RESEARCH ARTICLE

PRO Research Process 4Ing: Project- and RDM-Oriented Research Process for Engineering Sciences

Validation of a Newly Prosed Process to Combine Data Life Cycles and Project-Oriented Research Processes in Engineering Sciences

Tobias Hamann [1]
Michèle Robrecht [2]
Marcos Alexandre Galdino [1]
Anas Abdelrazeq [1]
Robert H. Schmitt [1, 2]

1. Laboratory for Machine Tools and Production Engineering | Intelligence in Quality Sensing (WZL|IQS), RWTH Aachen University, Aachen.
2. Fraunhofer Institute for Production Technology IPT, Aachen.

Date Published:
-

Reviewer:
-

Keywords:
Engineering Research, Research Data Management, Data Lifecycle, FAIR Data, Research Data Management Process

Data availability:
Data can be found here:
10.5281/zenodo.17100240

Software availability:
Software can be found here:

Abstract. In a first paper [1], we have shown the absence as well as the importance of a project-oriented research data management (RDM) process for engineering sciences. We argue that the integration of RDM into the practice would greatly benefit by the availability of such project-oriented RDM process. While many RDM and engineering research processes already exist, a combination of both was only recently designed in our previous paper [1]. Seeking for validity of the proposed RDM process, the research question in this article is: "Can our proposed RDM process be deployed into engineering research projects?" We interviewed ten (n=10) experts on their research practices and the fit of the proposed process to their research activities. Overall, we found a broad acceptance of the proposed process, as well as several possible potentials for enhancements, e.g. regarding wording and loops. As a result, this paper presents a validated project-oriented RDM process for engineering sciences.

1 Introduction

Integrating RDM processes into day-to-day work is still a hurdle for researchers in engineering sciences. [2], [3] We showed the reasons for that in our previously published paper *Matching Data Life Cycle and Research Processes in Engineering Sciences* [1]: While many suggestions on how to carry out RDM already exist [4], [5], [6], their integration into practice is not feasible on a one-to-one basis focusing on a strong theoretical reference [7]. In this context, the data life cycle (DLC) model is often used to refer to only research data. However, it cannot be transferred to research activities in general, yet it is frequently used for this purpose [1], [7]. Furthermore, for engineering sciences, project-orientation and specificity in engineering are missing in RDM processes. Project-oriented research has been increasing rapidly for years [8], [9], [10], [11], which is enforced by the "greater third-party funding activity in the natural sciences and engineering" [11]. This project-oriented approach spans through all forms of research and occurs in basic research as well as in applied research. [12]

In our first paper [1], we addressed this challenge (lack of integration of RDM into day-to-day work and project structures) while also focussing on the demands of researchers in engineering sciences. The finding of this paper is a multi-layered RDM process, combining the overall structure of project-oriented research with the execution of research and the handling of data. While this process was built upon two data collections of researchers' practices and demands, the resulting process has to be cross-checked by its users to ensure that it is sufficient in filling the identified gap. On this basis, the applicability of the process has to be evaluated and validated in this article, leading to the research question:

Can our proposed RDM process be deployed into engineering research projects?

To answer the proposed research question, we conducted ten qualitative expert interviews. Experts were questioned on their day-to-day research activities and the fit of the proposed process – later called Project- and RDM-Oriented Research Process for Engineering Sciences (PRO Research Process 4Ing) – to these activities.

The structure of this paper follows the line of reasoning in the introduction. Firstly, we sketch out related work and present the identified research gap. We also introduce our previously designed process. After that, we elucidate the research methodology used to validate and enhance our process. The results of the validation are shown in section 4, which describes the collected feedback to the previously designed process for the possible enhancements. Section 5 discusses the methodology and limitations of the validation. The paper closes with a conclusion and discussion.

2 Related and previous work

To give an overview on the research work considered when creating the previously designed process, this section sketches out different approaches to RDM and engineering research process modelling. An extract from the most relevant literature gives an overview of the considered literature. This overview leads to a research gap (see section 2.2), identified in our first paper [1]. This gap was addressed with the newly proposed process, detailed in section 2.3.

2.1 Extract from the most relevant literature

The literature holds engineering research processes, DLCs and other concepts for RDM processes as three main fields. The processes considered were on the one hand based in engineering, while on the other hand, different approaches to RDM processes were included.

For **engineering research processes**, data management is a prevalent topic. The data generated in engineering research is often extensive and complex in nature. Hence, a careful management of data is implied. Furthermore, the division of complex problems into smaller parts is inherent to engineering sciences. This division of bigger parts into smaller sub-parts is also portrayed in the process proposed by Griem et al. [13], as shown in figure 1.

DLCs are closed loops that describe the lifespan of data throughout different stages of its life cycle.

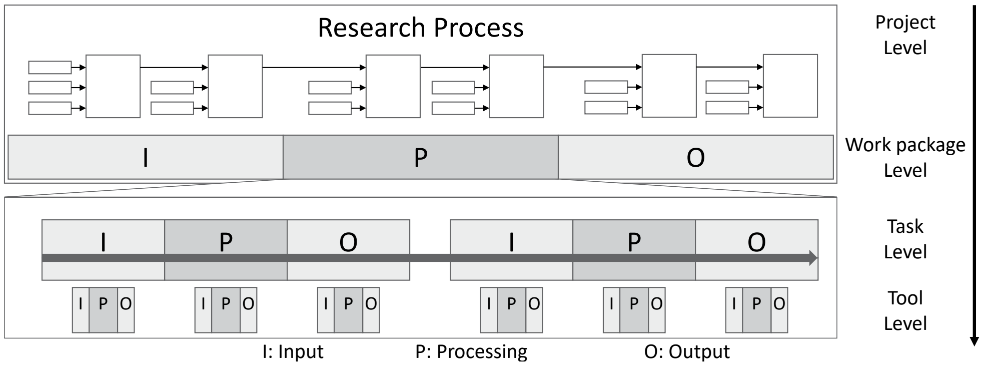
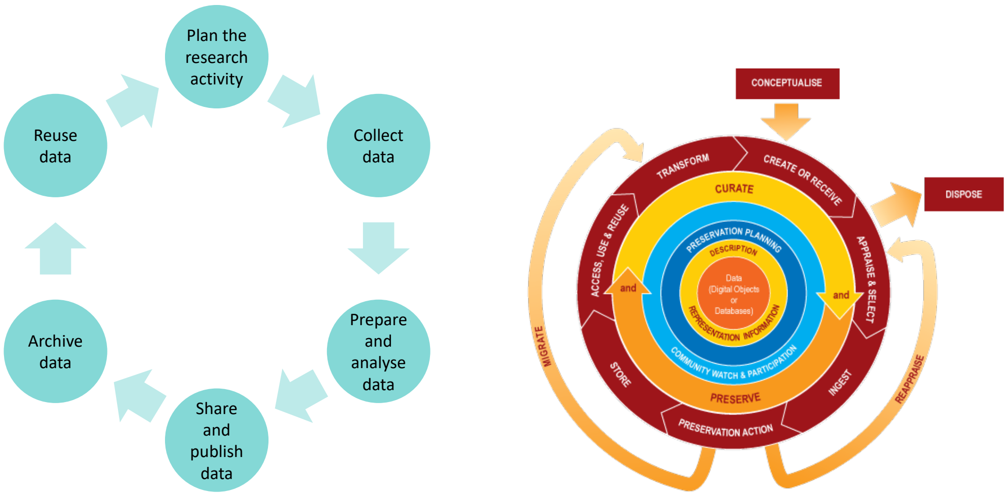


Figure 1: Research process as proposed by Griem et al. [cf. 13] [1]

DLCs ”ensure[...] that all the required stages are identified and planned, and necessary actions [are] implemented, in the correct sequence. This can ensure the maintenance of authenticity, reliability, integrity and usability of digital material” [6]. Many authors have already proposed DLCs as shown by Cox et al. [7] and Shah et al [14], e.g. ELIXIR [4] and Politze [5]. Their complexity differs from straight forward and plain simple (see figure 2a) to more complex and with input, output and feedback loops (see figure 2b).



(a) Data life cycle as model proposed by forschungsdaten.info [15]

(b) Data life cycle model as proposed by the Digital Curation Centre [6]

Figure 2: Contrasting a simple with a more complex DLC

Other concepts for RDM processes exist besides DLCs. These range from enhanced DLCs, as proposed by Tripathi and Pandy [16] to loose clusterings of activities like Putnings et al. [17] propose, and finally to “FAIR Workflow” [18] cycles. Still, none of these approaches considers both project related activities as well as iterative nature of research in certain stages. This leads to the identified gap.

2.2 Identified research gap

In our previous paper, we identified two main issues with the state of the art in project-oriented RDM processes. First, while the variations of detail and complexity in the processes as well as the definition of their sub-steps and activities are vast, there is not yet a process with a scope, that reaches from the project as a whole down to the data in detail.

Second, there is a gap between DLCs and other, more holistic models. [1] While DLCs support the management of data along its life cycle, they are not supposed to be utilised as a support structure for project management. DLCs are a "useful metaphor, but tend to encourage thinking that research processes are highly purposive, uni-directional, serial and occurring in a closed system. Research is often not like this" [7]. Contrasting this insufficiency, engineering research processes often lack RDM context and applicability. Hence, a process combining both project-oriented research and RDM did not exist to the best of our knowledge. We aimed to close this gap with our proposed process.

2.3 Previously designed process

The process proposed in [1] combines three levels of research, along with waterfall project management and agile approaches. Besides the relevant literature, it is also based on eight problem-centred interviews to gather requirements of researchers regarding the application of RDM. Furthermore, eight engineering researchers were interviewed on their research and typical workflows. The result is the three layered, multi-staged process as shown in figure 3.

