instruments, and on the other hand, existing objects from internal or external research can be reused and gathered from databases. Digital objects can also be generated in the form of software models or existing software models can be extracted and further developed. In this way, different processing steps and aggregations of digital objects can be demonstrated in a research process with which research results can be achieved. The focus of the developed model is on the traceability and reusability of the digital objects in the research process, bounded by a defined or realised research result. Starting from a research result, it is important to maintain the traceability and reusability of the digital objects by preparing and handling them following the requirements of RDM. Considering the further processing of digital objects in the research process as individual artefacts, there is a logical connection to a published research result.

4.2 Digital Objects

As described in the previous section, digital objects are collected as artefacts at various points in the research process described. Depending on the research method, the digital objects are created or collected and further processed for this purpose. A final digital object usually reflects the research result to be published and provides the framework for the research process in this work. Therefore, in the workflow model, the term "digital object" is used for all digital entities that arise in the course of research. According to Wittenberg et al., (Wittenburg, 2019) digital objects are the core of a proper data organization, which binds entities that are necessary for a stable reusable data domain. They consist of a sequence of bits (content) that can be stored in different repositories and are referenced by unique persistent identifiers and described by different types of metadata. They consist of a sequence of bits (content) that can be stored in different repositories and are referenced by unique persistent identifiers (ID) and described by different types of metadata. Metadata can also be digital objects themselves. Digital objects can also be combined into collections, which also leads to the creation of digital objects. Digital objects contain different types of information, such as data, software, configurations, and representations. Figure 3 shows the associated core data model for visualising digital objects.

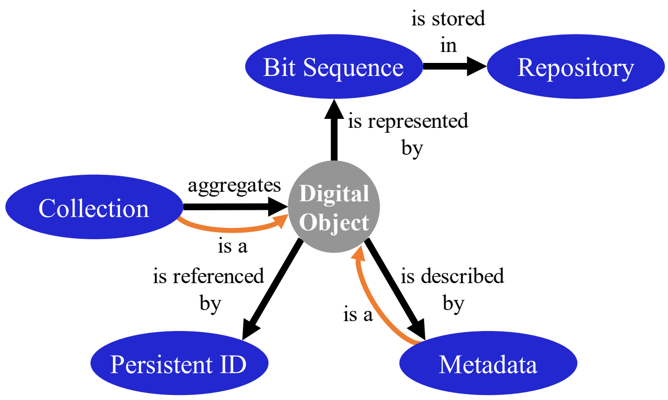

Figure 3. Core Data Model of a digital object (according to (Wittenburg, 2019))

4.3 Research Data Management Tasks

The aim of RDM is to ensure that the digital objects collected as part of the research process are prepared in such a way that they are reusable and comprehensible. RDM, described in the literature, generally refers holistically to the handling of data artefacts in research organisations, represented as life cycle models. For a process-oriented implementation of RDM in research, an application-oriented representation of data management tasks is used, in which the digital objects are placed at the centre. To this end, the life cycle models are broken up for the process view to enable the arrangement and combination of different digital objects. The requirements for the RDM can be translated into activities and presented as tasks along the research process, orientated towards the digital artefacts. Based on the findings of the two previous subsections, the digital object is used as the reference point for RDM in the workflow model. The research process and the resulting data outputs up to the research result are depicted to form the reference framework for the research process. This results in RDM tasks for the digital objects produced as part of the research process that relates directly to the artefact. The sequence of RDM phases shown in Figure 4 results for a single object. Each phase includes RDM tasks that can be assigned to the handling of a digital object.

Planning: In this phase, the creation of the digital object is planned as part of the FDM. This includes the planning of the collection for the design of a primary survey or collection of existing digital objects. Furthermore, the content planning for the creation of the digital object, the documentation of the collection process, the definition of process-relevant metadata and the determination of the occurring format are carried out.
Data collection: In this phase, the data collection is carried out according to the planning; this can be the generation, primary collection of the digital object, or the gathering, reuse, and aggregation of existing digital objects. The defined metadata is also collected in this phase and attached to the digital object. In addition, there is further documentation of the data collection, in which the collection process and the method as RDM tasks are documented alongside the actual collection of the digital objects.
Analysis: At the level of the digital object, the analysis does not mean the research content data analysis (cf. Figure 1) to answer the research question, but the analysis of the digital object and thus the review of the collected Digital object. Here, the quality of the digital object is analysed with regard to the planning and collection for suitability for subsequent data processing.
Archiving: This phase refers to the transfer of the digital object to a suitable repository for archiving. The digital object

