Staying Flexible and Transparent in a Changing World: Agile RDM Service Development

Petar Hristov [1], Katja Jansen [1], Ilona Lang [1]

[1] IT Center, RWTH Aachen University

No matter which area you consider, the world is constantly changing, so requirements must be continuously adapted. This is especially true for Research Data Management (RDM), and particularly for RDM platforms like Coscine. The increasing number of users and external stakeholders, as well as the balancing act between maintenance of existing and the development of new features, represent just some of the challenges faced. In this paper, we illustrate how these challenges can be managed in practice to keep pace with current RDM trends and not miss important developments such as FAIR Digital Objects (Politze et al. 2025) or advancements in querying technologies like SPARQL. We also examine how our anchors – FAIR principles, organization and operational processes – help maintain focus, and what adapting methods to real-world circumstances can look like.

Keywords: Research Data Management, Coscine, FAIR Principles, Software Development, Scrum

1 Introduction

Coscine[1] is a generic platform for research data management (RDM) that has been under development at the IT Center of the RWTH Aachen University since 2018. Within just a few years, the platform has evolved from a start-up service for the RWTH to fulfilling

1 re3data.org: Coscine; editing status 2024-03-15; re3data.org – Registry of Research Data Repositories. http://doi.org/10.17616/R31NJNJZ; *Visited on March 28, 2025.*

Published in: Vincent Heuveline, Philipp Kling, Florian Heuschkel, Sophie G. Habinger, and Cora F. Krömer (Hrsg.): E-Science-Tage 2025. Research Data Management: Challenges in a Changing World. Heidelberg: heiBOOKS, 2025. DOI: https://doi.org/10.11588/heibooks.1652.c23919 (CC BY-SA 4.0).

a mandate for nationwide service, thus highlighting the importance of adaptability to changing challenges. In terms of Coscine, it is necessary to consider not only challenges common to other software solutions, such as balancing maintenance and new features, but also emerging RDM trends.

The mission of Coscine is to provide a generic platform for German researchers and their collaboration partners to access storage resources, jointly work on datasets online in research projects, and retain datasets for up to ten years after the end of the project. We promote high quality research by fostering the FAIR-Principles (Wilkinson et al. 2016) for discovering and accessing data during the active project phase through the implementation of the FAIR Digital Object concept (for more details see our mission statement[2]). The maintenance of core functionalities includes, for example, user login, which supports authentication via DFN-AAI and via ORCID, ensuring easy and barrier-free access. After logging in, projects and sub-projects can be created in a tree structure in Coscine. Access management is ensured via a regulated role structure. Resources within projects can either provide access to storage space from the Research Data Storage (RDS-Web, -S3 or -WORM) or be linked to external storage systems (e.g. GitLab, LinkedData). In future, the RDS storage will be replaced by DataStorage.nrw.

Metadata management plays a central role in Coscine; therefore, entering metadata is mandatory when creating projects and resources. For the use of RDS-Web resources via the UI or the API, it is also essential to enter metadata before successfully uploading a file. To ensure that metadata is entered when using S3 resources, the storage space allocation for this resource type is carried out exclusively via an application procedure. Metadata is entered at the data level using metadata profiles. Metadata profiles can be selected from a list or created individually through the AIMS Metadata Profile Service (Grönewald et al. 2022). All created metadata profiles can be reused by researchers within Coscine. In addition, Coscine assigns Persistent Identifiers (PIDs) based on the ePIC handle system to ensure unique and permanent identification at both project and resource level.

Below we present the challenges that we have encountered during the development of Coscine in recent years, starting from 2018 till today with almost 3,300 registered users. In addition, we provide an insight into how we have effectively overcome these challenges and what our anchors are for flexibility in an ever-changing world.

2 Our Anchors in a Changing World

Anchors are crucial for addressing the challenges of a constantly evolving world. Therefore, we have identified three main anchors that enable us to remain true to our Coscine mission and effectively respond to the changing challenges:

2 https://about.coscine.de/en/about/mission-statement/; *Visited on March 28, 2025.*

- We aim to provide a FAIR environment for research data.
- We operate at the IT Center of RWTH Aachen University.
- We apply agile methods for maintenance and ongoing development.