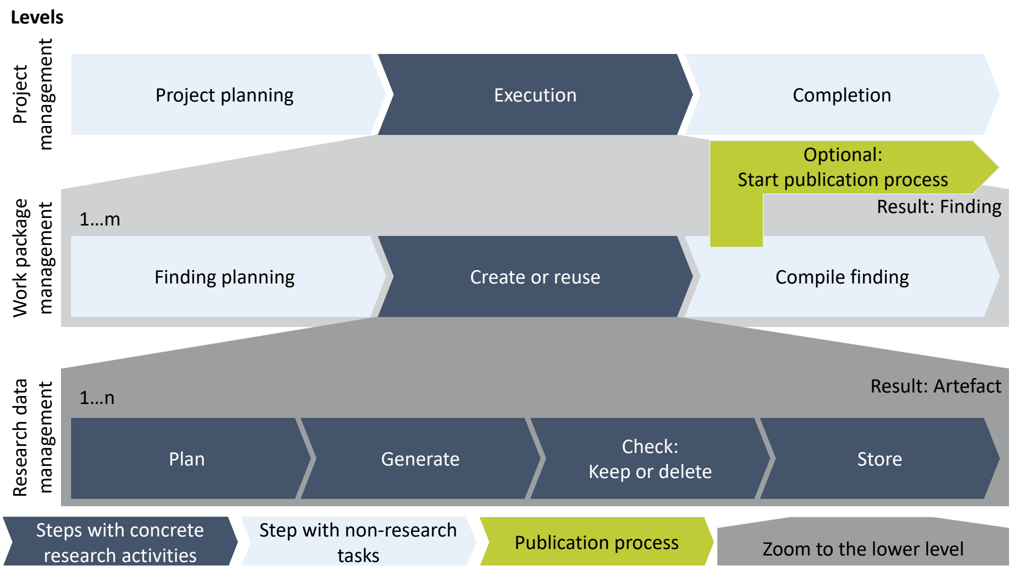


Figure 3: General structure of the previously designed RDM process

The **project management level** encompasses a waterfall-like approach. The project planning includes all activities that happen before the research project is accepted, either in form of a provided grant or accepted research proposal. Then the project is executed. This execution includes the content-wise execution, ongoing project management and other activities. The project completion starts when the project is coming to an end. It contains activities such as the writing of final reports or archiving of data. This level of the process is depicted in greater detail in figure 4.

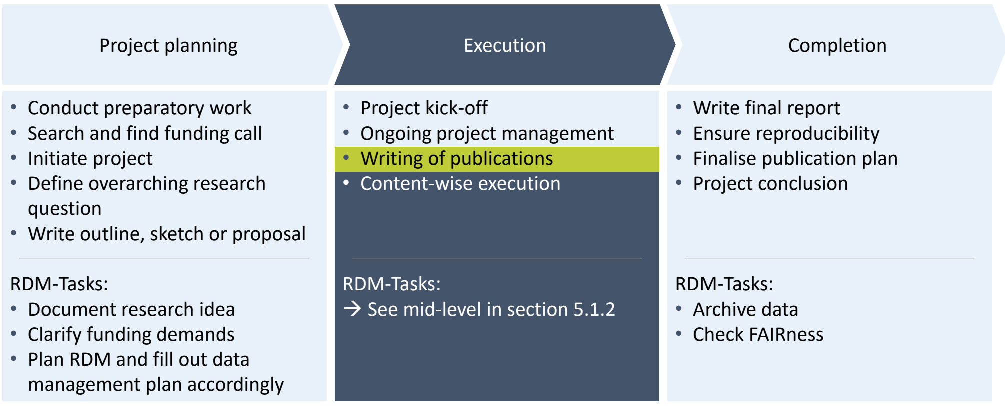


Figure 4: Project management level of the proposed research process

Below the project management level, the **work package level** is placed, where work packages are processed. Hence, this level is conducted *n* times in an iterative approach. For finding planning, firstly, a work package is chosen and divided into sub-steps needed to reach its goal. Afterwards, artefacts are created or reused - what researchers consider "actual research" and happens on the research data management level. The creation of artefacts has to happen *n* times to finish the work package. Afterwards, the so created artefacts are compiled to a finding in the finding compilation step. The finding is the result of the work package. The work package management level is depicted in figure 5.

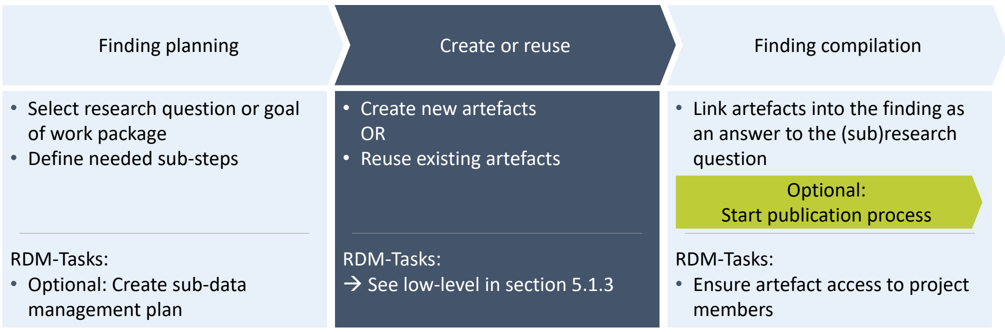


Figure 5: Work package management level of the proposed research process

Artefacts are the smallest unit of data in a research project. Depending on the use case, it can be a single file, a data set of several files or physical object, e.g. a demonstrator or a workpiece. If it is a digital piece of data, it is considered to be a digital object. [19] However, it is not necessarily a FAIR Digital Object (FAIR-DO) by the definition of Schwardmann [19], as a Persistent Identifier (PID) might be missing. An artefact is always the single outcome of one research method applied.

On the **research data management level**, each artefact is created by firstly planning its generation. This planning happens immediately before it is generated. It is not a experimental design, rather the experimental design would be a previous artefact. Instead, the planning of the data generation encompasses the gathering of the needed artefact, the methodical setup and the check if everything is properly configured. Then, the data is generated. Optimally, metadata is recorded along with the data generation, annotating the data, transferring it from data to data with context. Afterwards, a check of the data ensures the data was recorded at all, is reasonable and within expectations and not corrupted. Otherwise, it can be deleted immediately. If the data was recorded properly, it can be stored for further usage. This process revolves around each artefact, as shown in figure 11 and in figure 6.

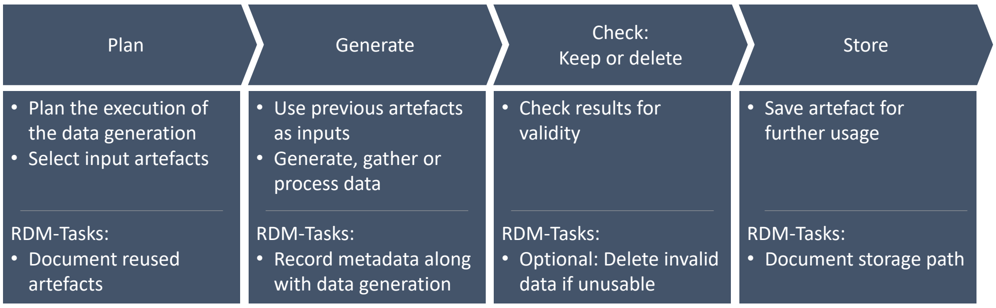


Figure 6: Research data management level of the proposed research process

Additionally to the finding compilation, a **publication process** can be started. It starts optionally from the work package level and runs in parallel. The topic definition, the first step in the process, is determined by the finding compilation. The publication is then prepared, which includes, for example, finding a suitable medium. This is followed by editing: Artefacts, findings and the knowledge generated as a result are converted into text and - as an RDM task - published. In the submission step, the publication is submitted, subjected to a quality check through a peer review process and finally published. This is at last followed by dissemination as illustrated in figure 7.

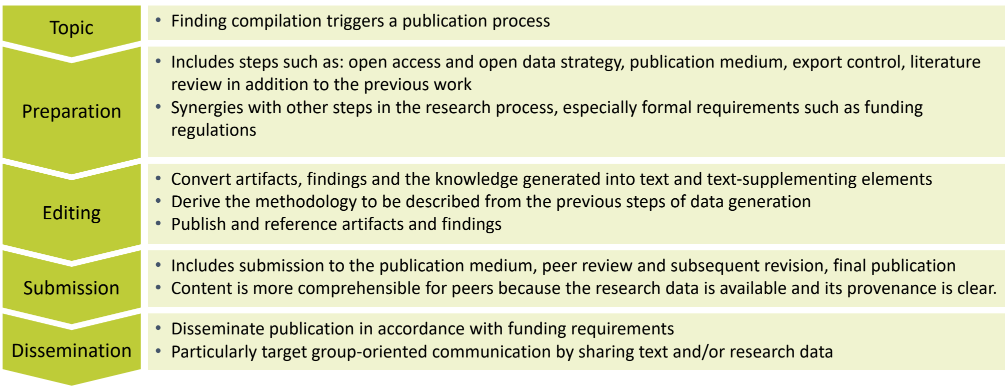


Figure 7: The publication process

The goal of the developed process is the combination of project-oriented research, demands of engineering researchers and RDM processes. This combination could solve the problem of lacking applicability of RDM found in [1]. However, it has to be validated, if this goal could be reached by the proposed process.

Parts of the proposed process has also already been taken up and further investigated and implemented by Wawer et al. [20]. However, in their paper, Wawer et al. [20] focus on the research data management level, elaborating on the concept of artefacts, their linking and compilation into a research result. In addition to their process, Wawer et al. [20] point out, “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”. This establishment is the focus of our proposed process both in this and the previous paper. Yet, to this point the process has not been proven to be sufficient, which makes a validation of this process necessary as presented in the next section.

3 Research methodology

As research methodology, we conducted expert interviews and analysed using qualitative content analysis. We conducted ten interviews, transcribed, anonymised, coded and aggregated. In the following, the exact procedure is explained.

3.1 Collecting the data: Planning and conducting the interviews

Expert interviews were conducted to validate the process. Interviews in general are a qualitative method of data collection that originates from the social sciences. [21], [22] The question posed was processed using non-standardised interviews - more precisely, guided interviews. In this approach, interviewers use a guideline that complies topics and questions they are expected to introduce during the interview. In order to create a natural flow of conversation, they can use the guidelines flexibly, adapt the order of questions and ask follow-up questions. This should allow the interviewees to answer as freely as possible. The expert interview method was chosen because a special type of knowledge was required for the validation: not only specialised knowledge, but also practical knowledge. In this type of interview, the experts themselves are not the object of investigation, but the medium through which the information for answering a question is collected. [23] The guideline for the expert interviews has been designed following the remarks of Gläser and Laudel [23].

Our goal was to have at least one interviewee - resp. expert - per TU9 – German Universities of Technology e.V. (TU9)-university, the Alliance of leading Universities of Technology in Germany, representative of engineering research at universities in Germany. In order to find experts who actually have the type of knowledge required, there were requirements for the interviewees: As the proposed process focusses mainly on engineering sciences, researchers had to have a corresponding engineering background. Furthermore, we demanded at least four years of experience in research. After some enquire, we could reach our goal of one interviewee per TU9-university with the required background.