will be transferred into a format compatible with long-term archiving and its storage should be secured for a defined period of time. In addition, the digital objects should be linked to technical and descriptive metadata. Archiving is aimed at physical preservation (repository), technical interpretability (formatting), and content interpretability (documenting).
Access: Access often takes place with simultaneous archiving, as this is usually included with the transfer to a repository. In addition, there is the definition of further bibliographic metadata, which is used to define reuse. Access and archiving usually take place in datasets and are not carried out individually for each digital object, which is why storage usually takes place first in the research process (see Figure 4). Archiving and access then only take place when the research result is achieved or at the end of the project.
Reuse: In this phase, no further RDM tasks are carried out, but shows that a digital object is reused. In this respect, all previous phases and tasks should be implemented as part of the research to master RDM.

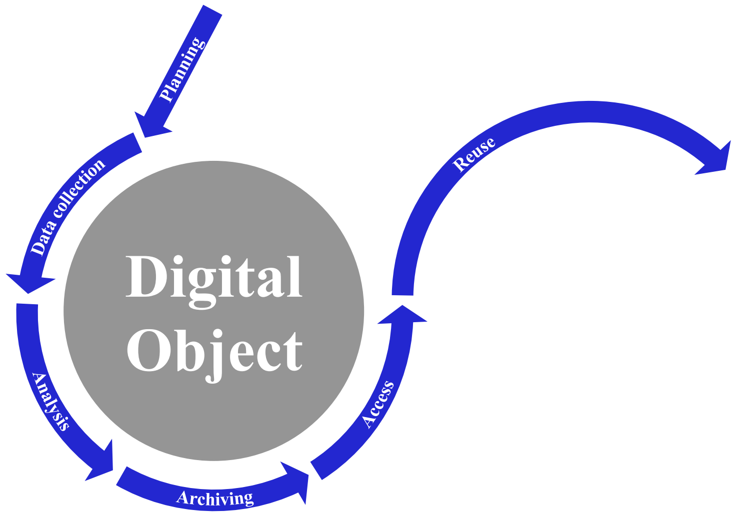

Figure 4. RDM tasks oriented to a digital object

If each digital object is now documented, processed, and saved in accordance with the RDM task, this results in a workflow that can be used to describe the implementation of RDM in the research process. Based on the life cycle representation of data, these are arranged to take into account further processing and enable the RDM to be orientated towards the digital artefacts that occur, regardless of the collection method or content origin.

The implementation of RDM is influenced by research-specific and project-specific requirements, which must be taken into account in the projects. In engineering research, for example, external partners are often involved in research projects, so-called third-party funded projects, which often do not allow open access to data. Here, academia and industry find themselves in a conflict of objectives when it comes to the accessibility of digital objects. Based on the FAIR Data Principles, RDM requires a community-specific approach to data management in the form of standards in documentation and interoperability. As these standards have not yet been established for the most part, RDM is still mostly individualised, which affects the traceability and reusability of the data. This also applies to the formats of the digital objects, which are usually provided in proprietary file formats, as these are system and software-specific and can usually only be reused with the associated software.

Therefore research data management systems must be available to researchers in order to carry out RDM integrated into the research process. System interfaces must be created for the research data management systems and existing research software in order to avoid creating redundant solutions. The systems must also ensure the quality of the RDM and corresponding digital objects (cf. (Müller et al., 2023)).

4.4 Generic Workflow Structure

Based on the research process shown in Figure 1, a research result usually refers to the last digital artefact (e.g. described in a publication). The contents of the previous data generation steps are usually only presented in text form as part of the methodology. To fulfill the requirements of RDM, it is important to establish the provenance of the data and make it traceable following the further processing of the digital object. In this way, the data processing steps for achieving the research result can be reconstructed through the combination of artefact further processing, and the associated artefacts can be reused. For this purpose, the digital objects must be prepared by the FAIR Data Principles (cf. (Schultes and Wittenburg, 2019)). Each digital object stands on its own and can be further processed in research in terms of reuse. In engineering research, different research processes occur depending on the research method, which can be mapped by the developed generic workflow structure, as it allows a combination of individual digital objects. Figure 5 schematically illustrates the formation of the workflow for a data-driven research process.

