

Management Concept

Research Data Storage Infrastructure

for North Rhine-Westphalia (FDSI.nrw)



DataStorage.nrw



DataArchive.nrw

Guidance and explanation on the allocation of DataStorage.nrw and DataArchive.nrw storage space by FDM state base services of DH.NRW

Authors:

Ilona Lang¹, Florian Claus¹, Stephanie Rehwald², Magdalene Cyra-Wolf³, Katja Jansen¹, Florian Willems⁴, Michael Werger², Thomas Eifert¹, Marius Politze¹

License: CC-BY 4.0

Version: 1.0

Note: Please note that the original German text of these guidelines is the only binding version. In the event of any discrepancies between the English translation and the original German text, the German text shall always prevail.

¹ RWTH Aachen University

² Universität Duisburg-Essen

³ fdm.nrw

⁴ Universität zu Köln

Table of Contents

Introduction	3
Target audience	3
Governance	4
Definitions	4
Goals and values of FDSI.nrw	5
Framework Conditions	5
Science-Driven Allocation Procedure	5
Possible Storage Duration	6
Supported Research Initiatives	6
Research Projects	7
Data Pools	7
Enabling Cooperative Use	8
Compliance with FAIR Principles	9
Publication of Selection Policy	9
Reporting Data Transmission	9
Participation of Employee Representatives	10
Publication of acknowledgements	10
Accreditation of FDM State Basic Service	11
Acknowledgements	11

Introduction

This document defines the management guidelines for the research data storage infrastructure in North Rhine-Westphalia (FDSI.nrw), which currently consists of DataStorage.nrw and DataArchive.nrw.

The management of FDSI.nrw is led and coordinated by FDSI.nrw Service Management at RWTH Aachen University. The governance of the management is integrated into the structure of Coscine.nrw². Part of the governance structure is the advisory board, which includes representatives of the users of the Digital University of North Rhine-Westphalia (DH.NRW) and the spokesperson of the FDSI.nrw operating committee. The DH.NRW is a collaboration group of 42 universities in North Rhine-Westphalia and the Ministry for Culture and Science of the North Rhine-Westphalia state. The operator committee is the governance level of the operation, which is composed of all participating institutions of the FDSI.nrw consortium³.

DataStorage.nrw is a redundant S3 object storage system operated at four DH.NRW university locations: the University of Duisburg-Essen, the University of Paderborn, the University of Cologne, and RWTH Aachen University. As part of FDSI.nrw, DataStorage.nrw supports the implementation of the DH.NRW's state RDM concept⁴. The object storage was designed for particularly large amounts of data that need to be stored long-term and offers a capacity of up to 24 petabytes.

DataArchive.nrw is another storage platform that is to be set up in 2026. It is specifically designed for “cold data,” i.e., data that users rarely access. In order to enable an economically and ecologically sustainable offering for the storage of even large amounts of data over longer periods of time, this platform uses tape storage technology. Storage takes place at two universities (RWTH Aachen University, University of Cologne) in order to achieve the highest possible data persistence through spatial distribution and duplicate data storage, so that the stored data remains reliably available over a long period of time without loss.

Access to and use of FDSI.nrw is provided via services that are offered as FDM state-based services⁵ (e.g., repositories, cloud offerings) by DH.NRW universities for other DH.NRW universities for cooperative use (see also the chapter “Cooperative Use”).

Target audience

This document describes the conditions under which an FDM state base service of DH.NRW may deliver data sets to FDSI.nrw. It is therefore intended for IT service providers who operate such an FDM state-based service. The distribution guidelines were developed by FDSI.nrw Service Management and approved by the Advisory Board.

In addition to FDM state-level services, this management concept is also aimed at large-scale research projects at DH.NRW universities. Large-scale research projects are research infrastructures that open up unique research opportunities and make them available to a broad

² See <https://about.coscine.de/> (accessed on 07.08.2025)

³ See <https://www.fdm.nrw/forschungsdatenspeicherinfrastruktur> (accessed on 07.08.2025)

⁴ DH.NRW I AG FDM. (2024). Digital Ecosystem DH.NRW – State Concept FDM. Zenodo. <https://doi.org/10.5281/zenodo.11242683>

⁵ See <https://www.fdm.nrw/technische-fdm-landesdienste> (accessed on 23.09.2025)

group of researchers. Examples include telescopes and particle accelerators. Such research infrastructures usually have the funding to ensure the management of the data generated. Direct use of FDSI.nrw by large-scale research projects must be applied for at an early stage at RWTH Aachen University and must comply with the guidelines set out here. A review will be conducted to determine whether cost sharing is required. Indirect use of FDSI.nrw via already connected FDM state base services can only take place if the respective selection criteria of the FDM state base service (so-called selection policies) are also complied with.