The created interview guideline contained five parts (see table 1).

The free descriptions of their research activities were intended to ensure that the interviewees did not orientate themselves to the process before the interviewers asked for direct comparisons. In this way, the actual state was queried and not implicitly associated with a target state. Nevertheless, it was necessary to make the process accessible to them before the interview so that they could prepare themselves. The demographic section (part 4) asks respondents to categorise their technical and professional situation so that their answers can be put into context.

This guideline was reviewed by an expert for survey methods form the Chair of Intelligence in

Part of the guideline	Content of the part
Part 1: Experiences with RDM	Previous experiences and attitudes towards RDM to contextualise the answers
Part 2: Presentation of the process	Interviewees were allowed to ask questions about the process and clarify ambiguities before the discussion
Part 3: Survey on the fit of the process to the interviewees' own research	Interviewees outlined the example project and were asked about the fit in this project for each process level as well as the publication process and the overall process. They first freely described how the research was conducted, after which the process was shown and they were asked for concrete comparisons to their descriptions.
Part 4: Information on the interviewees	A few demographic questions about the interviewees' careers and specialised contexts
Part 5: Conclusion	Space for concluding remarks and questions as well as an outlook on the further validation process

Table 1: Structure of the interview guideline

Quality Sensing (WZL-IQS) at the RWTH Aachen University (RWTH) and adapted accordingly. Additionally, a pretest was performed with a researcher from the RWTH. The learnings of the pretest were considered when conducting further interviews: minor improvements in the technical setup, interviewer behaviour trough gaining experience, minor changes in the wording of questions. However, as these learnings were not focussed at the core content of the guideline, the pretest was still considered usable and was also included into the final results.

The interviewees received for preparation an introduction to the interview and the declaration of consent to the interview as well as a description of our proposed RDM process in form of an excerpt from the previous article about it [1]. They were asked to sign the declaration of consent, read the article and to think about a research project in which they were involved, ideally from the application phase to the project completion. If they were not involved in a research project from start to finish, they were supposed to choose a research project in which they were involved for as long as possible and in which they experienced either the application phase or the project completion.

The interviews lasted approximately one and a half hours each and were conducted by two interviewers simultaneously, one researcher and one member of the research supporting staff. This made it possible to divide the interviews according to background knowledge and to ask differentiated questions. The interviews were conducted online and recorded.

3.2 Preparing the data: anonymization and stock of data

After each interview, the recording was transcribed with the support of a custom python pipeline for audio transcription. The transcription was anonymised partly automated using QualiAnon[1]. [24] Afterwards, the anonymised transcripts were imported into MaxQDA[2].

1. https://github.com/pangaea-data-publisher/qualianon
2. https://www.maxqda.com/

Ten interviews were included in the subsequent analysis: one from each of the TU9 Universities and the pretest at RWTH Aachen University. The interviews contain a total of 117,066 words.

The demographic questions make it possible to categorise the interviewees in terms of their technical and professional backgrounds: The interviewees' scientific carriers are mostly set in mechanical engineering, but also in industrial engineering and economics. Data science and civil engineering are also represented. This means that the engineering sciences are thematically diverse in the interviews, which increases the quality of the validation. The interviewees already have a doctorate degree or are about to complete it. They have experience as research assistants, but also in research and transfer management. Some also fulfil management roles such as project and group leaders. Also, one professor was participating in the interviews. The minimum required four years of professional experience was mostly exceeded, in some cases significantly with eight years or more. These professional backgrounds guarantee the availability of the required expert knowledge and enable differentiated perspectives on the process.

All interviewees had already had contact with RDM. Their activities were partly of an administrative nature, partly they had used it themselves - both in solo work while writing their dissertations and in projects of different sizes. In the discussions, they mentioned many aspects of RDM that they had already encountered: archiving (hardware and software), publication, creating guidelines, Documentation, data management plan (DMP)s. They often did not work systematically, but intuitively and according to the ”trial and error” principle, because knowledge was often lacking. Due to this lack of knowledge, tools that are not yet fully developed, external and internal guidelines that must be adhered to and poor integration into day-to-day work, their experiences with RDM have so far been predominantly bad. However, there are also interviewees who have had good experiences because they have benefited from the RDM activities after enduring initial stress.

Also, the interviewees were asked to bring an example project which would function as a mental guideline to describe their research processes. These projects were also surveyed in the interviews. These ranged from very personal projects with only a single person involved, e.g. doctoral degree projects, to big scale projects with multiple institutions involved. Both projects with and without industry participation were brought up by the interviewees. In terms of content, the projects described ranged from research on artificial intelligence, data modelling and processing, simulation, quality and process management, robotics, circular economy, computational science, manufacturing technology and electronics.[3]

3.3 Analyse the data: qualitative content analysis

In MaxQDA, the transcripts were analysed based on the qualitative content analysis method as proposed by Mayring [25]. Qualitative content analysis analyses material with regard to its content. [26] Since this material represents fixed communication (texts, images, music, etc.), the communication itself is also the subject of the analysis. It is characterised by a systematic, rule- and theory-based approach. The following section describes the practical implementation of this validation.

3. This categorization was simplified for anonymisation reasons.

Modelled after Mayring [25], a process adapted to this validation was created as shown in figure 8: The code system was formed deductively based on the theory and the question: one code each for agreement and disagreement per level and sub-step (cf. figure 9) as well as seven other codes, which are listed in table 2.

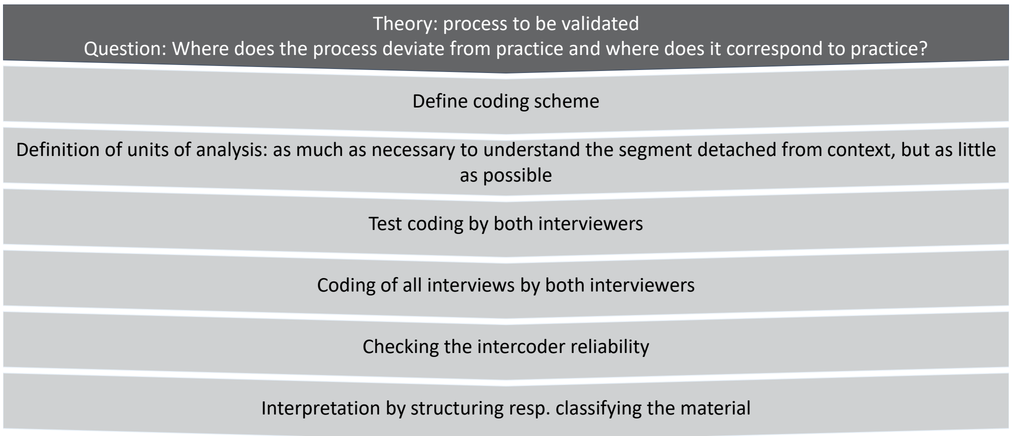


Figure 8: Content-analytical model based on [25]

A total of 66 codes were used, resulting on 741 coded segments in the ten interviews. Structuring resp. classification was used as the basic form of interpretation, i.e. filtering out certain aspects from the material by fitting its content into the code system.

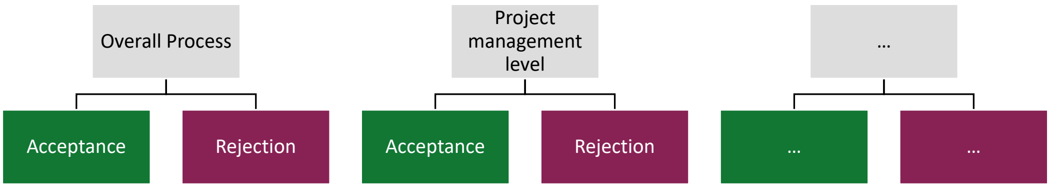


Figure 9: A sketch of the utilised code system

The coding was performed by the two interviewers individually to improve reliability and objectivity.

Afterwards, the codes were reviewed by a lead interviewer, using an intercoder analysis. In this analysis, similar coding, e.g. same codes with slightly different start and endpoints of the codes, was equalised. Additionally, conflicting codes, e.g. codes for acceptance and rejection for the same sections of transcripts, were identified. Conflicts were discussed between the reviewers and solved if possible, e.g. when a wrong coding was accidentally applied. However, only few conflicts could be solved that way. The conflicts that could not be solved in that way point towards inconsistencies in the proposed process and therefore show possible improvements. These inconsistencies and improvements are discussed in the results and changes to the process.

Concluding the analysis, the coded sections were summarised by their respective codes. In that way, statements to each code could be aggregated. These aggregations are further condensed into the results and changes to the process.

4 Results and changes to the process

In this section, the results are presented. Firstly, a descriptive feedback on the interviews is given. Afterwards, general feedback on the overall process is presented. Lastly, the different process levels are discussed regarding their inconsistencies in the proposed process and possible improvements. The coded segment are referred to as *statements* in the following. We have translated German statements into English.

4.1 Descriptive analysis of the interviews

The overall feedback given is positive. More than two thirds of the evaluative statements are expressing acceptance of the process in general or certain parts of it. The rest of evaluative statements express rejection. In total, about a third of all statements are indifferent, i.e. they are not giving evaluative statements on the process. These encompass the demographic descriptions of the interviewees along with their knowledge on RDM so far, descriptions of their projects used as examples, and statements on other topics on RDM. Furthermore, internal annotations were made to point out passages of interest amongst the reviewers. An overview is given in Table 2.