The workflow represents the aggregation of several individual digital objects in the research process, which are reused and analysed about a specific research question. For example, two digital objects are primarily collected and prepared for further processing. These resulting digital objects are then further processed and aggregated for analysis. This results in the research findings to be published. To ensure that further processing and the individual digital objects can be reused, the processes and further information on data processing must be documented and attached to the digital objects. This applies to the data collection processes as well as the further processing steps to be able to retrace and reproduce synthesis and analysis steps. Especially with qualitative data, which can be the basis for research results, the documentation and traceability of further processing are essential to be able to assess subjective influences. For these influences, making raw data accessible is also fundamental to be able to reproduce the processing steps according to the data provenance. When considering digital objects in the research process, archiving and access to artefacts usually takes place at the end of the process when the research result is achieved. During the research process, these two phases are replaced by storage. External reuse of the digital artefacts is made possible when they are made accessible (e.g. as part of a publication). In this way, the digital objects can be reused in other, possible external, research and therefore integrated as metadata in the form of a persistent identifier in new research processes.

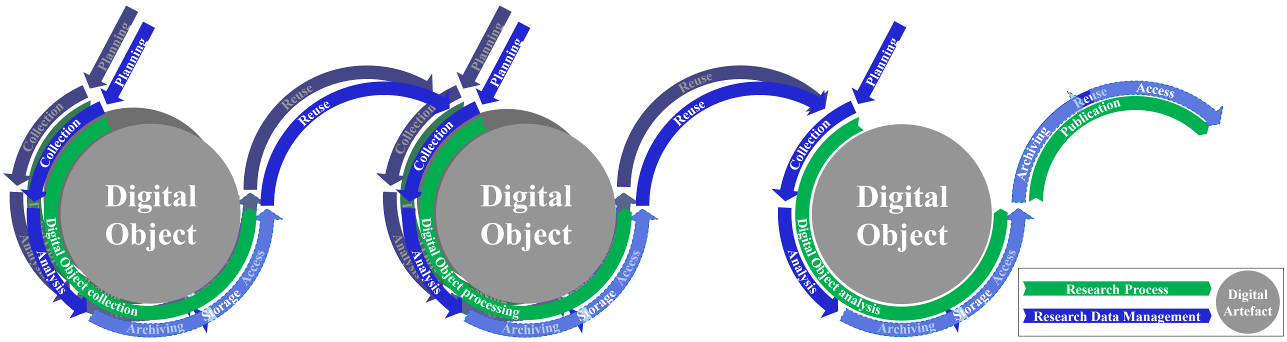

Figure 5. RDM workflow in engineering research based on digital objects

The use of data-intensive research methods such as machine learning in particular increases the number of digital objects in the form of databases that need to be aggregated, requiring the quality of each individual digital artefact to be ensured by secure and standardised RDM. In this respect, the preparation is directly linked to the data sets, which are then reused for further research. In this context, the need for external databases and the gathering of data for own research is also increasing, resulting in a new digital object for further processing in the own research process. The development (primary survey) and application (reuse) of software programs as digital objects can also be represented in this way. Furthermore, the interaction of different types of digital objects with the use of software for analysing databases can be represented by the generic workflow structure. In this way, two digital objects (software programs and databases) are reused for further processing in the research process as part of the generation of a digital object. Even research synthesis can be mapped with the workflow. Data collection, in this case, involves the gathering and reuse of existing data (such as publications), which is processed and analysed with regard to a defined research question. The resulting database allows the described research results to be traced back to the gathered publications, which can then be reused for further research purposes. In this way, qualitative research results can be traced back to databases, using publications as an example. The digital objects can also appear as metadata if they are accessible and referenced by a persistent ID. In this way, RDM is aligned to individual digital objects within the research process and implemented in the research process. In addition to the activities described, resources and systems must be incorporated into the workflow to enable the requirements and implementation of RDM in research processes, as described at the end of chapter 4.3. This presents both research institutions and projects with new challenges to improve RDM in research. The research processes described show the need for a research-orientated RDM, which can be generically mapped with the developed workflow model, thereby providing a secure RDM. This allows increasingly complex research processes to be broken down into individual digital objects and visualised by these.

