

Repository Policy - Radboud Data Repository

Date: this Policy came into effect on 06 March 2025

Version: 1.0 (06 March 2025)

Aim and scope of this Policy

This Policy concerns the [Radboud Data Repository](#) (RDR): Radboud University's institutional [repository](#) for research data. The aim of this Policy is to ensure that the contents of the RDR reflect and strengthen Radboud University's mission. This Policy also establishes the framework within which the RDR is managed.

The **target audience** of this Policy is:

- Researchers from Radboud University or the Radboud University Medical Centre (Radboudumc) that use the RDR to [archive](#) or [publish](#) their research data ('depositors')
- External users of the RDR
- Policy makers and support staff within and outside Radboud University
- Other stakeholders in the Research Data Management field

It governs the obligations, responsibilities, and rights of these players.

About Radboud Data Repository

The RDR supports researchers of the University and the Radboudumc in the long-term archiving and FAIR publication of their research data, in compliance with [Radboud University's Research Data Management \(RDM\) Policy](#).

The **mission of the RDR** is to [preserve](#) Radboud University's research data for the [long-term](#) and to make these data accessible according to the [FAIR principles](#), in order to:

- Enhance the impact of research executed at Radboud University by supporting **reuse** of research data **within and outside of Radboud University**
- Support **research quality and integrity** by facilitating the documentation of the research process

With these objectives, the RDR contributes to the mission of Radboud University:

"to contribute to a free and healthy world with equal opportunities for all and to make a significant impact on a regional and an international level."

The primary focus of the RDR is digital research data from all disciplines.

*By **research data**, we mean all data, digital and non-digital, that was generated or reused during the research process and used to form scientific conclusions. This encompasses information that is used to form and test hypotheses, as well as information upon which conclusions are based. This definition of research data includes observational, experimental,*

simulated, derived, and compiled data. It also includes software generated during the research process or the workflow describing the research process.

Research data are organised in the RDR in packages called 'data collections'.

*A **data collection** is defined as a set of research data files related to a specific topic, project, or publication. It always includes metadata that support the findability and accessibility of the research data. A data collection is sufficiently well documented to allow end users to use the data collection for its intended purpose. This intended purpose may be internal reuse of research data, external reuse of research data, and/or verification or replication of published research.*

The RDR was originally developed at the Donders Institute for Brain, Cognition, and Behaviour and is now widely used by many of Radboud University's research institutes and the Radboudumc. The RDR is developed and maintained by the University's Information & Library Services division.

Relevant documents:

- [RDR User Manual](#)
- [Radboud University Guidelines for Research Data Management](#)

User groups of the RDR

The RDR includes the following user groups:

- **Depositors:** researchers of Radboud University and Radboudumc. They are called "collection managers" in the RDR. They can:
 - 1) [Deposit](#) their research data in the RDR
 - 2) Download and edit content from data collections dedicated to their project(s)
 - 3) Browse and search the [metadata](#) and download the content of published data collections
- **Associates:** people who have joint projects with depositors. They can:
 - 1) Download and edit (if edit rights are granted by the depositors) content from the collection dedicated to their project
 - 2) Browse and search the metadata and download the content of published data collections
- **External users:** professional researchers, citizen scientists, teachers, students, informed citizens, the general audience, etc. They can:
 - 1) Browse and search the metadata and download the content of published data collections

Content Policy

The RDR archives and provides (public) access to research data collections of Radboud University's and Radboudumc's researchers from all disciplines. The definitions of research data and data collections are provided above. If the deposited data files do not

meet the definition of research data, or cannot be considered a data collection, they are not accepted in the RDR. The table in the [Appendix](#) includes a non-exhaustive list of examples of research data, and describes in which collection type in the RDR they should ideally be archived. The table also includes examples of data that are **not** to be deposited in the RDR.

Research data may be sensitive, for example in the case of personal data. The RDR offers various [access levels](#) that are appropriate for different levels of sensitivity.

Data collections in the RDR must be accompanied by sufficient documentation files and/or metadata (for examples, see the RDR [best practice page](#)) to allow for the intended use of the data collection (i.e. internal reuse, external reuse, and/or verification or replication of published research).

The RDR accepts data in all file formats but has [recommendations on preferred file formats](#) that offer the best long-term guarantees for usability, accessibility, and sustainability.

Data collections may be deposited in any language.

The RDR has the right to refuse and retract data collections that are not in scope of this Content Policy.