Governance

Important decisions regarding the management of FDSI.nrw are made by FDSI.nrw Service Management on the recommendation of the Advisory Board.

To this end, FDSI.nrw Service Management prepares a decision template in which the subject of the decision is explained, the interests and positions of the parties involved are presented, and a reasoned proposal for a decision is made. Decision templates can be initiated or created by both state-level services and FDSI.nrw service management itself.

Decision templates are made available to the advisory board at least 7 days before its next meeting. The advisory board makes its recommendation for a decision. It is free to endorse a decision option already mentioned in the template or to propose a decision option itself. It may justify its recommendation but it is not required to do so.

Important decisions to which this process applies include the accreditation of a state base service or a large-scale research project, the withdrawal of accreditation, and the confirmation of data recovery. It is also applied to other issues of sufficient importance that are not currently foreseeable. FDSI.nrw service management decides whether an issue is of sufficient importance.

Definitions

FDM state-level services are applications in accordance with the state concept of fdm.nrw that make FDSI.nrw storage resources available to end users.

For the purposes of this document, **research data** refers to digital data that has been collected, generated, or used as part of a research initiative.

The research data storage infrastructure of DH.NRW (FDSI.nrw) currently consists of DataStorage.nrw and, in the future, DataArchive.nrw.

FDSI.nrw service management consists of at least two people (one main person in charge, one deputy) who work as FDM staff at the IT Center of RWTH Aachen University. RWTH Aachen University is the consortium leader of FDSI.nrw.

A **selection policy** for a state-based service is a guideline describing the selection criteria that determine which types of files are included in and managed by the service and which criteria must be considered. This guideline helps to ensure the quality and relevance of the stored data.

Goals and values of FDSI.nrw

FDSI.nrw serves as a capacity and persistence layer for storing research data for a maximum of 10 years after the end of the research initiative in which the data was generated. DataStorage.nrw offers a low-threshold option for storing data sustainably early in the research data lifecycle, ideally immediately after it has been obtained.

FDSI.nrw is intended exclusively for research data and is open to research data from all disciplines. FDSI.nrw is committed to good scientific practice (GWP) and the FAIR principles. Data stored there must be structured, annotated with sufficient metadata, documented, and referenced by persistent identifiers (PID) to ensure its reusability. This is ensured by the applications through which the storage space is provided and the review of storage space requests.

FDSI.nrw creates the infrastructural framework (GWP, Guideline 3) that enables the management of research data in accordance with the FAIR principles (GWP, Guideline 13) and protects against the manipulation of research results (GWP, Guideline 12).

FDSI.nrw is committed to transparency. Statistical information on usage is made publicly available. The allocation of storage space must also be carried out in a transparent, non-discriminatory manner.

Framework Conditions

FDM state base services of DH.NRW conclude contracts with the operator committee of the storage consortium for the use of FDSI.nrw and undergo accreditation based on the management concept mentioned here. Each state base service names two contact persons to FDSI.nrw service management.

The target group for FDM state-based services is researchers from participating universities. Participating universities must first conclude appropriate contracts with the respective FDM state-based service. All member universities of DH.NRW that are state-funded or state-refinanced are eligible to participate.⁶ FDSI.nrw can also be used via FDM state-level services to support researchers through NFDI consortia based at participating universities.

Science-Driven Allocation Procedure

The technical provision and management of storage space on the FDSI.nrw must be carried out through uniform, science-driven procedures. These procedures must be identical for researchers from all participating universities and include:

- The application for storage resources per research initiative,
- The evaluation of applications through a science-driven process by a pool of experts from all participating universities, and
- The provisioning of the storage resources.

For each research project, up to 100 GB can be allocated to researchers without further application through FDM state basic services, provided that the FAIR principles are technically ensured

⁶ At present, this excludes only Witten/Herdecke University as a purely privately funded institution of higher education.

by the FDM state basic service according to the concept outlined in the accreditation. The provision of up to 100 GB of storage space without an application promotes user-friendliness and flexibility by allowing users quick and uncomplicated access to resources. This facilitates the implementation of small projects and balances administrative effort with storage needs. The limit of 100 GiB is based on years of experience from the RDS.nrw context. A change in automatic storage allocation can occur through an application by a state basic service via the FDSI.nrw decision-making process.

Additional storage requirements must be applied for and reviewed by at least two people from a pool of experts from the participating institutions for compliance with the FAIR principles and good data organization. At least one external reviewer from another participating institution within the FDM context must be involved.