5 Use Case – Engineering Science

In this chapter, the developed workflow model is illustrated using a use case from engineering research. As part of the case study, design guidelines for the additive manufacturing of multi-material components were established as a research result (Meyer et al., 2023). The design guidelines represent the published research results and are the final digital object (in the format of a table) in the closed research process. The guidelines are based on a parameter study of the scan process with optical inspection procedures as the measurement procedure, such as in-process visual inspection and post-process scanning electron microscope (SEM), as well as post-process radiographic inspection using micro-computed tomography (µCT) in the manufacturing process. The data are primarily generated in the research process and thus represent the first digital objects. For the documentation of the scanning process, machines and process parameters were defined and annotated. Also, for the measurement procedure and the derivation of the design guidelines, the scanning parameters were annotated with the boundary conditions and the parameter sets of the parameter study to the data for reusability. The visual data was then processed for analysis. The data analysis was then the basis for the synthesis and derivation of the design guidelines. The test specimens were analysed on a microscopic level for internal defects on all sectional planes. Based on these digital objects from the visual inspections, machine/process-specific and part/material-specific design guidelines were obtained. For this research process, the data processing workflow shown in Figure 6 is based on the research results. The digital artefacts are presented as the output of each research step.

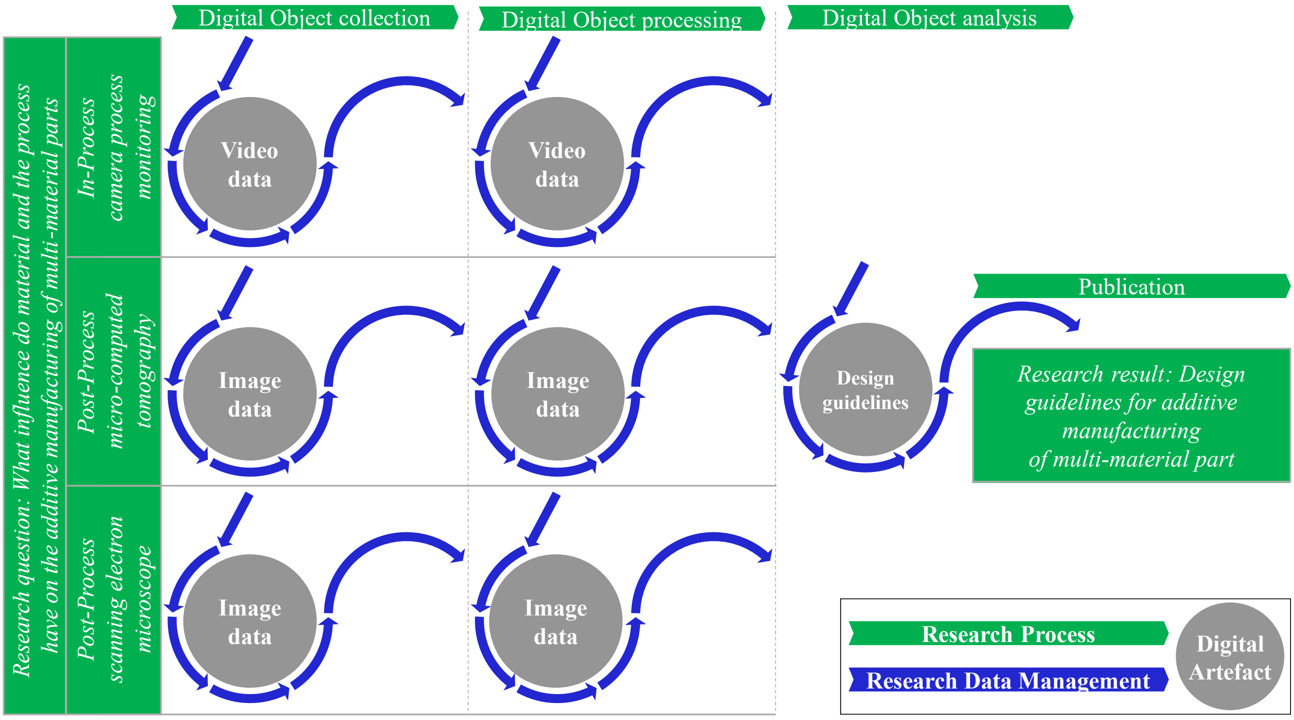

Figure 6. Use case workflow for the data-based derivation of design guidelines

The use case illustrates the aggregation of different digital objects in the research process and the highly specialised and individualised approaches to research results. The synthesis of design guidelines is not atypical in the engineering sciences and can be derived based on the consideration of different properties. The research process is based on a total of 7 digital objects, three of which were primarily collected. The digital objects were then further processed so that they could be used for the analysis and synthesis of the design guidelines. The design guidelines were derived from the digital objects of the process monitoring and were documented in material- and process-specific design guidelines in the form of a catalogue. The individual artefacts were documented and processed following the defined RDM tasks in the research process and were saved during the research process; at the end of the process, the underlying digital objects that led to the derivation of the design guidelines were published and made accessible. The research results were then published as part of a publication, which can also be seen as an additional digital artefact. In most cases, publications are regarded as the primary digital artefact in reuse and cited in the context of further research.

6 Discussion and Conclusion