The RDR is not responsible for the content of any data collection and is not liable for damages that occur as a result of the use of any data collection, nor for (incorrect) conclusions based on a data collection.

Submission Policy

Data files may only be deposited by authenticated and authorised users of the RDR. Within the RDR, Radboud University's research institutes and Radboudumc are accommodated as Organisational Units. Users can be made eligible collection manager within an Organisational Unit by that Organisational Unit's research administrator, if they:

- Have Radboud University or Radboudumc credentials, and;
- Are employed by the Organisational Unit's research institute, and;
- Meet any other criteria that have been drawn up by the Organisational Unit

The submission process must conform to the submission guidelines in the [user manual](#).

The following conditions apply to a deposit in the RDR:

- Collection managers shall ensure that they or their Organisational Unit have all necessary rights and consents required to deposit and -if applicable- publish the data
- Collection managers are responsible for compliance with any national or international laws and regulations, such as the General Data Protection Regulation ([GDPR](#)), ethical guidelines such as the [Netherlands Code of Conduct](#)

[for Research Integrity](#), and publisher embargoes. Collection managers shall ensure that deposited content is not unlawful, threatening, abusive, deceptive, fraudulent, erroneous, or invasive of another person's privacy

- The validity and authenticity of the data collection are entirely the responsibility of the collection manager
- The collection manager shall ensure that the deposit and publication of a data collection does not infringe (intellectual property) rights of third parties. The RDR is not liable for damages that result from any copyright or other intellectual property right infringements
- All data collections are automatically checked for the presence of obligatory metadata
- Submission of a data collection is non-exclusive, meaning that the collection manager is free to submit the data collection to other publishers or repositories
- Collection managers shall ensure that deposited data collections do not contain software viruses or any other programs that have been designed to disrupt the proper function of any software, hardware, or telecommunications equipment or to damage or obtain unauthorized access to any system, data files, or other information of the RDR or any third party
- There is no size limit for data deposit into the RDR. If the data are too large for the RDR's upload functionality, the RDR team will work with the collection manager to find a solution
- All data collections that are to be published are subject to a review process by Research Data Management experts of the university. The review process includes advice to improve the completeness and clarity of documentation and metadata as well as a limited scan for the presence of sensitive data that should not be publicly shared. Reviewers only edit a collection's data or metadata in agreement with the collection manager
- The collection manager must specify an access level and user terms (i.e. a licence or [Data Use Agreement](#)) prior to a data collection's publication
- Data can be deposited at any time but will not be made publicly accessible by the RDR before the collection manager publishes the collection and any embargo period selected by the collection manager has expired. Items will be made publicly accessible under the access level and licence or Data Use Agreement selected by the collection manager
- The RDR will provide a Digital Object Identifier (DOI) for every deposited data collection
- The RDR will publish the metadata of publicly available collections and export them to relevant repositories and aggregators in several standard formats, such as schema.org JSON-LD, DublinCore, and DataCite Metadata Schema. The RDR is dedicated to providing high-quality and machine-readable metadata
- If a data collection is published under restricted access conditions, the RDR forwards access requests to the collection manager. The collection manager is responsible for responding to such an access request within a reasonable period. By submitting a restricted access data collection to the RDR, the collection manager accepts the responsibility for handling access requests

- In the event that the RDR finds a violation by an end user of the applicable licence, Data Use Agreement, or access restrictions of a data collection, the RDR aims to notify the collection manager and to suspend the user from further use of the data collection

If the RDR has reasons to suspect that a deposited data collection constitutes a copyright violation, data breach, falsified research, or noncompliance with the conditions listed above in any other way, a relevant body (within or outside of Radboud University) will be identified based on the specific case and contacted to conduct an investigation. Depending on the results, the relevant data collection may be retracted.

By depositing (meta)data into the RDR, collection managers accept the terms and conditions included in this Policy.

Versioning and Curation Policy

The RDR records the version and status (editable, reviewable, or published/archived) of every data collection.

If necessary, an updated version of a published or archived data collection may be deposited. This may be necessary when the researcher wishes to add data to the collection, correct an error in the collection, or update it in any other way. The repository uses explicit version numbers. Different versions of the same collection link to each other, with the most recent version clearly identified. The collection's persistent identifier (the DOI) will always resolve to the latest version.

The access level and terms of use (i.e. licence or Data Use Agreement) of a published data collection can be updated only upon approval from Radboud University's Digital Competence Centre. Otherwise, the data collection can be retracted by the Digital Competence Centre upon request.