For larger storage requirements, additional scientific quality reviews must be conducted by suitable, independent reviewers. Suitable reviewers have relevant professional experience in the field or corresponding further education. For applications exceeding 125 TiB per project, a review by external scientific reviewers must take place (cf. Jansen et al., 2025). Scientific reviewers must be active researchers in the field of applied projects. The review can be replaced if it refers to an application already reviewed through a science-driven peer-review process (e.g., DFG), if this directly results in a specific need for additional storage space. Additionally, reviewers ensure that applications meet criteria for either research projects or data pools (see chapter on Research Initiatives).

Possible Storage Duration

FDSI.nrw offers pure bit stream preservation (further details are available in its operating concept) for up to ten years after completion of any given project during which research data were generated. Longer retention periods can only be realized within potential future DH.NRW offerings concerning long-term availability when measures such as risk management, monitoring, and curation are demonstrated. The providers of FDM state basic services offering accessible storage assume technical responsibility for data while maintaining references to their context over necessary migrations including responsible parties and descriptive metadata. State basic services ensure processes deleting orphaned data within accreditation frameworks.

FDSI.nrw requires periodic updates/expansions; thus, funding commitments made by consortium universities (University Duisburg-Essen; University Paderborn; University Cologne; RWTH Aachen) under DataStorage.nrw and DataArchive.nrw stand contingent upon continued financial support from NRW State Government. In cases where funding ceases unexpectedly - termination protocols outlined within operational guidelines will govern decommissioning procedures. FDSI.nrw Service Management notifies designated contact persons at respective state basic services immediately following decisions regarding decommissioning events.

Supported Research Initiatives

FDM state basic services may only allocate storage space on the FDSI.nrw within defined research initiatives: these are research projects and data pools. Accordingly, storage space may not be allocated to individual researchers, institutes, or core facilities without reference to a specific research initiative. All research initiatives (research projects & data pools) for which an FDM state basic service provisions storage space on the FDSI.nrw must exhibit the following characteristics:

- **Timeframe:** The start of an initiative is clearly defined in the associated metadata.

- **Data:** The collected data of an initiative have a clear thematic coherence. This must be reflected in uniform metadata for research initiatives. The conditions and prerequisites for contributing data are defined. Quality assurance processes are defined and implemented.
- **Participants:** Initiatives are carried out by a defined group of people whose composition can change over time. There are one or more principal investigators (Principal Investigator, PI), who are usually constant throughout the project duration. People who primarily apply for and use storage space on the FDSI.nrw must be employed at a participating university or have a direct connection to an NFDI consortium based at a participating university.
- **Organizational Location:** A research initiative is located at one or more research institutions. In cases where multiple institutions are involved, one of them is consortium-leading and must primarily be assigned to the initiative.

These characteristics enable the handling of research data according to the research data lifecycle. Research data from research initiatives that are already made permanently available elsewhere by a recognized organization should not be stored again on the FDSI.nrw.

Research Projects

The primary mission of FDSI.nrw is to provide research projects with the necessary storage space to secure the research data generated or used during the project for its duration as well as up to 10 years after its completion.

Research projects exhibit (in addition to those mentioned under research initiatives) the following characteristics:

- **Timeframe:** A research project is time limited. The end of the project is clearly defined in the associated metadata. Subsequent changes to the project duration must be registered by the FDM state basic service.

Data Pools

Data pools are often long-term collections (i.e., often without a defined endpoint) of thematically related data that have been generated or obtained in different contexts and reused in different contexts. The thematic coherence of the data is also reflected in metadata fields that all stored data have in common. For storage on FDSI.nrw, (in addition to those mentioned under research initiatives) the following characteristics are relevant:

- **Timeframe:** Both the collection and reuse of data can be unlimited or have defined endpoints. Therefore, applications for data pools are initially approved for a collection phase of a maximum of five years and a subsequent reuse phase of another five years. At the end of the collection and/or reuse phase, an extension by another three years can be applied for. The extension can be applied for repeatedly without limit. In case of extension applications, it is particularly evaluated by the respective FDM state basic services whether the criteria specified in the concept have been met.
- **Storage Requirement Over Time:** A storage application must be submitted per collection phase. The expected storage requirement must always be specified for the end of each year within the collection phases.
- **Usage:** Data pools have a developed usage concept in which the expected reuse is presented. The usage concept defines the group of persons authorized for reuse. If there are prerequisites for reuse by external persons, these must also be presented. The

usage concept describes all relevant legal requirements that must be observed for re-use. Additionally, it specifies which technical prerequisites are required for reuse and how these are ensured and continuously verified. The creation, implementation, and transmission of the usage concept is defined in the Selection Policy of the respective state basic service.