In this paper, we describe an approach for an RDM orientated research workflow to handle digital objects generated in research to ensure good scientific practice in the context of RDM. The workflow is generically structured so that different research methods and approaches with the requirements of engineering research can be applied to it. This workflow is intended to raise awareness for researchers in the implementation of RDM directly in research, as these topics are usually still considered too general and not application-oriented. With this approach, a research-oriented implementation of RDM in the context of research processes could be presented. The focus of this workflow is on the research result and the underlying digital objects and their provenance. The workflow enables to handle the digital objects that are created as part of the research to be further processed and serve to answer the research question and represent the research result. This approach also follows the FAIR Data Principles when dealing with digital objects that have led to the research result in order to ensure the traceability and reusability of these artefacts in the context of the data provenance. As part of this workflow, a direct link was established to the research process in the context of achieving research results. To this end, RDM was directly aligned with the digital artefacts produced. The RDM tasks identified for the preparation and further processing of digital objects in the research process deal with basic content. However, there are a number of research-specific and project-specific aspects in the implementation of RDM as mentioned before, such as data licence models or domain-specific metadata, which must be developed specifically. For this reason, the content is described descriptively in the workflow model, as it must be specifically designed and implemented. The model can represent areas for a holistic RDM in the research process and at which point which content must be fulfilled. The implementation also requires a large number of interfaces and systems in order to ensure effective implementation in the research process without producing obstructive additional work. Here, a direct exchange between the data management level and the research level must be made possible in order to link the artefacts of data management directly with the research artefacts and organise them along the process. In addition, integrating RDM into the research process leads to more work in the short term. Still, it enables the effectiveness of research and exchange to be improved in the long term and the quality of the reusable digital objects to be ensured. This opens up the potential for more effective collaborative research.

For further consideration, the context of RDM should be established at project level, as data is usually archived at the end of the project and a data management plan is created at the start of the project. Thus, RDM is generally defined at project level and implemented in the research process. For RDM, this means that it is considered at different levels; in this work, the focus was placed directly on the research process. These approaches are also used to develop evaluation models for RDM processes; they represent defined phases and tasks that need to be implemented in research and whose execution can be ensured using assessment models (Wawer et al., 2023).

Acknowledgment

The authors would like to thank the Federal Government and the Heads of Government of the Länder, as well as the Joint Science Conference (GWK), for their funding and support within the framework of the NFDI4Ing consortium. Funded by the German Research Foundation (DFG) - project number 442146713.

References

Altun, O., Oladazimi, P., Wawer, M.L., Raumel, S., Wurz, M., Barienti, K. et al. 2023. ENHANCED FINDABILITY AND REUSABILITY OF ENGINEERING DATA BY CONTEXTUAL METADATA, in: Proc. Des. Soc. 3, pp. 1635–1644. DOI: 10.1017/pds.2023.164.