Code Group	**Acceptance**	**Rejection**	**Indifference**	**Total group sum**
Overall process	31	7	0	38
Project management level	92	40	11	143
Work package management level	74	22	8	104
Research data management level	73	28	10	111
Publication workflow	55	46	31	132
Other topics sum	0	0	213	213
Project description	0	0	46	46
RDM-application	0	0	27	27
Demographics	0	0	23	23
Possible support structures	0	0	14	14
Project structures	0	0	12	12
Influencing factors	0	0	7	7
Others	0	0	16	16
Internal annotations	0	0	68	68
Total sum	325	143	273	741
Total sum percentages	43.9%	19.3%	36.8%	100%

Table 2: Codes on the feedback received

4.2 General feedback on the process

The general perception of the process is very positive. 31 of 38 statements, or 81.6%, of the interviewees are positive about the fit of the process to their day-to-day work. One consensus by the interviewees was the positivity about how the structure of the process is set up from the whole project to the smallest units of research. As pointed out in section 2.1, the division of complex problems into smaller parts is inherent to engineering sciences. One statement in particular summarised this:

> *“ In other words, we try to break this task down into smaller subtasks. (...) This often works relatively well because this entire engineering structure is actually organised in this way. In other words, when we look at a system, I can actually always break this system down further into smaller subsystems with corresponding system boundaries etc. and actually go deeper, deeper, deeper until I just have a single element.”* - Interviewee

The combination of waterfall-like structures with iterative elements was perceived as a strength of the process. Along with this, the linking of artefacts to an enriched research product were praised. Also, the completeness and end-to-end approach for research projects was acclaimed by the interviewees.

As an attention point on the proposed process, to some degree, interviewees see a discrepancy between the proposed process and the reality of research. While the process aims to depict the practice of research, it also aims to provide a *good practice* example on how the processes should look like. This however can never be expected from reality and if the process matches reality, it will rather be by chance than by careful planning, because there are many internal and external influencing factors. Some interviewees argued that these factors necessitate more iterations in the process. In contrast, others pointed out the research proposal would hold all the work packages already, which is why there would be no iteration at all on the work package level. In any case, interviewees claimed that the process should not enforce a strict timeline, as this would increase overall stress.

Additionally, there were conflicting statements, when exactly research results are generated, what these are and what form they come in. In some interviewees’ opinions, only research results in form of manuscripts were relevant, others saw intermediate results as results already. Furthermore, some statements hint towards results being already generated before the research proposal is accepted. These are however often not sufficient for publishing but are rather quick, but not neat solutions to prove feasibility.

Another attention point regarded the wording of certain parts of the process. For instance, while there is a research data management level, still every level has its own RDM tasks. This causes unnecessary confusion and needs to be addressed. Specifically the expressions of artefacts and findings have to be concretised and enriched with examples. Similarly, parts of the proposed process should be made clearer.

Besides the overarching demarcations on the process, there were also specific statements on how certain parts of the process could be improved or do not fit the research process of the interviewees. These rejecting statements are further discussed in the following sections to improve the proposed process. Positive statements will mostly be neglected as they offer little to no impetus for improvement.

4.3 Changes to key wordings: artefact and finding

Artefacts remain the same as in our previous publication. [1] However, they are specified more and implications are made explicit. Additionally, the demarcation of two artefacts are depicted more clearly.

An artefact is the result of one performance of the artefact level. It is the result of a single method applied, for example: When a test bench for an experiment is designed, the plans for it are one artefact (A1). When it is build, the test bench itself is another artefact (A2), building upon the first artefact (A1). To measure something on this test bench, a software script has to be written to record data from sensors. This script is an artefact as well (A3). The recording of raw data is a new artefact (A4), resulting from artefacts (A2) and (A3). Another script (A5) has to be programmed to analyse the recorded raw data (A4). The analysed data (A6) then answers the research question. This example linking of artefacts is depicted in figure 10. How exactly artefacts are linked is depicted in figure 11.

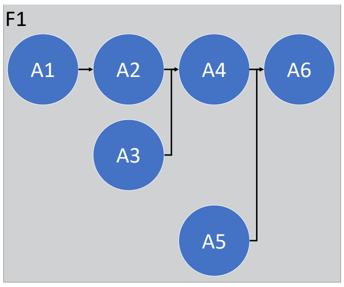


Figure 10: Linking of artefacts in the proposed process

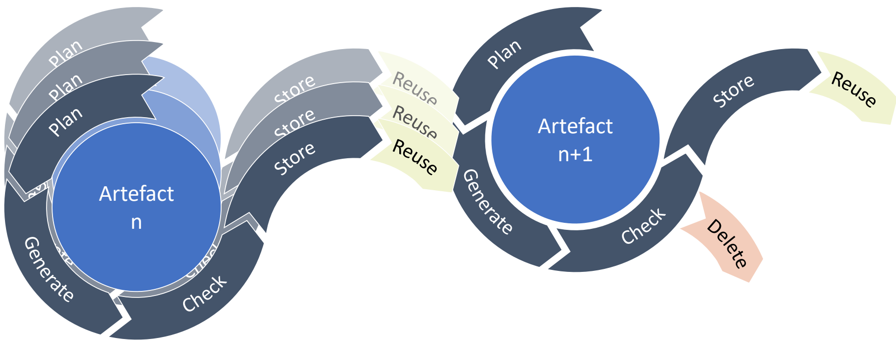


Figure 11: Procedure of artefact linking: Many to one relation

An artefact may be a primary data collection, a physical object or the implementation of source code, or a collection of secondary data, e.g. from a systematic literature review, as it follows the definition of research data by the German Research Foundation (DFG):

> " *Research data includes measurement data, laboratory values, audiovisual information, texts, survey or observation data, methodological test procedures and questionnaires. Compilations, software and simulations can equally represent a central result of scientific research and are therefore also included under the term research data. Research data in some subject areas is based on the analysis of objects (such as tissue, material, rock, water and soil samples, test specimens, installations, artefacts and art objects), so its handling must be just as careful and consideration must be given to a technically adequate option for subsequent reuse whenever meaningful and possible. Should subsequent reuse of the resulting research data be closely associated with objects, then please also elaborate on this by providing all relevant information.* " [27]

To elaborate further: As defined by North [28] with the model of the knowledge stairway, data is a number of characters (numbers, letters, special characters) that are related to each other by syntax (ordering rules). The metadata recorded along with the generation of the data adds context - i.e. a meaning - to it, so data becomes an artefact, which is on the information step at the knowledge stairway.

No changes were made to the definition of findings. However, and according to North's knowledge stairway [28], the finding compilation was more explicitly focussed on the generation of knowledge: If several pieces of information (artefacts) are networked and related to each other, knowledge is created. So a finding is knowledge, as, according to North [28], knowledge is a semantically interconnected information. As a finding is compiled by the interconnection of artefacts, which are bits of information, it is knowledge.

4.4 Feedback on the project management level

One of the most controversially discussed parts of the proposed process is the project management level, now adjustified to project level. Table 3 shows the codes found for this level in general as well regarding certain parts of it.

Code Group	Acceptance	Rejection	Indifference	Total group sum
Project management level	92	40	11	143
General	17	7	5	29
Project planning	30	13	5	48
Execution	37	8	1	46
Completion	8	12	0	20

Table 3: Codes on the feedback received on the project management level

Most of the attention points on the project level focussed on its waterfall-like design. Some interviewees claimed, that it would still be iterative. These comments especially refer to the demarcation between execution and completion. Furthermore, some interviewees annotated, that this distinction is not as clear as between project planning and execution, while others emphasised it particularly as clear-cut. The acceptance of a research proposal marks the transition. A research proposal may be not just a funding proposal but any proposal made to anyone to conduct a research project, e.g. to industry partners. It could also be seen as, for example, a proposal for a dissertation made to a professor. The acceptance of a research proposal or the determination of the projects completion phase can be seen as a kind of quality gate, similarly to the concept of quality gates promoted in Prefi [29] or Hawlitzky [30]: When a quality gate is reached in the process (e.g. the transition from project planning to execution), a check is performed to determine whether the activities carried out so far are sufficient (e.g. acceptance of the research proposal). If not, the quality gate cannot be passed, meaning activities must be repeated before the next phase can begin.

Every part of work performed before the acceptance is seen as preparatory work: may be an early start, if the project is very likely going to be accepted, may be "quick and dirty" data, not worth of publishing, yet able to give away findings. Preparatory work follows the principles of content-wise execution in its approach. Consequently, the data generated in preparatory work can be handled the same way as data created in the execution. Furthermore, data used in preparatory work can also be a result of a predecessor project, which means, artefacts are reused for the new project's planning.

For the transition between execution and project completion, the interviewees agreed that research is a never ending process; its completion is only achieved when the relevant demands are satisfied. There will always be something to look deeper into or reconsider or research more on. Hence, the transition to the project completion phase takes place once the supervisors for the ongoing project management determine that one or more of the following conditions are met:

- The findings generated in the project suffice to answer the overarching research question
- The project's time frame has ended
- The project funding has run out

In the project completion, the ensuring of reproducibility is an important task as it has to be seen as re-considering the work done and validating that it was done correctly: The term ensure was specified and means, that a verification and affirmation, if an activity has been performed, is conducted. If the activity was not conducted at the point of ensuring it, it has to be performed latest at this point when the ensuring takes place. Ideally, the ensured activity was performed beforehand.

In the execution phase only few attention points were brought up by the interviewees. Most importantly, the ongoing project management has to be specified. It holds not only typical tasks like comparing the project plan to the current status but also encompasses the writing of intermediate reports and considerations regarding follow up projects. It also has to be considered that unforeseen challenges arise, demanding adaption of the original plans. The ongoing project management defines the next finding that should be generated based on the work done with previous findings, which are adapted, if needed. This decision is done at the project management when a finding is compiled.

Based on the interviewees, ensuring reproducibility can be challenging to impossible some times. Some experiments are unique and will never be able to be reproduced, e.g. the recording of a supernova[4]. Hence, the exact wording of this point has been changed. It also has to be annotated that some interviewees do not see the ensuring of reproducibility as important. This can be condensed to one statement in particular: "*What you haven't done by then, you won't be able to save later on either.*" Still, ensuring of reproducibility is an important task, yet coming back to the aforementioned wording of *ensure* as re-considering the work done and validating that it was done correctly.

Furthermore, the interviewees annotated that the project completion phase often already considers follow-up projects. For instance, if the prolongation of a project is already considered guaranteed, the project completion is not seen as strict as without a follow-up project.

Also, the interviewees pointed out that publications could be started at any given stage of the research project. Specific triggers for publications are discussed in section 4.7. After its redesign, the new process for the project level of the proposed research process is shown in figure 12.