- **Participants:** The actor groups of contributing and reusing researchers are defined in the aforementioned concepts. The researchers applying for a data pool demonstrate that they have sufficient competencies and resources to maintain it properly; one leading person as well as at least one deputy leading person will be named—these should hold permanent positions within their respective organizations.

Data Rescue

In cases of data rescue for datasets of exceptional scientific value for which no other suitable storage location is available short-term, the criteria of the aforementioned types of research initiatives can be temporarily suspended. Upon appropriate application, the respective service management of the state basic services will contact FDSI.nrw Service Management. The assessment of whether it constitutes a data rescue is made by FDSI.nrw Service Management, which then allocates the corresponding storage space. The decision on long-term storage on FDSI.nrw is made according to the FDSI.nrw decision-making process.

Enabling Cooperative Use

FDSI.nrw is a cooperative information infrastructure⁷. Es muss durch den FDM-Landesbasisdienst sichergestellt werden, dass Personen, die Speicherplatz auf der FDSI.nrw hauptverantwortlich beantragen und nutzen, eine Anstellung an einer teilnehmenden Hochschule oder einen direkten Bezug zu einem NFDI-Konsortium, das an einer teilnehmenden Hochschule angesiedelt ist, aufweisen. Ein direkter Bezug ist dann gegeben, wenn eine direkte Betreuung durch NFDI-Personal bei der jeweiligen Forschungsinitiative nachweisbar ist.

It must be ensured by the FDM state basic service that persons who primarily apply for and use storage space on FDSI.nrw are employed at a participating university or have a direct connection to an NFDI consortium based at a participating university. A direct connection exists if direct supervision by NFDI staff in the respective research initiative can be demonstrated.

FDM state basic services allocating storage space from FDSI.nrw must be open to all participating universities. They may also be offered to other research institutions in NRW, subject to cost participation covering additional costs incurred (personnel costs for operation and service management). When applying for access to FDSI.nrw, FDM state basic services must demonstrate that they have qualified according to funding rules for DH.NRW projects⁸. In their public appearance, they must clearly target user groups beyond local institutions. For authentication, a procedure supported by IDM.NRW must be used.

⁷ See also DFG discussion paper “Digital Research Practice and Cooperative Information Infrastructures” <https://www.dfg.de/de/aktuelles/neuigkeiten-themen/info-wissenschaft/2025/ifw-25-08> (accessed on October 13, 2025)

⁸ See funding for digitization-related cooperation projects, DH.NRW, <https://www.dh.nrw/foerderung> (accessed on August 7, 2025)

Compliance with FAIR Principles

Research data stored on FDSI.nrw must be stored in accordance with the FAIR principles, tailored to the respective use case and stage in the research data lifecycle. FDM state basic services enable compliance with all elements of the FAIR principles, require their users to adhere to them, and review the corresponding concepts as part of their allocation process. These criteria include, among others, the assignment of PIDs per initiative and associated datasets (buckets), linkage with a minimum set of metadata, and enabling role structures. The minimum set of metadata per research initiative is as follows:

- Name of the initiative
- Description of the initiative (Abstract)
- Name(s) of the initiative's leadership (PIs) including email address(es)
- Start date
- End date (for data pools: end after 5 years)
- DFG subject discipline
- Responsible organization
- If applicable: Additional organizations
- Keywords

The specific support for FAIR principles by each access layer must be outlined in a published document. Part of FDM state basic service development must be regular derivation and implementation of measures to improve services in line with FAIR principles.

Publication of Selection Policy

Each FDM state basic service must publish a Selection Policy (description of selection criteria) from which allocation guidelines for storage space on FDSI.nrw for end users are derived.

Reporting Data Transmission

Each FDM state basic service must provide at least monthly reporting on storage allocation, which is publicly accessible to universities, broken down by DFG subject disciplines and participating universities. The monthly reporting enables FDSI.nrw Service Management to monitor trends and developments promptly, helping to identify deviations early and take measures if necessary (e.g., acquiring new resources). All FDM state basic services collect key metrics at the level of individual research initiatives for this purpose.

The purpose of reporting is to create transparency about the use of FDSI.nrw. It should demonstrate how often and to what extent FDSI.nrw is used to emphasize its importance for research at participating universities. Additionally, reporting serves to provide decision-makers in FDSI.nrw governance (Service Management, Advisory Board, Operating Committee) with data that contribute to improvements, targeted investments, and further development of the infrastructure. DH.NRW and MKW are also part of the target group as funding providers.