The RDR has the right to modify the format and/or functionality of published and archived data collections if this facilitates the digital sustainability, distribution, interoperability, or re-use of the data collection. In that case, initial items are retained unchanged and edits are only made on copies of originals.

Preservation, Deletion, and Retraction Policy

The RDR aims to ensure continued content readability and accessibility of the deposited data for as long as indicated by the depositor. It has a Preservation Policy in place that details the repository's approach to long-term preservation.

In the event of the RDR being closed down, Radboud University will endeavour to transfer the database to another trusted archive.

The collection manager must select a preservation period for their data collection. After this period, the collection manager may remove the content of the collection.

The content of a data collection may be retracted upon request if it has been published under a wrong access level, licence, or Data Use Agreement.

The RDR may retract the content of a published data collection if it is found to be out of scope of the repository's Content Policy or does not comply with the conditions listed in the Submission Policy.

The metadata of a removed or retracted data collection are retained indefinitely and the Digital Object Identifier (DOI) will continue to resolve to a "tombstone" page. The "tombstone" page contains a message stating that the data collection has been retracted or removed and -if applicable- refers to the data collection's successor.

Metadata Access Policy

Access to some metadata is public, whereas for other metadata it is restricted. Metadata of published collections (open access, open access for registered users, or restricted access) are always made public. Metadata of internally archived (closed access) collections are made public only if the collection manager chooses to do so.

Public metadata are released under the Creative Commons [CC0 Public Domain Dedication licence](#). This dedication applies exclusively to the metadata associated with the deposited items and under no circumstances is a substitute for or affects their licences or permits.

Anyone can access, copy, share, modify, harvest, and re-use the public metadata for any purpose, free of charge, and without requiring permission from Radboud University or the depositor. Mention of the RDR is appreciated but not mandatory.

Data Access Policy

The access policy for the deposited research data varies. Restrictions may apply to allow the collection manager to prevent harm to individuals, communities, ecosystems, etc. The RDR therefore offers various access levels:

- Closed access
- Open access
- Open access for registered users
- Restricted access
- Embargoed access
- Double anonymous access to unpublished data collections for peer review

Internally archived (closed access) collections are accessible only to the collection manager and associates.

The access level and the conditions and restrictions on re-use are specified in the collection's metadata. The [user guide for visitors](#) provides instructions for end users on how to access and/or request access to the repository's published collections.

Users shall acknowledge the contribution of the data producer if they use a repository collection in a publication in whatever form (e.g. journal publication, poster, presentation), even if the licence does not explicitly require them to do so, in accordance with the scientific standards for research integrity. The reference should include the names of the data producer(s), the publication year of the data, the title of the data collection, the name of the publisher, the DOI of the collection as a full URL, the version number, and -optionally- the resource type (e.g., Dataset, Software, Workflow).

The collection's reference can be found and downloaded (as a BibTex file) on the landing page.

Harvesting Policy

The metadata of published collections and internally archived (closed access) collections with public metadata are made available to harvesters and aggregators under the Creative Commons [CC0 Public Domain Dedication licence](#).

Privacy Policy

The repository's privacy policy can be found on the web portal:
https://data.ru.nl/doc/privacy_policy.html.

Policy Revisions

This Policy is subject to revision whenever deemed necessary by changed circumstances or changes in best practice.

Terms and Definitions

Archive: an organisation that intends to [preserve](#) information for access and use by a [designated community](#). From <https://public.ccsds.org/pubs/650x0m2.pdf>.

The verb, i.e. "to archive", refers to placing information in an archive.

Data repository: a data repository is an archival service providing [long-term](#) care for digital objects with research value. For the RDR, this concerns research data and corresponding metadata.

Data repositories are also commonly referred to as 'digital repositories', 'institutional repositories', or 'digital archives'.

Data Use Agreement: a Data Use Agreement governs the sharing of data between the data sharing party and the user of those data. It establishes how information in a dataset may be used by the intended recipient and how it is protected.

Deposit: the act of placing materials in custody of a data repository.

A '**depositor**' is the role played by those persons which provide the information to be preserved.

Designated Community: a group of potential users of the information held by the archive/data repository, who should be able to understand that information independently. The designated community is defined by the archive and this definition may change over time. From <https://public.ccsds.org/pubs/650x0m2.pdf>.