4. Example changed due to anonymisation.

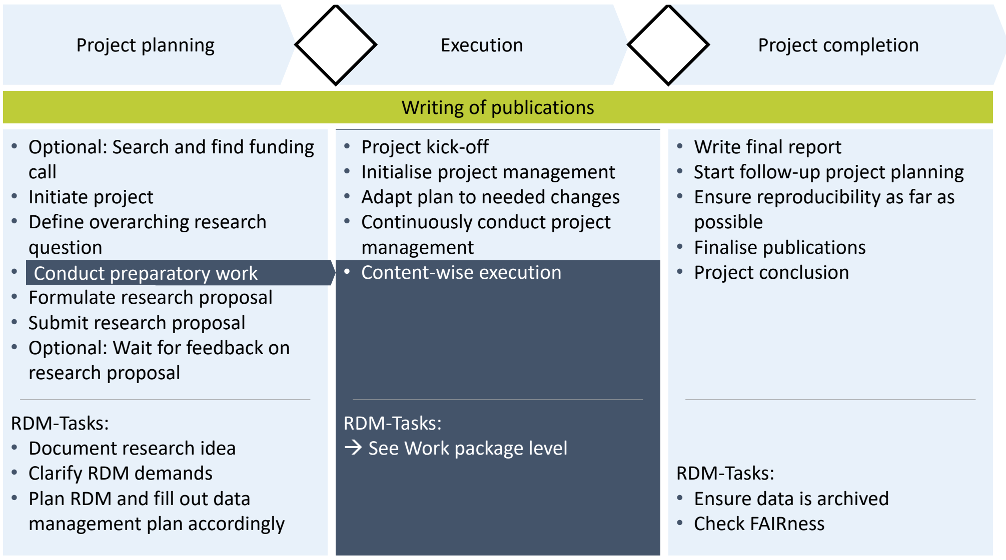


Figure 12: The redesigned project level of the proposed research process

4.5 Feedback on the work package management level

The work package management level also offers some opportunities of improvement, both in wording as well as overall process structure. The discussions with the interviewees focussed on two points in particular. On the one hand, they empathized the importance of this level being of an iterative nature, as sometimes findings will result in discovering new directions worth investigating. On the other hand, the finding compilation is seen as debatable in the proposed state. Table 4 shows the codes found for the level.

Code Group	Acceptance	Rejection	Indifference	Total group sum
Work package management level	74	22	8	104
General	28	5	5	38
Finding planning	13	5	0	18
Create or Reuse	22	2	2	26
Finding compilation	11	10	1	22

Table 4: Codes on the feedback received on the work package management level

The work package management level is seen under a completely new perspective: As the interviewees pointed out, the former designation does not properly represent the nature of research. Rather, the level needs to be separated from the perspective of work packages and has to focus on the finding aspect of research instead. This way, uncertainties with wording and the scope of the level are addressed: The level is now called finding level.

The finding level has an iterative nature as the findings compiled at the end of the level are reported back to the project level. This may either complete a work package or research question, rise a new research question, cause the need for further investigation or all of the aforementioned. It is the only level with an iterative nature. Hence, the work package perspective is now included in the ongoing project management. This also specifies, that although work packages are defined in the research proposal, research is still prone to unforeseen findings and changing plans during the project. Furthermore, this level allows for multiple parallel findings worked on simultaneously by the same or different researchers.

The plan approach to finding step is triggered by the supervisors of the ongoing project management. Firstly, a question to be answered is selected or formulated (see plan approach to finding). To answer this question, needed sub-steps are defined, each with an artefact as outcome. These artefacts are created or reused as described in section 4.3. After enough artefacts have been generated or reused, a finding can be compiled. This finding can be a significant discovery or just the fact, that the research question can not be answered with the chosen approach. In both cases, the finding is reported to the ongoing project management. Optionally, a finding may be worth publishing, either as data, software or manuscript, e.g. in form of a scientific article (cf. Feedback on publication workflow of the process).

In the finding compilation, artefacts are linked to each other. Following the definition of North [28], new knowledge is created by this linking process. As data plus its meaning (metadata) form artefacts, these artefacts are information. The linking of information forms knowledge. [31] The new knowledge, whether expected, anticipated or unforeseen is reported to the ongoing project management, leading to the answering of a research question or further investigation. If this knowledge is sufficient to answer the (sub-)research question selected, this (sub-)research question can be seen as solved. If the knowledge is insufficient, more artefacts are needed, and are therefore created or reused. This causes a new loop on the finding level. This nature of research of having to search again for something that was not yet found was well described by one interviewee:

> *" That's why we call it* ***re****search. We look for something and then realise we have to look for it again because we haven't found it. "*

Additionally, at the beginning and the end of such a research loop on the finding level, the DMP is updated to document the intention of artefact and finding creation as well as its outcome. The updated finding level is depicted in figure 13.

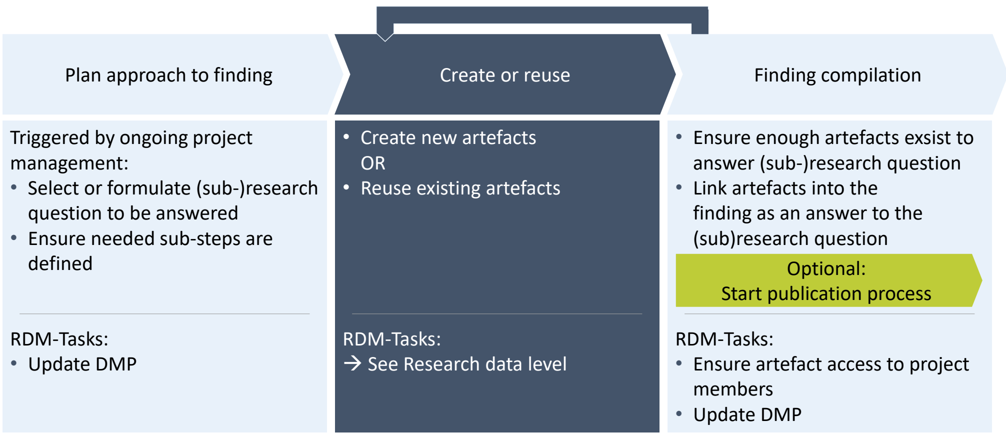


Figure 13: The redesigned finding level of the proposed research process

4.6 Feedback on the research data management level

On the research data management level, interviewees mostly agreed. Their biggest concerns were clustered at the *check* step. This can also be seen in table 5.

Code Group	Acceptance	Rejection	Indifference	Total group sum
Research data management level	73	28	10	111
General	33	6	7	46
Plan	8	7	2	17
Generate	10	2	1	13
Check	19	12	0	31
Store	3	1	0	4

Table 5: Codes on the feedback received on the research data management level

The wording was adjusted to fit the finding level. As the research data management level is focussed on the generation of artefacts, it is renamed to artefact level.

The artefact level happens to be of a very short nature in general, compared to the duration of work packages, as most activities can be completed within a few hours. Yet, the generation of data may take up to days or even weeks depending on the setup. The tasks performed by researchers are able to be performed and finished immediately and without accompanying activities. Hence, the artefact level is not iterative, but straight forward. Once performed, the outcome is either stored or deleted. In both cases the result is taken to the finding compilation to consider if more artefacts need to be generated or the new knowledge generated is sufficient to answer the research question.

In the preparation phase of the data generation, the wording is adjusted to fit the actual scope of the step, namely the generation setup. The data generation is set up, meaning that the exact procedure of data generation is planned and checked. This, for example might be the inspection of a test bench, if all cables are connected correctly and if the software is set up correctly for the recording of data. For interviews, this might also be the check if the recording devices are running properly, the interview guideline is present and known to the interviewers. Furthermore, input artefacts are selected. This step happens moments before the generation of data.

Then data is generated. Ideally, metadata is automatically recorded with data generation. This might encompass environmental conditions, labels for data or demographic data. The form of metadata to be used is highly dependant on the subject of investigation and the (sub-)research question to be answered.

Afterwards, the data is checked for correct recording, i.e. correct set up of the data generation. So this step is not a validation of data but ensures data correctness. Although interviewees stated that this step was often not performed, we recommend it highly to filter out unusable data right away. If the data is corrupted in any way, e.g. a sensor was not connected or malfunctioned, the data can be deleted, because no information can be drawn from this failed data generation.

If the data was recorded correctly, it can be stored along with the corresponding metadata. The artefact should be stored in a location that is backed up and ideally offers access to project members. The updated artefact level is shown in figure 14.

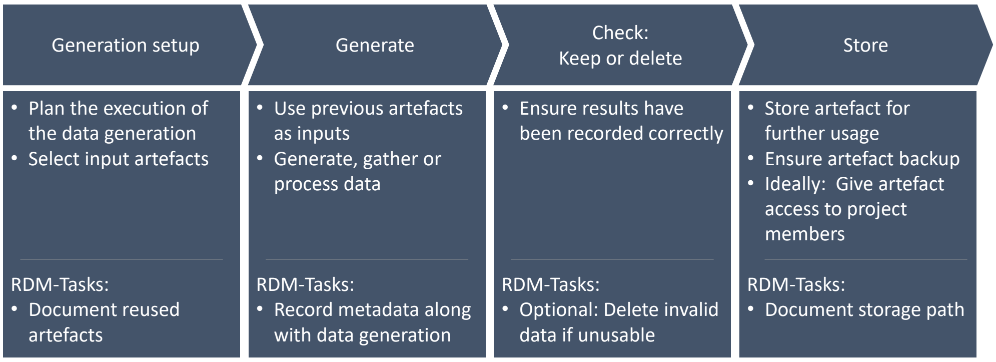


Figure 14: The redesigned research data management level of the proposed research process

4.7 Feedback on publication workflow of the process

The interviewees generally found the publication process to be appropriate. Nevertheless, comments such as "*[the publication process] basically depicts reality as it is and should be*" illustrate the discrepancy between theory and practice. For this reason, isolated differences to the model process emerged several times in the interviews. They can be attributed to time pressure and external factors (such as project partners or funding regulations) in day-to-day research. Table 6 shows how this feedback was distributed over the proposed phases.