The published data must be anonymized (no direct reference to research initiatives or involved persons should be recognizable) and include the following information:

1. **Used Storage Space:** Amount of storage space currently used by the FDM state basic service.

2. **Storage Space Utilization:** Ratio of used to provisioned storage space (in percent).
3. **Number of Users:** How many users have an account with the service and how many interact actively with it.
4. **Average Consumption per Organization:** Average storage consumption per participating university.
5. **Data Growth:** Rate at which used storage space increases over the past week.
6. **Storage Space Utilization by Categories:** Breakdown of used storage by DFG subject discipline at both service and initiative levels.

For the bi-annual meetings of the Advisory Board (held in January and June each year), reporting figures must be provided. A reminder will be sent one month in advance by FDSI.nrw Service Management.

Participation of Employee Representatives

FDM state base services that wish to access FDSI.nrw must provide all necessary information for the participation of the respective staff representatives of the user universities. This usually includes clear documentation (including concepts for authorizations, guidelines for deleting data, data protection concepts) to ensure compliance with legal and organizational requirements based on the State Personnel Representation Act. This provision can be fulfilled as part of the DH.NRW application process for state-based services.

Publication of acknowledgements

Each state-based service must indicate the use of FDSI.nrw storage space on its website. This information can be presented as follows:

A central element of [FDM state-based service name] is access for researchers to storage space on the DH.NRW research data infrastructure (FDSI.nrw). FDSI.nrw is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia (DataStorage.nrw: MKW: 214-76.01.09-7-7937; DataArchive.nrw: MKW: 76.04.10-06 – 2024-6188) and the German Research Foundation (DataStorage.nrw: DFG: INST 222/1530-1; DataArchive.nrw: DFG: INST 222/1581-1). The joint, distributed object storage system for central applications is managed by the FDSI.nrw consortium. The participating universities are RWTH Aachen University as consortium leader, FH Aachen, Ruhr University Bochum, TU Dortmund, University of Duisburg-Essen, University of Cologne, and University of Paderborn.

Researchers who use FDSI.nrw should be asked to include the following statement in publications and conferences:

If you publish your research (meta)data with [FDM state base service name] and publish a publication in this context, we would be grateful if you could also mention this in your acknowledgments. Simply use the following citation template:

The data used in this publication is managed via [service name], whose storage space is part of FDSI.nrw. FDSI.nrw is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia (DataStorage.nrw: MKW: 214-76.01.09-7-7937; DataArchive.nrw: MKW: 76.04.10-06 – 2024-6188) and the German Research Foundation (DataStorage.nrw: DFG: INST 222/1530-1; DataArchive.nrw: DFG: INST 222/1581-1).

Accreditation of FDM State Basic Service

FDM state base services that allocate storage space to FDSI.nrw are accredited by FDSI.nrw service management on the recommendation of the advisory board. Accreditation is valid for an unlimited period of time, provided that the criteria specified in the management concept are met. However, FDSI.nrw Service Management or the Advisory Board may, if there is a legitimate interest, initiate an audit of the FDM state base services to review the accreditation. An unsuccessful audit may result in conditions being imposed on the continued provision of the FDM state base service or in the immediate withdrawal of accreditation by FDSI.nrw Service Management as a result of the FDSI.nrw decision-making process.

Compliance with the management concept is monitored and ensured by FDSI.nrw Service Management. If FDM state-level services wish to access FDSI.nrw, the advisory board must issue a recommendation to this effect. The advisory board thus advises FDSI.nrw service management at RWTH Aachen University in its role as the consortium leader, which makes the final decision.

The following additional documents are required for accreditation:

- Concept description (FDM state-level service and assessment procedure)
- Mission statement
- Selection criteria (selection policy)
- Statement on the FAIR principles
- LOIs from the universities in the DH.NRW funding application
- Staff council vote

After successful accreditation, FDSI.nrw Service Management informs the FDSI.nrw Operator Committee about the new FDM state base service.

Acknowledgements

We would like to thank the advisory board of FDSI.nrw and the operator committee for their support in developing this management concept. FDSI.nrw is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia (DataStorage.nrw: MKW: 214-76.01.09-7-7937; DataArchive.nrw: MKW: 76.04.10-06 – 2024-6188) and the German Research Foundation (DataStorage.nrw: DFG: INST 222/1530-1; DataArchive.nrw: DFG: INST 222/1581-1). The joint, distributed object storage system for central applications is managed by the FDSI.nrw consortium. The participating universities are RWTH Aachen University as consortium leader, FH Aachen, Ruhr University Bochum, TU Dortmund, University of Duisburg-Essen, University of Cologne, and University of Paderborn.