Long-term: a period of time long enough for there to be concern about the impacts of changing technologies (including support for new media and data formats) and of a changing [designated community](https://public.ccsds.org/pubs/650x0m2.pdf), on the information being held in a data repository. This period extends into the indefinite future. From <https://public.ccsds.org/pubs/650x0m2.pdf>.

Long-term Preservation: the act of maintaining information understandable over the [long-term](https://public.ccsds.org/pubs/650x0m2.pdf). Long-term preservation entails:

- 1) Secure, backed-up storage of digital data
- 2) Procedures and activities (a.k.a. data curation) that ensure that the information in digital data remains understandable over time

Long-term preservation is necessary since technologies and the requirements of the [designated community](https://public.ccsds.org/pubs/650x0m2.pdf) of the data repository may change over time. For example, software to open certain file formats may become unavailable, or the research community may change their definition of a concept.

Metadata: data about other data. In this context, data about an RDR data collection and the data files of that collection. Metadata in the RDR include the information on the collection's landing page as well as files in the collection that have been labelled as documentation files.

Publish: to make information available to the public or a specific group of people.

Appendix

Table: what goes where in the Radboud Data Repository?

	Data Acquisition Collection (DAC)	Research Documentation Collection (RDC)	Data Sharing Collection (DSC)
Function:	To preserve research data long-term for internal reuse and to support research quality and integrity	To preserve and document the research process and thereby support research quality and integrity	To ensure the accessibility and reusability of digital research data for the wider scientific community and society at large and thereby enhance the impact of research executed at Radboud University
When:	During the data collection phase	Throughout the research project	At the end of the research project, prior to publication of the journal article based on the research data
Form:	Raw data in its original form, meaning there has been no manipulation that would limit further analyses	Processed data that has been digitised, translated, transcribed, cleaned, or validated. Analysed data in the shape of models, graphs, tables, texts and so on that are based on the raw and processed data	This can be identical to the Data Acquisition or Research Documentation Collection, or a combination of both
Primary content:	Raw, unprocessed research data, such as: ✓ Wet-lab experiments ✓ Measurements ✓ Questionnaire responses ✓ Audiotapes ✓ Videotapes ✓ Photographs ✓ Fieldwork notebooks	Any data that document the research process, such as: ✓ Processed or cleaned data ✓ Processing pipelines ✓ Intermediate results ✓ Laboratory notebooks ✓ Transcripts ✓ (Preliminary) analysis scripts ✓ (Preliminary) models and algorithms ✓ Audit trails ✓ Qualitative data analysis	Any data that may be useful for replication studies or reuse, such as: ✓ Final figures, tables, etc. ✓ Raw data ✓ Processed or cleaned data ✓ Processing pipelines ✓ Transcripts ✓ Analysis scripts ✓ Models and algorithms ✓ Qualitative data analysis

Identifiability:	Identifiable research data that has not (yet) been anonymised, although it can be pseudonymised to protect the privacy of participants	Data is identifiable, pseudonymised, or anonymised depending on the research needs	Data is identifiable, pseudonymised, or anonymised depending on the informed consent and needs of data re-users
Secondary content:	✓ Documentation: files that explain the context (how the research was done), the structure (overview of folders and files), and the content of the collection (what data can be found in the dataset) ✓ Codebooks: explains all the concepts and/or variables that are present in the data files (e.g. meaning, numerical values, units of measurement, classification schemes used, labels of questions asked) ✓ README: provides information about a dataset and is intended to help ensure that the data can be correctly interpreted, by yourself at a later date or by others when sharing or publishing data ✓ Informed consent template: this provides information on the informed consent that was obtained and what re-users can and cannot do with the data ✓ Version logs: records the changes that are made in new versions ✓ Methodologies: describe the collecting, cleaning, transforming, aggregating, or calculating procedures or methods ✓ Software: information on the software (version) that was used to analyse and/or compress files ✓ General information: Overview of project history, objectives, or hypotheses		
Data that are not to be deposited in the RDR:	? Data generated by students (BSc or MSc) should only be deposited in the RDR if they are of scientific value and serve as the foundation of a scientific report other than the student's thesis ✗ Applications for the ethical committee, research program, or grant provider ✗ Approvals of ethical committee ✗ Copyrighted materials for which the researcher does not own the rights ✗ Correspondence with other people (e-mail, social media, blogs, wiki) ✗ Signed informed consent forms (<u>archive elsewhere for the same duration as the dataset is archived!</u>) ✗ Data Management Plans ✗ Journal articles or other (versions of) academic publications ✗ Key files of pseudonymised datasets		