Ashiq, M., Usmani, M.H., Naeem, M. 2022. A systematic literature review on research data management practices and services, in: GKMC 71 (8/9), pp. 649–671. DOI: 10.1108/GKMC-07-2020-0103.

Ball, A. 2012. Review of Data Management Lifecycle Models. Available online at https://researchportal.bath.ac.uk/en/publications/review-of-data-management-lifecycle-models, checked on 1/8/2024.

Brase, J. 2009. DataCite - A Global Registration Agency for Research Data. in:, 2009 Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology. 2009 Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology (COINFO). Beijing, China, 21.11.2009 - 23.11.2009: IEEE, pp. 257–261.

Briney, K., Coates, H., Goben, A. 2020. Foundational Practices of Research Data Management, in: RIO 6, Article e56508. DOI: 10.3897/rio.6.e56508.

Carlson, J. 2014. The use of life cycle models in developing and supporting data services. in: Joyce M. Ray (Ed.), Research data management. Practical strategies for information professionals. West Lafayette, Ind.: Purdue Univ. Press (Charleston insights in library, archival, and information sciences), pp. 63–86.

Clements, A., McCutcheon, V. 2014. Research Data Meets Research Information Management: Two Case Studies Using (a) Pure CERIF-CRIS and (b) EPrints Repository Platform with CERIF Extensions, in: Procedia Computer Science 33, pp. 199–206. DOI: 10.1016/j.procs.2014.06.033.

Corti, L.,van den Eynden, V.,Bishop, L.,Woollard, M.,Haaker, M.,Summers, S. 2020. Managing and sharing research data. A guide to good practice. 2nd edition. Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE.

Gupta, A. 2009. Data Provenance. in: LING LIU, M. TAMER ÖZSU (Eds.), Encyclopedia of Database Systems. Boston, MA: Springer US, p. 608.

Iglezakis, D., Schembera, B. 2018. Anforderungen der Ingenieurwissenschaften an das Forschungsdatenmanagement der Universität Stuttgart - Ergebnisse der Bedarfsanalyse des Projektes DIPL-ING, in: o-bib 5 (3), pp. 46–60. DOI: 10.5282/o-bib/2018H3S46-60.

Kowalczyk, S.T. 2017. Modelling the Research Data Lifecycle, in: IJDC 12 (2), pp. 331–361. DOI: 10.2218/ijdc.v12i2.429.

Kumar, A., Boehm, M., Yang, J. 2017. Data Management in Machine Learning. in: Rada Chirkova, Jun Yang, Dan Suciu (Eds.), Proceedings of the 2017 ACM International Conference on Management of Data. SIGMOD/PODS'17: International Conference on Management of Data. Chicago Illinois USA, 14 05 2017 19 05 2017. New York, NY, USA: ACM, pp. 1717–1722.

Meyer, I., Oel, M., Ehlers, T., Lachmayer, R. 2023. Additive manufacturing of multi-material parts - Design guidelines for manufacturing of 316L/CuCrZr in laser powder bed fusion, in: Heliyon 9 (8), e18301. DOI: 10.1016/j.heliyon.2023.e18301.

Michener, W., Vieglais, D., Vision, T., Kunze, J., Cruse, P., Janée, G. 2011. DataONE: Data Observation Network for Earth Preserving Data and Enabling Innovation in the Biological and Environmental Sciences, in: D-Lib Magazine 17 (1/2). DOI: 10.1045/january2011-michener.

Mozgova, I., Jagusch, G., Freund, J., Kraft, A., Glück, T., Herrmann, K. et al. 2021. Product Life Cycle Oriented Data Management Planning with RDMO at the Example of Research Field Data. With assistance of Vincent Heuveline, Nina Bisheh. in: Vincent Heuveline, Nina Bisheh (Eds.), E-Science-Tage 2021: Share Your Research Data, Heidelberg, pp. 105–118.

Mozgova, I., Koepler, O., Kraft, A., Lachmayer, R., Auer, S. 2020. Research Data Management System for a large Collaborative Project. in:, Balancing Innovation and operation. Proceedings of NordDesign 2020, 12th - 14th August 2020: The Design Society.

