

DataDOI Policy

1. Collection policy

Data stored in DataDOI can be open access, embargoed, restricted or closed. Open access means that the dataset is freely accessible and downloadable. Other options, embargoed, restricted or closed have some specific limitations.

Data from all fields and subjects can be uploaded to DataDOI collections.

2. Submission policy

The data should be in final publishable state. It is in the responsibility of the submitters to guarantee the rights and permissions for uploading and sharing the data. Datasets should contain documentation (README.txt file) describing the purpose and terms of reuse as well as explaining how data was collected and how to use this specific dataset.

The maximum upload size of one dataset is 2GB. For larger datasets please contact the repository administrator.

Although DataDOI accepts all kinds of well-known file formats, we recommend using those listed below. This ensures that datasets are suitable for long-term preservation and are also usable in the future.

Recommendable formats:

- textual data
 - TXT, ODT, DOCX, PDF (A-2a or A-2u)
- tabular data
 - ODS, XLSX, CSV
- presentations
 - ODP, PPTX, PDF (A-2a or A-2u)
- markup documents
 - HTML+CSS, XML+XSD+XSL, JSON
- photos
 - TIFF, PNG, JP2, DNG
- audio
 - AVI, MOV, MP4, MXF
- digital signature format
 - ASICE

Submitted datasets should have the standard metadata to assure that data is findable, accessible, interoperable and reusable and effectively curatable. Metadata is openly accessible. DataDOI's metadata is reachable via [OAI-PMH](#) and is available for harvesting.

Metadata may be reused in any medium without prior permission for not-for-profit purposes. Metadata must not be reused in any medium for commercial purposes without formal permission.

Metadata is exported in several standard formats, for example Dublin Core, DataCite Metadata Schema, etc.

Datasets may be submitted only by registered members. A dataset can be submitted at any time but will not be publicly visible if the publisher or the funder has established an embargo period. Data will be publicly available when publishers' or funders' embargo period has expired.

If there is a new and improved version of a dataset then it should be uploaded next to the previous version(s) and given a version number (e.g. dataset_name_v1, dataset_name_v2 etc.).

Any copyright violations are entirely the responsibility of the datasets' authors. If DataDOI administrators receive proof of copyright violation, the relevant dataset will be removed immediately.

The uploaded content should not consist of private, confidential, sensitive or restricted information that might harm others.

3. Deposit license and policy

When a dataset is submitted, a license must be included to give directions on what terms and conditions data can be used.

The proprietary rights of the data are transferred to the university by a contract of employment (academic staff) or to another umbrella organization by a written document (act of handing over intellectual property).

Data will be made public by the Creative Commons license [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/).

A third party whose data has been used for the results of the grant may set restrictions for the usage of the data. In this case, these limitations must be taken into account when licensing the data, i.e. the use of the data can only be licensed to the extent of the rights granted by the third party (i.e. the amount of rights received by the university from third parties).

4. Re-use policy

Authors who upload their datasets should consider that their data is reused at some degree. It is important to add a license to a dataset in order to explain terms and conditions for using a dataset. To make reusing a dataset more efficient it is advisable to add a citation to make citing easier and faster for other researchers.

5. Take-down policies and embargoes

Once a dataset is published and given a DOI it won't be removed from the repository because it may have been cited and reused. Although there can be some exceptions and special case, for example cases of data falsification and corruption. A dataset can be removed from the repository wholly, partly or with

only some parts being restricted. Access can be restricted temporary or permanently. DOI remains publicly available

Datasets can be deposited under open, closed, restricted or embargoed access. Datasets deposited under an embargo are provided with an end date for embargo. After the embargo period has ended, the dataset will be automatically made public. Closed or restricted datasets need access rights given by the dataset owner.

If the University or a third party, whose data has been used for creating the grant results, wishes to submit a patent or useful model application in order to protect the data, the publication of the data must be postponed until such application has been submitted.

6. Preservation policy

Dataset replicas are stored in University of Tartu servers and is backed up on a nightly basis. Virtual servers are backed up every 10 days.

Datasets are retained for the lifetime of the repository. In case of closing the repository, all efforts will be done to find a replacement institution and/or repository. DataDOI provides the longevity of its collections in any circumstances and does everything possible to retain its content.

7. Rights

DataDOI has a right to remove or restrict datasets that violate the copyright or consist of private, confidential, sensitive or restricted information that might harm others.

DataDOI gives itself a right to change terms of this policy according to the best practices of open research data.

8. Responsibilities

DataDOI provides access, retaining and longevity to all submitted datasets and its metadata.

DataDOI ensures that submitted datasets follow the rules of open access and FAIR principles.

In case of any questions please contact: datadoi@datadoi.ee