

Storage and Publication of Research Data

What is a Repository?

<latin> storage

Repository system:

hardware & software combination to store data in order to make the **accessible**

Archive

the purpose of an archive is **long time storage**

Research data repositories provide

Research data repositories provide

- a searchable catalogue of research data

Research data repositories provide

- a searchable catalogue of research data
- data saved in a format that can still be opened in the future

Research data repositories provide

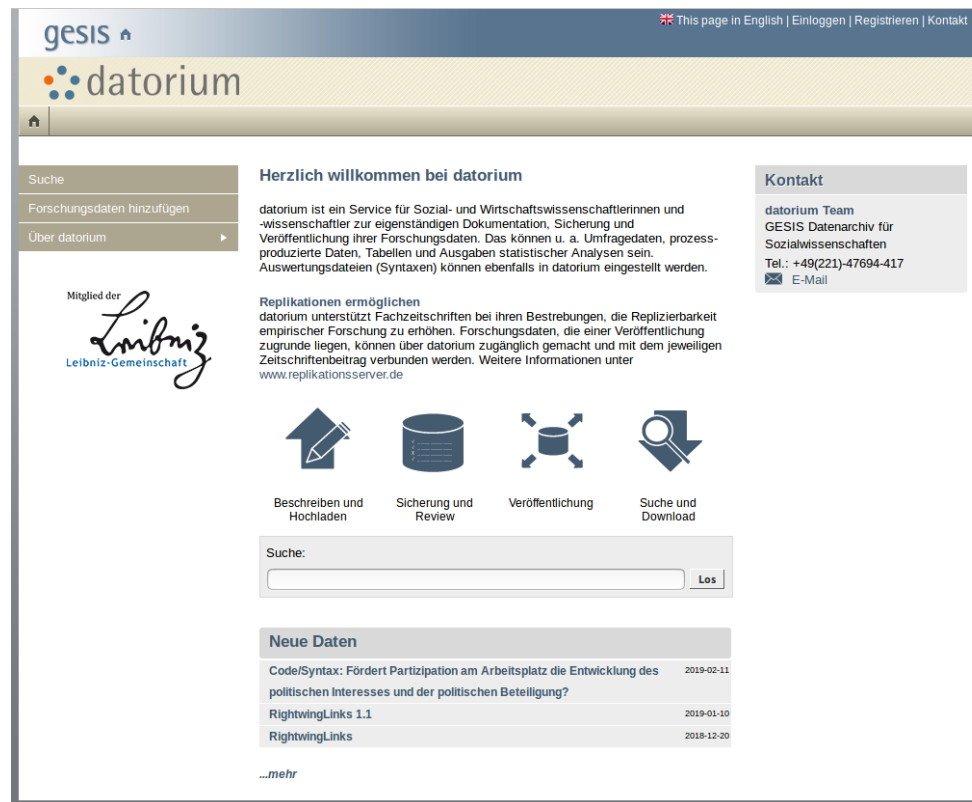
- a searchable catalogue of research data
- data saved in a format that can still be opened in the future
- combined information packages of data and documentation

Research data repositories provide

- a searchable catalogue of research data
- data saved in a format that can still be opened in the future
- combined information packages of data and documentation
- traceable data

How to choose a repository

Discipline-specific repositories are usually the best place for your data:



<https://datorium.gesis.org>

Other examples of discipline specific repositories

- **arXiv**: physics, math, computer science
- **CLARIN-D**: linguistic and language data
- **VerbundFDM**: Research and Information in Education
- **GenBank**: genetic sequence database ...

List of accredited data centres for social and economic data

<https://www.ratswd.de/forschungsdaten/fdz>

Trusted repositories: Data Seal of Approval)



Trusted repositories: Data Seal of Approval)

- Data can be **found** on the Internet



Trusted repositories: Data Seal of Approval)

- Data can be **found** on the Internet
- Data is **available** (unique rights and licenses)



Trusted repositories: Data Seal of Approval)

- Data can be **found** on the Internet
- Data is **available** (unique rights and licenses)
- Data can be **reused** (format)



Trusted repositories: Data Seal of Approval)

- Data can be **found** on the Internet
- Data is **available** (unique rights and licenses)
- Data can be **reused** (format)
- Data is **reliable/trustworthy**



Trusted repositories: Data Seal of Approval)

- Data can be **found** on the Internet
- Data is **available** (unique rights and licenses)
- Data can be **reused** (format)
- Data is **reliable/trustworthy**
- Data have PIDs so that they can be **referenced**



Research data repository @ UHH



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

ZENTRUM

FÜR NACHHALTIGES

FORSCHUNGSDATENMANAGEMENT

FDR @ UHH

FDR @ UHH

Open

When uploading, you decide whether your files are made available as open, restricted or closed, or whether they are provided with an embargo period.