Code Group	Acceptance	Rejection	Indifference	Total group sum
Publication workflow	55	46	31	132
General	10	11	25	46
Topic	11	13	2	26
Preparation	12	6	1	19
Editing	12	8	2	22
Submission	8	7	1	16
Dissemination	2	1	0	3

Table 6: Codes on the feedback received on the publication workflow

Evaluating the first two phases, a need for a clearer distinction has emerged: While the preparation phase sets formal framework conditions, the focus in the topic phase is on planning the content. So they are renamed to topic and scoping and formal preparation. In addition, a publication is not only triggered by the finding compilation, as previously stated. In practice, there are two types of triggers:

- Self-motivated publication triggers, e.g.: intrinsic motivation of researchers to share a finding that they have compiled; gaining an overview of the state of the art in the context of a review paper; turning a thesis into a paper together with the student
- Externally motivated publication triggers, e.g.: invitations to conferences and associated deadlines; finalising the publication plan from the project proposal; doctoral degree regulations; writing a cumulative dissertation; supervisors push for publication

Even if there are various triggers for publications, it remains clear that finding compilation is essential for this process:

> *“ Let me put it this way: If I don’t have any findings, then it will be difficult [...] writing a publication. But of course I can try to make findings possible in a more targeted and less targeted way. You know exactly that some things might work and then you do something in that direction, because you know that if you do this or that, then I will at least manage to generate enough new knowledge to at least justify a publication. (...) At the end of the day, of course, it would be nicer if I had a specific research question that I try to answer or was trying to answer and then say: Hey, now the research question has been answered and now I’m going to write a publication about it. Yes, there are – in practice, I would say there are both. ”*

This description fits with interviewees’ statements, that writing publications is a reflective activity, where research data and text production are closely linked. When converting artefacts and findings into a written form (to create a medium that passes on the knowledge from the finding compilation to other people), gaps may become visible that need to be filled with additional content, or new contexts may emerge.

> *“ I always find that it all makes a lot of sense in my head. But it’s only when you put it down on paper that you realise: Oh no, that’s somehow, there’s no thread running through it. So that’s actually always a point where you have to think more deeply about it when you’re writing what have you actually just done here, because you have to get your own thoughts down somehow, and so there’s a lot of reflection involved. ”*

At the same time, the interviewees emphasised that there are also publication types such as concept papers that contain less data. Interviewees stated that publications could occur at any time over the course of the whole research project. Another aspect that the interviewees repeatedly mentioned was preprints: By publishing publications that have not yet been peer-reviewed, they want to position themselves quickly in the dynamic academic discourse and make their work referencable so that they can be followed up in subsequent publications.

Finally, the interviewees addressed the communication of scientific results to the public - usually referred to as “science communication” - which is becoming increasingly established in the job description of researchers [32]. Communication within the scientific system is correspondingly “scholarly communication”. (cf. e.g. [33], [34]) Since these two sub-areas of publishing differ in terms of target group and communication goal and form, but are similar looking at the basic process, the revised graphic is divided into two parts, as shown in figure 15.

The content on science communication is essentially based on the explanations of Konneker [32], Frick and Seltmann [35] and Bertemes et al. [36]. A key difference between science and scholarly communication lies in the support. Depending on the involvement of the respective institution’s communication department, process steps in the area of science communication (e.g. parts of editing or submission) are handled by the department rather than by the researchers themselves - in contrast to scholarly communication.

In addition to the highlighted aspects so far, the publication process has undergone further minor adjustments, such as the addition of internal reviews and (external) approval processes.

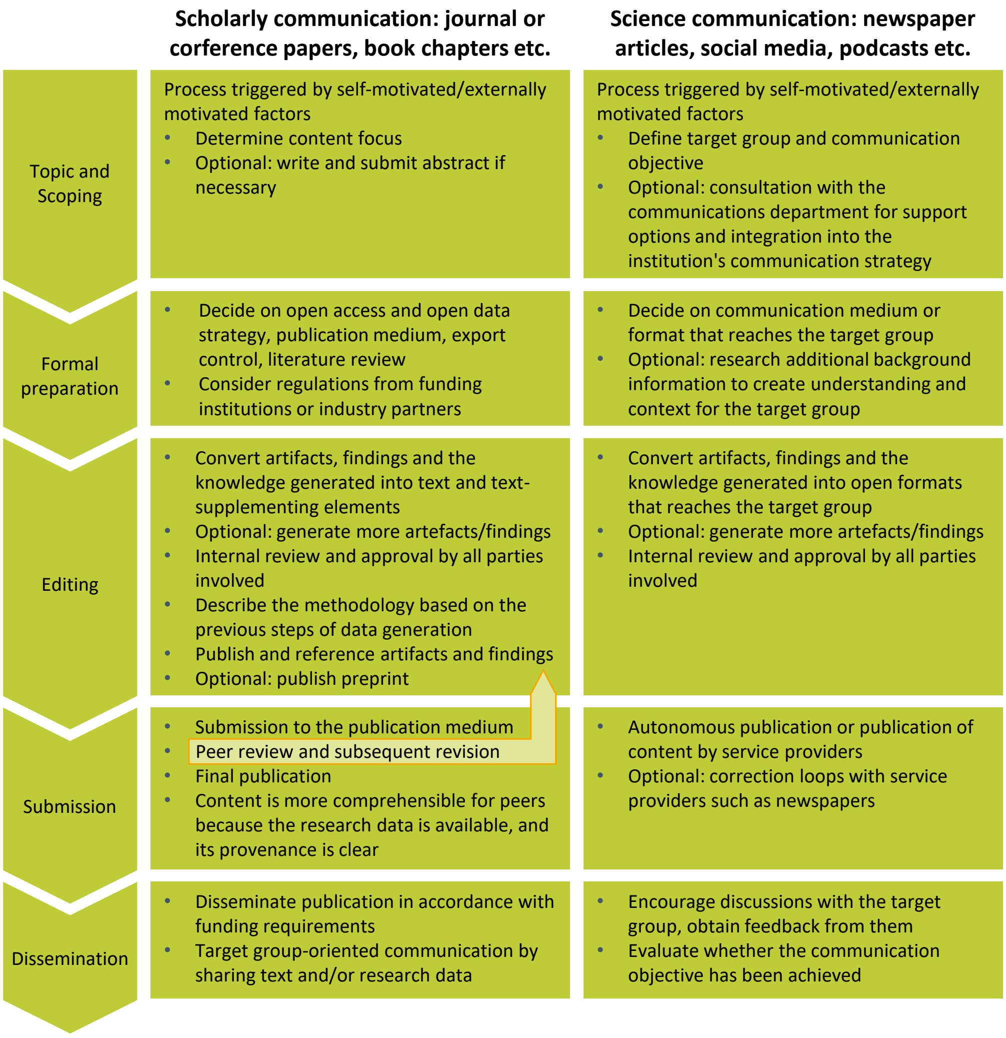


Figure 15: The redesigned publication process

4.8 Overall process

The overall process was met with general approval by the interviewees. Still, the refinements made to the process on the basis of the interviewees feedback improved it significantly. Wordings have been specified, steps and tasks have been made more explicit and the structure of the process was adapted to the reality of research while also upholding standards of RDM. The final process is depicted in figure 16.

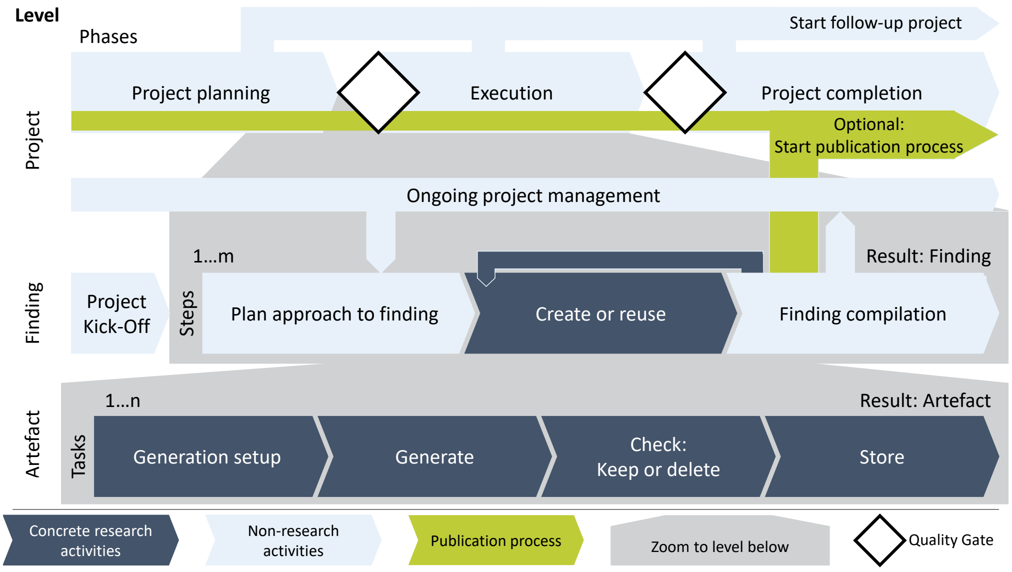


Figure 16: General structure of the proposed RDM process

In comparison to figure 3, the biggest changes are:

1. The option to start follow-up projects in any of the process' phases
2. The option to start a publication process in any of the process' phases
3. The option to generate findings in the project planning phase
4. The project kick-off was moved to not be included in the content-wise execution
5. A feedback loop between finding compilation and plan approach to finding
6. A feedback loop between finding compilation and create or reuse
7. Wording changes, especially on the names of the levels:
   - Project management level changed to project level
   - Work package level changed to finding level
   - Research data management level changed to artefact level

This process validated the previously identified needs, demands, and practices of researchers from engineering sciences. It aims to resemble the day-to-day research activities of researchers in engineering sciences as best as possible while also integrating RDM activities as seamlessly as possible.

5 Discussion and Limitations

The answer to the research question can be given with a high certainty. The evidence for this statement is high, as it is based on the consolidation of eight problem-centred interviews with 19 researchers, the surveying of 13 research projects by another eight researchers on their workflows and this validation with another ten researchers in expert interviews. This total of 37 different researchers questioned on their research's demands and structure.

While the first two methods were biased on the RWTH and Leibniz University Hannover (LUH), the validation took into consideration the whole TU9. While, smaller universities or universities of applied sciences might have different approaches to research, the approach of validation via TU9 allows for insights into a big part of engineering sciences.

Furthermore, the changes and adaptions to the process, while seemingly big, are not crucial. The biggest changes were made due to wording inconsistencies. Certain parts of the process were explicated and defined more precisely, or in some cases, more openly to allow for a better situational fit. The overall process structure with its levels and their contents remain unchanged.

The validation, which was urgently missing in the last paper, was now performed. It confirmed the good reception the process got in several events where it was presented (see [37], [38], [39], [40]). Overall, it affirmed the process with minor changes made to it.