Müller, L., Wawer, M.L., Heimes, N., Uhe, J., Koepler, O., Auer, S. et al. 2023. Data quality assurance in the research process using the example of tensile tests. in:, DS 125: Proceedings of the 34th Symposium Design for X (DFX2023). 34th Design for X Symposium (DFX 2023), pp. 143–152. Available online at https://www.designsociety.org/publication/46856/data+quality+assurance+in+the+research+process+using+the+example+of+tensile+tests.

Schmelzer, H.J.,Sesselmann, W. 2020. Geschäftsprozessmanagement in der Praxis. Kunden zufriedenstellen, Produktivität steigern, Wert erhöhen. 9., vollständig überarbeitete Auflage. München: Hanser (Hanser eLibrary).

Schmitt, R.H., Anthofer, V., Auer, S., Başkaya, S., Bischof, C., Bronger, T. et al. 2020. NFDI4Ing - the National Research Data Infrastructure for Engineering Sciences.

Schultes, E., Wittenburg, P. 2019. FAIR Principles and Digital Objects: Accelerating Convergence on a Data Infrastructure. in: Yannis Manolopoulos, Sergey Stupnikov (Eds.), Data Analytics and Management in Data Intensive Domains. 20th International Conference, DAMDID/RCDL 2018, Moscow, Russia, October 9–12, 2018, Revised Selected Papers, vol. 1003. 1st ed. 2019. Cham: Springer (Springer eBooks Computer Science, 1003), pp. 3–16.

Sheveleva, T., Wawer, M.L., Oladazimi, P., Koepler, O., Nürnberger, F., Lachmayer, R. et al. 2022. Creation of a Knowledge Space by Semantically Linking Data Repository and Knowledge Management System - a Use Case from Production Engineering, in: IFAC-PapersOnLine 55 (10), pp. 2030–2035. DOI: 10.1016/j.ifacol.2022.10.006.

Sinaeepourfard, A.,Masip-Bruin, X.,Garcia, J.,Marín-Tordera, E. 2015. A survey on data lifecycle models: Discussions toward the 6vs challenges. Technical Report (UPC-DAC-RR-2015–18).

Smedt, K. de, Koureas, D., Wittenburg, P. 2020. FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units, in: Publications 8 (2), p. 21. DOI: 10.3390/publications8020021.

UK Data Service 2019. Data liecycle. Available online at https://ukdataservice.ac.uk/learning-hub/research-data-management/, checked on 2/8/2024.

Wawer, M.L., Wurst, J., Lachmayer, R. 2023. Quality Assessment for Research Data Management in Research Projects, in: Proc Conf Res Data Infrastr 1. DOI: 10.52825/cordi.v1i.420.

Weber, T., Kranzlmüller, D. 2019. Methods to Evaluate Lifecycle Models for Research Data Management, in: Bibliothek Forschung und Praxis 43 (1), pp. 75–81. DOI: 10.1515/bfp-2019-2016.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J.J., Appleton, G., Axton, M., Baak, A. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship, in: Scientific data 3, p. 160018. DOI: 10.1038/sdata.2016.18.

Wissik, T., Ďurčo, M. 2015. Research Data Workflows: From Research Data Lifecycle Models to Institutional Solutions. in:, Selected papers from the CLARIN annual conference, pp. 94–107. Available online at https://ep.liu.se/ecp/123/008/ecp15123008.pdf.

Wittenburg, Peter, Strawn, George, Mons, Barend, Boninho, Luiz, Schultes, Erik 2019. Digital Objects as Drivers towards Convergence in Data Infrastructures. Available online at https://www.rd-alliance.org/sites/default/files/Digital_Objects_as_Drivers_towards_Convergence_in_Data.pdf, checked on 1/30/2024.

Wolf, A.H., Leppla, C. 2020. Harmonisierung von Datenlebenszyklus-Modellen: Nutzung von Synergien für optimierte Anwendungen im FDM, in: BFDM (2), pp. 1–19. DOI: 10.17192/bfdm.2020.2.8281.

Contact: Wawer, Max Leo, Leibniz University Hannover, wawer@ipeg.uni-hannover.de