2.1 Our Stakeholders – A Changing World

If we look at the various stages of Coscine's development, we recognize three characteristic phases: *Development Phase*, *Pilot Phase* and *Regular Operation* (see Figure 1). At the beginning of the Coscine development phase in 2018, there was only one stakeholder determining the requirements for the RDM platform – the IT Center of RWTH Aachen University. With the transition to the pilot phase in 2020, the number of stakeholders increased significantly, leading to close collaborations, for example, with NFDI-MatWerk or NFDI4ING. Since entering regular operation in April 2023, the number of stakeholders has continued to grow steadily, along with increasing requirements.

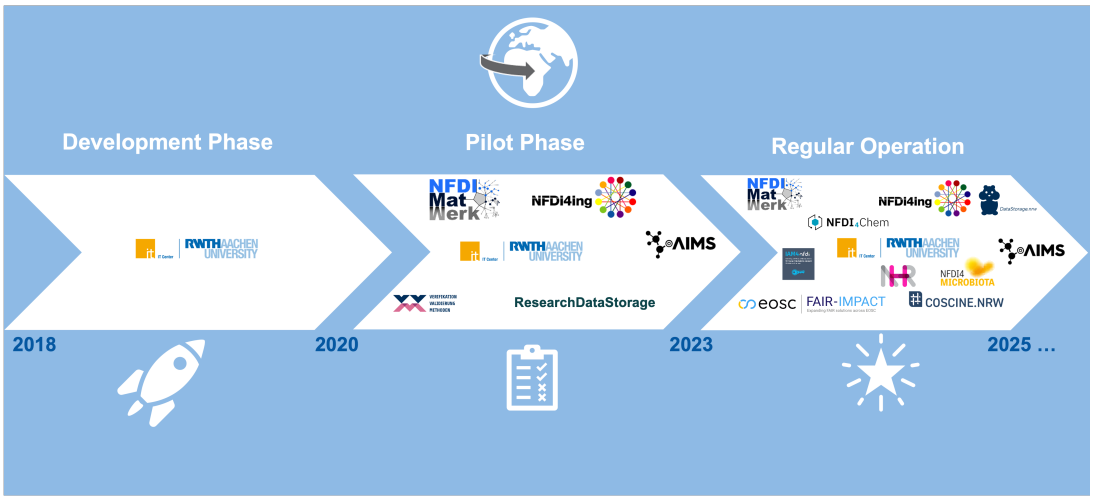

Figure 1: Coscine Stakeholder development over the years.

2.2 Our Development Phases in a Changing World

For an overview, we describe the three phases in more detail: the *Development Phase*, *Pilot Phase* and *Regular Operation*.

Development Phase (2018–2020)

The development of Coscine started in 2018 and continued for two years, characterized by a start-up mentality in the development team. The primary aim was to develop a

platform that focuses on RDM and supports the implementation of the FAIR principles (Lang, Jansen, and Politze 2025) in everyday life. In addition, it was important that the RDM platform not only covers a single section of the research data lifecycle, but can, at best, accompany researchers throughout the entire research data lifecycle.

This started with a single stakeholder – RWTH Aachen University. The primary requirement was to ensure that the RDM needs of RWTH researchers were met. As a result, the first decisions for the Coscine platform were made very low-threshold and the foundation for a generic RDM platform could be laid quickly.

The Coscine team, which was initially very small, chose aspects of Scrum and Kanban that were tailored to individual needs to organize development. Initially, however, the focus in the development phase was on the development team mastering the use of GitLab, for example, and building up and further optimizing the planning of the Scrum sprints step by step.

Pilot Phase (2020–2023)

The pilot phase began in 2020 at RWTH Aachen, transitioning the start-up mentality into a testing mentality. The purpose of the phase was to find out where the strengths and weaknesses of the platform lie and to incorporate improvements based on feedback from pilot users.

During this phase, many new stakeholders were added who were involved in the development of Coscine (see Figure 1) through funding and some of whom are still involved today (e.g., NFDI4ING and NFDI-MatWerk). This increased the need to collect ideas and feature requests, requiring clear prioritization for the implementation. We will go over how exactly this was achieved and what has changed in our prioritization process in the next chapter (see Section 3).

As the number of stakeholders increased, the Coscine development team gradually grew larger, as did the number of users.

Regular Operation (since 2023)

Coscine has been in regular operation since March 2023. This marked a milestone in development. As part of Coscine.nrw, the platform is made available to all DH.NRW universities as Software as a Service. As a result, the number of stakeholders and their needs continues to rise, and a balance needs to be found between maintenance of existing and the release of new features. During regular operation, the prioritization of new feature requests was finely adjusted to reflect real-world conditions.