Nevertheless, the methods used should be critically contextualized: The usual quality criteria such as objectivity, reliability and validity in the sense of definitions from natural sciences can only be applied to expert interviews to a limited extent. For example, the interviews themselves take place in social situations that create a dynamic between interviewers and interviewees or even unintentional mutual influences, which can lead to bias. [41] The objectivity of qualitative content analyses - a method which core is the interpretation of the material by people with different knowledge and views - can at least be strengthened by instruments such as intercoder analysis. [25]

6 Conclusion and Outlook

In this article, we validated our proposed process for project-oriented RDM in engineering sciences. This process, now called Project- and RDM-Oriented Research Process for Engineering Sciences (PRO Research Process 4Ing), is the result of a thorough investigation of the problems, requirements, workflows and practices of engineering researchers. In total 37 different researchers were consulted by the means of different methods.

PRO Research Process 4Ing is validated in this article by a series of ten expert interviews. The result is a validated process that received minor adjustments to better fit the reality of day-to-day work and facilitate understandability. This also leads the the answer to the proposed research question:

Can our proposed RDM process be deployed into engineering research projects?

With the changes made to it, yes.

PRO Research Process 4Ing will provide the foundation of Jarves, the Joint Assistant for Research in Versatile Engineering Sciences [5]. This digital data steward guides engineering researchers in their RDM throughout their whole research project. Jarves will also contain the management of RDM guidelines within a decision support system. [1]

5. https://jarves.nfdi4ing.de/

PRO Research Process 4Ing offers a missing link between theoretical RDM and actual research practices. It offers a bridge between research projects and their structure and the life cycles of research data. We hope, that PRO Research Process 4Ing facilitates the application of RDM in engineering sciences. Additionally, in future research PRO Research Process 4Ing could be transferred to other disciplines that utilize project-oriented research.

A Appendix

This appendix contains a glossary on the most important terms in this paper and the process it describes. This glossary aims to facilitate understandability for those who would like to implement, utilise or further develop this process. The glossary is an updated version of the previous paper's glossary. [1]

Glossary

Analysed data	Any data that is processed in any way that allows for new information to be derived from it. **On artefact level. Straight forward.**
Artefact	The result of one iteration of the artefact level. An artefact is planned, contextualised and validated research data in its smallest unit. It may be a primary data collection or the implementation of source code. It was generated by creating a generation setup, generating the data and performing a check to ensure the correct recording and, if valid, storing the data. An artefact is information. **On artefact level. Straight forward.**
Artefact level	The level on which artefacts are generated. It contains the tasks generation setup, generate, check and store. It has a straight forward structure with distinct tasks following one after another. **On artefact level. Straight forward.**
Check	Check the generated data for corruption. This task ensures data correctness. **On artefact level. Straight forward.**
Compile finding	Term from the previously designed process. See now finding compilation. **On finding level. Iterative.**
Completion	Term from the previously designed process. See now project completion. **On project level. Waterfall structure.**
Content-wise execution	All activities on the finding level containing the generation of findings and knowledge. **On finding level. Iterative.**
Core research activities	All activities on the artefact level containing the generation of research core data. **On artefact level. Straight forward.**

Create or reuse	The step of creating a new artefact or reusing an existing one by importing it into the current project. The creation of a new artefact can be done by using any amount of existing artefacts. **On artefact level. Straight forward.**
Dissemination	Last of the five publication process phases. Dissemination includes all activities that promote the distribution of a publication, such as inclusion in the personal, publicly accessible publication list, social media posts or even personal recommendations via mailing lists. **On publication process. Straight forward.**
Editing	The third of the five publication process phases. Can only be started if the topic and scoping phase is finished. Includes all activities that revolve around the actual production of text. **On publication process. Iterative.**
Ensure	Ensuring an activity has been performed and, if not, performing it latest at this point when the ensuring takes place. Ideally, what is ensured was performed beforehand.
Execution	The second of the three research project phases. Can only be started if the acceptance of the research proposal. **On project level. Waterfall structure.**
Externally motivated publication trigger	Trigger for publication initiated by the external parties or factors, e.g.: invitations to conferences and associated deadlines, finalising the publication plan from the project proposal, doctoral degree regulations, writing a cumulative dissertation, supervisors push for publication
Finding	The result of one iteration of the finding level. It may be the result of one work package or the answer to one research question. It was generated by planning what should be found out, compiling one or several artefacts and hereby deriving new knowledge from this process. A finding is knowledge. **On finding level. Iterative.**
Finding compilation	Compile a finding from artefacts. In this way, new knowledge is generated. **On finding level. Iterative.**
Finding level	The level on which findings are generated. It contains the steps plan approach to finding, create or reuse and finding compilation. **On finding level. Iterative.**
Finding planning	Term from the previously designed process. See now plan approach to finding. **On finding level. Iterative.**
Formal preparation	Second of the five publication process phases. Formal preparation includes all activities that relate to the formal framework conditions of the publication process and trigger corresponding planning. **On publication process. Straight forward.**

Generate	In the generate task, new artefacts are brought into existence. This can happen by using existing artefacts or without them. New data is generated, gathered or gained by processing of existing artefacts. Due to the annotation of the data with metadata, meaning is provided, adding meaning to the data and therefore turning it into information. **On artefact level. Straight forward.**
Generation setup	The data generation setup happens immediately before the generate task. It is the planning and affirmation of the exact procedure of data generation. Furthermore, input artefacts are selected. **On artefact level. Straight forward.**
Ongoing project management	Over the course of the project, the ongoing project management tracks the process of the project in accordance to the project plan. The ongoing project management defines the next finding that should be generated based on the work done with previous findings, which are adapted, if needed. It also contains other project management tasks, such as the writing of reports and project meetings. **On project level. Waterfall structure.**
Plan	Term from the previously designed process. See now generation setup. **On artefact level. Straight forward.**
Plan approach to finding	When the supervisors trigger the investigation of a specific topic, a corresponding research question is either selected (if existent, e.g. in the proposal) or defined (if not existent yet). Additionally, needed sub-steps to answer the research question are defined. **On finding level. Iterative.**
Planning of the data generation	Term from the previously designed process. See now generation setup. **On artefact level. Straight forward.**
Preparation	Term from the previously designed process. See now formal preparation. **On publication process. Straight forward.**
Preparatory work	Preparatory work is content-wise work conducted in a research project before the research proposal is submitted. In preparatory work, content-wise execution can already be conducted and hence, findings can already be generated. **On project level. Waterfall structure.**
Project	Unique activity with a defined outcome and predetermined resources in terms of costs and time [42]. **On project level. Waterfall structure.**
Project completion	The last of the three research project phases. Can only be started if the execution is finished. **On project level. Waterfall structure.**

Project kick-off	At the beginning of the execution phase, this step occurs to actually start the project after the sometimes long time passed since the project planning phase. Here, the project team is formed if not yet existent. Also, the proposal is consulted to get an overview over the goals and needed approach. **On finding level. Straight forward.**
Project level	The level on which projects are conducted. It contains the phases project planning, execution and project completion. It has a waterfall-like structure with distinct phases. **On project level. Waterfall structure.**
Project management level	Term from the previously designed process. See now project level. **On project level. Waterfall structure.**
Project-oriented research	Research with a planned and defined outcome and predetermined resources in terms of costs and time (c.f. [42]).
Project planning	The first of the three research project phases. The research project starts with this phase. **On project level. Waterfall structure.**
Publication	A publication is a written record of a finding. The knowledge that is generated during the finding compilation can be passed on to other people through this medium.
Publication process	Independently running process that can start at any given stage of the research project. Its result is a publication that contains the knowledge generated during the finding compilation.
Raw data	Any data recorded or created, but not yet analysed. Does not generate any new knowledge. **On artefact level. Straight forward.**
Research core data	The research data collected in a research project, which generates or is used to generate new knowledge. **On artefact level. Straight forward.**
Research data	“Research data includes measurement data, laboratory values, audiovisual information, texts, survey data, objects from collections or samples that are created, developed or analysed in the course of scientific work. Methodological test procedures such as questionnaires, software and simulations can also represent key results of scientific research and should therefore also be categorised as research data” [27]. **On artefact level. Straight forward.**
Research data management (RDM)	The handling of research data (e.g. collection, organization, storage, and documentation) during and after a research process.
Research data management level	Term from the previously designed process. See now artefact level. **On artefact level. Straight forward.**
Research project	Project to investigate a specific research question. **On project level. Waterfall structure.**

Research project data	The data used to describe the information about a research project without the creation of new knowledge.
Research proposal	A proposal for a research project. This can but does not need to be a funding proposal. It can also be a proposal of a dissertation project to a professor or a proposal of an industry project to an industry partner. **On project level. Waterfall structure.**
Scholarly communication	Scholarly communication is a type of communication in the publication process. It takes place within the scientific system by researchers who disseminate their findings peer to peer. **On publication process.**
Science communication	science communication is a type of communication in the publication process. Researchers communicate their findings to the public, possibly with the support of communications departments. **On publication process.**
Self-motivated publication trigger	Trigger for publication initiated by the researcher, e.g.: intrinsic motivation of researchers to share a finding that they have compiled, gaining an overview of the state of the art in the context of a review paper or turning a thesis into a paper together with the student.
Store	When the artefact was checked, it can be stored. Ideally, it is stored in a way that project members can access it. **On artefact level. Straight forward.**
Submission	This phase comprises the activities surrounding the submission of the publication, e.g. to the publisher. The core is the peer review, which ensures scientific quality. **On publication process. Iterative.**
Topic	Term from the previously designed process. See now topic and scoping. **On publication process. Straight forward.**
Topic and scoping	This phase is triggered by either a self-motivated publication trigger or a externally motivated publication trigger. The availability of a finding is regardless of this crucial. In this phase, the content of the publication is determined. **On publication process. Straight forward.**
Work package level	Term from the previously designed process. See now finding level. **On finding level. Iterative.**

B Acknowledgements

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.

C Roles and contributions

Tobias Hamann: Idea, Conceptualization, Methodology, Writing

Michèle Robrecht: Conceptualization, Methodology, Writing

Marcos Alexandre Galdino: Writing - Review

Anas Abdelrazeq: Writing - Review

Robert H. Schmitt: Writing - Review, Supervision, Funding acquisition

References

[1] T. Hamann et al., "Matching Data Life Cycle and Research Processes in Engineering Sciences. To be published," *ing.grid*, 2024. Accessed: Sep. 10, 2025. [Online]. Available: https://preprints.inggrid.org/repository/view/42/

[2] A. Wuchner, M. Robrecht, P. Kehl, and H. R. Schmitt, "Challenges in publishing research data – a Fraunhofer Case Study. To be published," *ing.grid*, 2023. Accessed: Sep. 10, 2025. [Online]. Available: https://preprints.inggrid.org/repository/view/20/

[3] D. Ortloff, S. Anger, and M. Schellenberger, "An Empirical Study of the State of Research Data Management in the Semiconductor Manufacturing Industry: An analysis of industry and research institutions in the iDev4.0 project. To be published," *ing.grid*, 2023. Accessed: Sep. 10, 2025. [Online]. Available: https://preprints.inggrid.org/repository/view/19/

[4] ELIXIR-CONVERGE, *All training resources*, 2023. Accessed: Sep. 10, 2025. [Online]. Available: https://rdmkit.elixir-europe.org/all_training_resources

[5] M. Politze, *A reference architecture and implementation enabling data protection in distributed elearning and escience processes: Eine Referenzarchitektur und Implementierung unter Berücksichtigung des Datenschutzes in verteilten eLearning und eScience Prozessen* (Ergebnisse aus der Informatik). Aachen: Apprimus, 2019, vol. 10. DOI: 10.18154/RWTH-2019-11476

[6] S. Higgins, "The DCC Curation Lifecycle Model," *International Journal of Digital Curation*, vol. 3, no. 1, pp. 134–140, 2008. DOI: 10.2218/ijdc.v3i1.48

[7] A. M. Cox and W. W. T. Tam, "A critical analysis of lifecycle models of the research process and research data management," *Aslib Journal of Information Management*, vol. 70, no. 2, pp. 142–157, 2018. DOI: 10.1108/AJIM-11-2017-0251

[8] W. Krull and A. Tepperwien, "Neue Herausforderungen für die öffentliche und private Forschungsförderung," in *Handbuch Wissenschaftspolitik*, ser. Springer Reference Sozialwissenschaften, D. Simon, A. Knie, S. Hornbostel, and K. Zimmermann, Eds., Wiesbaden: Springer, 2016, pp. 447–463. DOI: 10.1007/978-3-658-05455-7_17

[9] N. Winterhager, *Drittmittelwettbewerb im universitären Forschungssektor* (Organization & Public Management). Wiesbaden: Springer, 2015. DOI: 10.1007/978-3-658-07188-2

[10] H. P. Bonn, “Die Vernetzung zwischen Hochschulforschung und Industrie: Facetten und Vorteile für den Wirtschaftsstandort,” in *Zukunft Ingenieurwissenschaften – Zukunft Deutschland*, M. Nagl, H.-J. Bargstädt, M. Hoffmann, and N. Müller, Eds., Berlin and Heidelberg: Springer, 2009, pp. 111–120. DOI: 10.1007/978-3-540-89609-8_12

[11] E. Berning and S. Falk, “Das Promotionswesen im Umbruch,” *Beiträge zur Hochschulforschung*, vol. 27, no. 1, pp. 48–73, 2005. Accessed: Sep. 10, 2025. [Online]. Available: https://www.bzh.bayern.de/fileadmin/news_import/1-2005-berning-falk.pdf

[12] L. Behlau, *Forschungsmanagement: Ein praktischer Leitfaden*. Berlin and Boston: De Gruyter, 2017. DOI: 10.1515/9783110517828

[13] L. Griem et al., “KadiStudio: FAIR Modelling of Scientific Research Processes,” *Data Science Journal*, vol. 21, 2022. DOI: 10.5334/dsj-2022-016

[14] S. I. H. Shah, V. Peristeras, and I. Magnisalis, “DaLiF: a data lifecycle framework for data-driven governments,” *Journal of Big Data*, vol. 8, no. 1, 2021. DOI: 10.1186/s40537-021-00481-3

[15] forschungsdaten.info, *Der Datenlebenszyklus: Stationen des Forschungsdatenmanagements*, 2023. Accessed: Sep. 10, 2025. [Online]. Available: https://forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/

[16] D. P. Tripathi and S. R. Pandy, “Developing a Conceptual Framework of Research Data Management for Higher Educational Institutions,” in *IEEE 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services [ETTLIS 2018]*, S. Kataria, J. P. Anbu K., R. Gartner, and G. Sandhu, Eds., Piscataway, NJ: IEEE, 2018, pp. 105–110. DOI: 10.1109/ETTLIS.2018.8485193

[17] M. Putnings, “Datenmanagement,” in *Praxishandbuch Forschungsdatenmanagement*, M. Putnings, H. Neuroth, and J. Neumann, Eds., Berlin/Boston: De Gruyter Saur, 2021, pp. 297–302. DOI: 10.1515/9783110657807-017

[18] M. Jagerhorn and X. Chen, “DataCite and FAIR Workflows,” *Procedia Computer Science*, vol. 211, pp. 83–92, 2022. DOI: 10.1016/j.procs.2022.10.179

[19] U. Schwardmann, “Digital Objects – FAIR Digital Objects: Which Services Are Required?” *Data Science Journal*, vol. 19, 2020. DOI: 10.5334/dsj-2020-015

[20] M. L. Wawer, J. Wurst, and R. Lachmayer, “Research Data Management Workflow – A Framework for Engineering Science,” in *Proceedings of the NordDesign 2024 conference*, J. Malmqvist, M. Candi, R. Sæmundsson, F. Bystöm, and O. Isaksson, Eds., 2024. DOI: 10.35199/norddesign2024.79

[21] A. Bogner, B. Littig, and W. Menz, “Introduction: Expert Interviews – An Introduction to a New Methodological Debate,” in *Interviewing Experts*, ser. Research Methods Series, A. Bogner, B. Littig, and W. Menz, Eds., London: Palgrave Macmillan, 2009, pp. 1–13.

[22] U. Froschauer and M. Lueger, “Expert Interviews in Interpretive Organizational Research,” in *Interviewing Experts*, ser. Research Methods Series, A. Bogner, B. Littig, and W. Menz, Eds., London: Palgrave Macmillan, 2009, pp. 217–234.

[23] J. Gläser and G. Laudel, *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*, 3., überarbeitete Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften, 2010.

[24] K. Mozygemba and B. Hollstein, *Anonymisierung und Pseudonymisierung qualitativer textbasierter Forschungsdaten – eine Handreichung*, 2023. DOI: 10.26092/elib/2525

[25] P. Mayring, *Qualitative Inhaltsanalyse: Grundlagen und Techniken*, 13., überarbeitete Auflage. Weinheim and Basel: Beltz, 2022.

[26] U. Flick, *An Introduction to Qualitative Research*, 4th ed. London: Sage Publications, 2009.

[27] German Research Foundation, *Leitlinien zum Umgang mit Forschungsdaten*, 2015. Accessed: Sep. 10, 2025. [Online]. Available: https://www.dfg.de/resource/blob/172112/4ea861510ea369157afb499e96fb359a/leitlinien-forschungsdaten-data.pdf

[28] K. North, *Wissensorientierte Unternehmensführung: Wissensmanagement im digitalen Wandel*, 7., vollständig überarbeitete Auflage. Wiesbaden: Springer, 2021.

[29] T. Prefi, "Qualitätsmanagement in der Produktentwicklung," in *Masing Handbuch Qualitätsmanagement*, T. Pfeifer and R. Schmitt, Eds., München and Wien: Carl Hanser, 2014, pp. 402–427.

[30] N. Hawlitzky, *Integriertes Qualitätscontrolling von Unternehmensprozessen: Methodische Gestaltung eines Quality-Gate-Konzeptes zur Planung, Messung und Steuerung der Prozessqualität: Teilw. zugl.: München, Techn. Univ., Diss., 2002* (Wissenschaft und Praxis). München: TCW Transfer-Centrum, 2002, vol. 20.

[31] DIN Deutsches Institut für Normung e. V., *Qualitätsmanagementsysteme: Grundlagen und Begriffe (ISO 9000:2015); Deutsche und Englische Fassung EN ISO 9000:2015*, Berlin, 2015.

[32] C. Könneker, *Wissenschaft kommunizieren: Ein Handbuch mit vielen praktischen Beispielen: Mit 15 Cartoons von Sarah Zimmemann*. Weinheim: Wiley-VCH, 2012.

[33] S. Hagenhoff, L. Seidenfaden, B. Ortelbach, and M. Schumann, Eds., *Neue Formen der Wissenschaftskommunikation*, Göttingen, 2007. Accessed: Sep. 10, 2025. [Online]. Available: https://publications.goettingen-research-online.de/bitstream/2/14204/1/fallstudien_wikom.pdf

[34] T. W. Burns, D. J. O'Connor, and S. M. Stocklmayer, "Science Communication: A Contemporary Definition," *Public Understanding of Science*, vol. 12, no. 2, pp. 183–202, 2003. DOI: 10.1177/09636625030122004

[35] C. Frick and M. E.-H. Seltmann, "Referenzrahmen für eigenständige digitale Wissenschaftskommunikation durch Forschende," *Zeitschrift für digitale Geisteswissenschaften / Working Papers*, vol. 3, 2024. DOI: 10.17175/WP_2023B_V2

[36] J. P. Bertemes, S. Haan, and D. Hans, *50 Essentials on Science Communication*. Berlin and Boston: Walter de Gruyter, 2024. DOI: 138961

[37] T. Hamann and M. Moser, *Workflows ingenieurwissenschaftlicher Forschungsabläufe: Auszug des aktuellen Forschungsstands*, 2022. DOI: 10.5281/zenodo.7125571

[38] T. Hamann, *Jarves: Der digitale FDM-Assistent: Joint Assistant for Research in Versatile Engineering Sciences*, 2023. DOI: 10.5281/ZENODO.7642481

[39] T. Hamann, *Jarves: The digital RDM assistant: Joint Assistant for Research in Versatile Engineering Sciences*, 2023. DOI: 10.5281/ZENODO.7715683

[40] T. Hamann and J. Werheid, *Jarves: The Digital Data Steward for Engineering Science Research*, 2023. DOI: 10.5281/ZENODO.8378834

[41] A. Bogner, B. Littig, and W. Menz, *Interviews mit Experten*. Wiesbaden: Springer Fachmedien Wiesbaden, 2014. DOI: 10.1007/978-3-531-19416-5

[42] T. Möller and F. Dörrenberg, *Projektmanagement* (WiSorium - Wirtschafts- und Sozialwissenschaftliches Repetitorium). München: Oldenbourg, 2003.