To further support the increasing number of users, in addition to an extensive range of training courses (e.g., workshops on the basic functions and creating metadata profiles), tutorial videos have also been published to help users get started[3]. Furthermore, RDM staff at the DH.NRW level are trained to advise researchers within their own organization and introduce them to using Coscine. A peer review process is employed to allocate storage space, ensuring that researchers receive storage allocations after a successful and independent review. This ensures that all universities and technical colleges are provided with storage space in accordance with funding providers' specifications. To manage this increased support effort, a dedicated support model was developed for Coscine, and both the development team and the service management teams have grown since the transition to regular operation.

3 Challenges & Approaches

Below we share some challenges we have encountered in recent years, along with our experiences and proposed solutions.

3.1 Scalability

Increasing user numbers naturally, with around 3,300 users now registered in Coscine, result in a rise in feature requests, requiring a robust concept to ensure scalability. For us, the key has been the prioritization of requirements to best accommodate stakeholder wishes according to their importance.

During the *Development Phase*, prioritization was done internally, i.e. the decision about which features to be developed next was made by one or two people. External stakeholder influence was minimal, enabling a strong focus on creating a generic RDM platform to support FAIR principles.

To effectively handle the growing number of stakeholders during the *Pilot Phase*, we applied a prioritization model from the business sector: the WSJF (Weighted Shortest Job First) formula, originating from the Scaled Agile framework (SAFe)[4]. WSJF is commonly used by project owners to prioritize backlogs and determine next steps.

However, during *Regular Operation*, it became clear that the WSJF method in its original form was unsuitable for Coscine. We therefore modified the method by removing "job size" from the calculation. This adjustment reflects the necessity to fulfill stakeholder

3 https://docs.coscine.de/de/videos/; *Visited on March 28, 2025.*
4 https://framework.scaledagile.com/wsjf; *Visited on March 28, 2025.*

requirements regardless of job size, ensuring continued funding. Additionally, certain adjustments were necessary to keep up with evolving RDM trends.

Table 1 summarizes our prioritization process across the different development phases.

Table 1: Prioritization Approach of Coscine Development.

Phase	**Changing World**	**Style**	**Decision Maker**	**Process**
Development	One Stakeholder	Internal	Product Owner, Group Lead Coscine	Discussion
Pilot	Stakeholders with individual RDM requirements	Industry	Steering Board (Product Owner, Group Leads: Coscine, Consulting, HPC, Service Management, Department Head)	Voting via Weighted Shortest Job First (WSJF)
Regular Operation	Many Stakeholders; focus on scalable features	RDM	Steering Board (see above)	WSJF Adaption: no division by job size anymore

Figure 2 shows the process from the creation of issues and epics to prioritization and development. The creation is undertaken by the product owner, people in service management (May 2025: two full-time positions), the Scrum Team itself or external stakeholders. Issues, i.e. small work tasks, are prioritized by the product owner. Epics comprise larger work packages and are a collection of issues on one topic. These are prioritized by the Steering Board. The Scrum Team (May 2025: two full-time and three part-time developers) implements the specific work tasks within a two-week sprint with the typical Scrum events (planning, daily, review, retro).

3.2 Transparency

We aim to develop Coscine transparently, clearly communicating our current projects and methods. A key step toward transparency was developing Coscine as an open-source platform, enabling other platforms to reuse our features. A notable example is the NFDI Login, added in late 2024 through the IAM4NFDI project. Coscine contributed to the first incubator phase, enabling exchanges about technical implementation, i.e. NFDI login and account merging.

We also communicate clearly how Coscine supports FAIR principles (Jansen et al. 2024). For instance, we published insights at Zenodo through the FAIR-IMPACT Program Work-

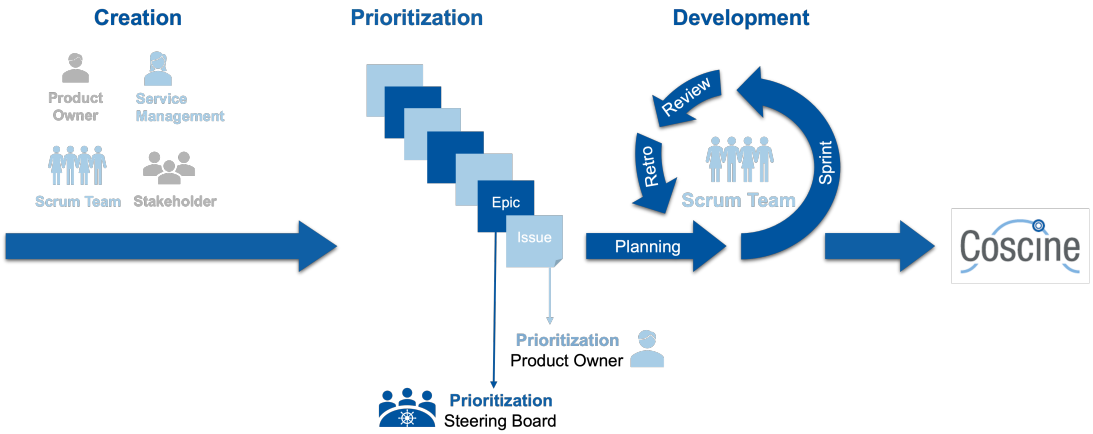

Figure 2: Agile Development Process for Coscine.

shop (Lang, Jansen, and Politze 2025). Additionally, a mission statement detailing our goals, mission, and values was published on our website[5].

Transparency extends to organizational practices. Our roadmap for larger work packages (EPICs) is publicly accessible, allowing users and stakeholders to view and directly submit feature requests[6]. At the development team level, work is tracked privately on a GitLab issue board, which reflects progress within each sprint. These issues represent smaller work packages that may collectively form an EPIC or address individual hotfixes and bug fixes. This level of development is not public to give the team space to work freely and to maintain a protected environment.

Transparency also involves reporting key performance indicators (KPIs). These metrics have been defined and are stored automatically. We have developed a visualization for external presentation, which is currently available only to selected individuals (e.g., stakeholders, DH.NRW staff). However, in coordination with the Coscine.nrw and DataStorage.nrw advisory board, we plan to make these publicly accessible over the course of 2025. This will provide insights into user numbers, quota usage, and login methods used.

3.3 RDM Streams

RDM is gaining significance, and trends in the field are constantly evolving. Coscine as an RDM platform is intended to grow with these trends and we therefore strive to keep an eye on the latest requirements and move forward as pioneers. For example, FAIR Digital Objects have been implemented (Politze et al. 2025) and continually adapted as part of the FAIR-IMPACT program.

5 https://about.coscine.de/en/about/mission-statement/; *Visited on March 28, 2025.*
6 https://git.rwth-aachen.de/groups/coscine/-/epic_boards/539; *Visited on March 28, 2025.*

The use of SPARQL queries for managing and retrieving metadata is becoming increasingly important for researchers. Recognizing this as a significant trend in RDM, we have identified SPARQL as a valuable addition to Coscine's capabilities. Offering researchers the possibility to use SPARQL enables more precise, efficient, and flexible querying of metadata, giving them finer control over their data. As part of our ongoing efforts to stay aligned with current and emerging RDM streams, we are actively exploring ways to integrate SPARQL query support into Coscine, empowering researchers to utilize advanced metadata management techniques in their everyday research activities.

Since excellent research depends on sufficient storage space, Coscine serves as a gateway to storage systems. Initially, we offered RDS, which will be replaced in 2025 by a new storage system called DataStorage.nrw. While this update introduces state-of-the-art technology for researchers, the transition poses a challenge: storage systems must typically be replaced approximately every five years, requiring repeated migrations. We are therefore currently in the migration phase from RDS to DataStorage.nrw and are establishing long-term migration plans for future cycles.

Coscine allows research data to be stored for up to 10 years after the end of a project. However, there are use cases where this period is not sufficient, and researchers require storage that guarantees long-term availability. To address this, we are currently collaborating with Long-Term Availability NRW (LZV.NRW) to explore integration options for long-term storage within Coscine.

3.4 Quality Management

To provide a functional and customizable platform for a large and diverse user base, a well-designed quality management approach is essential. Agile processes allow us to respond flexibly and confidently to changes, which is critical in the development area. We aim for a balanced focus between maintaining existing features and introducing new ones.

Development is inseparable from thorough testing, and the Coscine team has made several adjustments in this area in recent years. During the *Pilot Phase* (2018–2020), the priority was speed – many features were developed in a short time to deliver a basic RDM platform. Today, the focus has shifted to ensuring the quality of delivered features, even if this occasionally comes at the cost of development speed. Looking ahead, the team is placing greater emphasis on end-to-end (E2E) testing and adopting a test-driven development (TDD) approach.

We are also actively preparing for ISO/IEC 27001 certification to ensure security compliance – a major milestone in Coscine's development.

If users experience issues, our support team is available via a ticketing system, reachable by phone, chat, or email during service hours. Whether an on-call service will be needed in the future is currently under evaluation.

To continually improve the platform, we conduct an annual user survey to assess satisfaction and gather valuable feedback. We promote the survey through the Coscine newsletter, a banner on the platform, and the IT Center blog. For both the development and service management teams, these results are an important tool for understanding user needs and tracking changes compared to previous surveys.

4 Key Aspects

Mastering the challenges of a constantly changing world is not easy and requires continuous adjustments to methods and processes.

In essence, the following provides a summary of the biggest challenges faced by the Coscine RDM platform since its initial development until today:

1. **Scalability:** The platform should be usable by numerous stakeholders and therefore must fulfill a wide range of requirements and be quickly adaptable.

2. **Transparency:** We, as a team, clearly understand our mission and communicate it to the outside world. Additionally, we clearly define our processes and present them as openly and transparently as possible externally.

3. **RDM Streams:** The field of research data management is constantly evolving, and new changes are continually emerging. As an RDM platform, we aim to comply with these developments as effectively as possible and proactively move forward with the currents in the field.

4. **Quality Management:** The transition from a start-up to a permanent service requires constant attention to quality. We therefore do everything we can to deliver consistent quality while continuously improving.

We remain flexible and continuously adapt our workflows to changing requirements – like for example the adaptation of our prioritization method. We strive for active networking within the RDM sector, regularly participating in nationwide events, conferences, and similar activities (e.g. via fdm.nrw, NFDI, FAIR-Impact). Finally, we always remain true to our anchors: the FAIR principles, the IT Center of RWTH Aachen University, and agile methods.

Acknowledgements

The work on Coscine was partially supported with resources granted by DH.NRW for the project Coscine.nrw by Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen, by NFDI4Ing, funded by Deutsche Forschungsgemeinschaft (DFG) under project number 442146713 and NFDI-MatWerk, funded by Deutsche Forschungsgemeinschaft (DFG) under project number 460247524.

Authorship Contributions

- Katja Jansen: Conceptualization, Methodology, Writing
- Petar Hristov: Conceptualization, Methodology, Writing
- Ilona Lang: Conceptualization, Methodology, Writing

Conflict of Interest

The authors declare no conflict of interest.

Bibliography

Grönewald, Matthias, Patrick Mund, Matthias Bodenbrenner, Marc Fuhrmans, Benedikt Heinrichs, Matthias S. Müller, Peter F. Pelz, et al. 2022. "Mit AIMS zu einem Metadatenmanagement 4.0: FAIRe Forschungsdaten benötigen interoperable Metadaten". In *E-Science-Tage 2021: Share Your Research Data: Share Your Research Data,* edited by Vincent Heuveline and Nina Bisheh, 91–104. Heidelberg: heiBOOKS. ISBN: 978-3-948083-54-0. https://doi.org/10.11588/heibooks.979.c13721.

Jansen, Katja, Lukas C. Bossert, Amirreza Daei Rezaei Moghaddam, and Marius Politze. 2024. "Coscine – Makes your research data FAIR". In *From Data to Diamonds – Empowering Research with AI and RDM, Aachen (Germany), 14 Nov 2023 – 14 Nov 2023.* https://doi.org/10.18154/RWTH-2023-10428.

Lang, Ilona, Katja Jansen, and Marius Politze. 2025. *Implementing the FAIR Principles with Coscine.* Zenodo. https://doi.org/10.5281/zenodo.14886335.

Politze, Marius, Benedikt Heinrichs, Sirieam Hunke, Ilona Lang, and Thomas Eifert. 2025. "FAIR Digital Objects: FAIRtilizer for the Digital Harvest". In *EPiC Series in Computing,* 105:284–274. EasyChair. https://doi.org/10.29007/hfzk.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersbergand, Gabrielle Appleton, Myles Axtonand, Arie Baakand, Niklas Blombergand, et al. 2016. “The FAIR Guiding Principles for scientific data management and stewardship”. *Scientific data* 3 (1): 1–9. https://doi.org/10.1038/sdata.2016.18.