FDR @ UHH

Open

When uploading, you decide whether your files are made available as open, restricted or closed, or whether they are provided with an embargo period.

Citable

Each entry in the repository is provided with a persistent DOI (Digital Object Identifier) and can be found and cited via this.

FDR @ UHH

Open

When uploading, you decide whether your files are made available as open, restricted or closed, or whether they are provided with an embargo period.

Citable

Each entry in the repository is provided with a persistent DOI (Digital Object Identifier) and can be found and cited via this.

Secure

sets of up to 50 gigabytes can be uploaded via the web interface (more upon request) and are securely stored at the RRZ of the Universität Hamburg (S3 memory).

FDR @ UHH

Open

When uploading, you decide whether your files are made available as open, restricted or closed, or whether they are provided with an embargo period.

Citable

Each entry in the repository is provided with a persistent DOI (Digital Object Identifier) and can be found and cited via this.

Secure

sets of up to 50 gigabytes can be uploaded via the web interface (more upon request) and are securely stored at the RRZ of the Universität Hamburg (S3 memory).

Sustainable

The retention period of your files is at least 10 years.

FDR @ UHH

Open

When uploading, you decide whether your files are made available as open, restricted or closed, or whether they are provided with an embargo period.

Citable

Each entry in the repository is provided with a persistent DOI (Digital Object Identifier) and can be found and cited via this.

Secure

sets of up to 50 gigabytes can be uploaded via the web interface (more upon request) and are securely stored at the RRZ of the Universität Hamburg (S3 memory).

Sustainable

The retention period of your files is at least 10 years.

Flexible

All materials that are available in digital form are suitable for long-term storage in the repository. Any file format is possible.

Publication steps

1. Register with B identifier (Shibboleth)
2. Prepare & Upload Data
3. Enter Metadata
4. Save and publish

Prepare files

Prepare files

- File format: as open as possible

Prepare files

- File format: as open as possible
- File selection:
 - check for hidden files
 - check the first in order (this will show in the preview)
 - check for sensitive files / information you might not want to publish
 - delete obsolete backup or versioning information
 - check for big files (max 50GB / file)
 - unzip if necessary

Prepare files

- File format: as open as possible
- File selection:
 - [check for hidden files](#)
 - check the first in order (this will show in the preview)
 - check for sensitive files / information you might not want to publish
 - delete obsolete backup or versioning information
 - check for big files (max 50GB / file)
 - unzip if necessary

Prepare files

- File format: as open as possible
- File selection:
 - check for hidden files
 - [check the first in order \(this will show in the preview\)](#)
 - check for sensitive files / information you might not want to publish
 - delete obsolete backup or versioning information
 - check for big files (max 50GB / file)
 - unzip if necessary

Prepare files

- File format: as open as possible
- File selection:
 - check for hidden files
 - check the first in order (this will show in the preview)
 - **check for sensitive files / information you might not want to publish**
 - delete obsolete backup or versioning information
 - check for big files (max 50GB / file)
 - unzip if necessary

Prepare files

- File format: as open as possible
- File selection:
 - check for hidden files
 - check the first in order (this will show in the preview)
 - check for sensitive files / information you might not want to publish
 - delete obsolete backup or versioning information
 - check for big files (max 50GB / file)
 - unzip if necessary

Prepare files

- File format: as open as possible
- File selection:
 - check for hidden files
 - check the first in order (this will show in the preview)
 - check for sensitive files / information you might not want to publish
 - delete obsolete backup or versioning information
 - **check for big files (max 50GB / file)**
 - unzip if necessary

Prepare files

- File format: as open as possible
- File selection:
 - check for hidden files
 - check the first in order (this will show in the preview)
 - check for sensitive files / information you might not want to publish
 - delete obsolete backup or versioning information
 - check for big files (max 50GB / file)
 - **unzip if necessary**

Prepare files

- File format: as open as possible
- File selection:
 - check for hidden files
 - check the first in order (this will show in the preview)
 - check for sensitive files / information you might not want to publish
 - delete obsolete backup or versioning information
 - check for big files (max 50GB / file)
 - unzip if necessary
- once published the data can not be changed, every change creates a new version (DOI)

Prepare the documentation

Prepare the documentation

- ✓ Research project description

Prepare the documentation

- ✓ Research project description
- ✓ underlying hypotheses

Prepare the documentation

- ✓ Research project description
- ✓ underlying hypotheses
- ✓ information on the collection of data

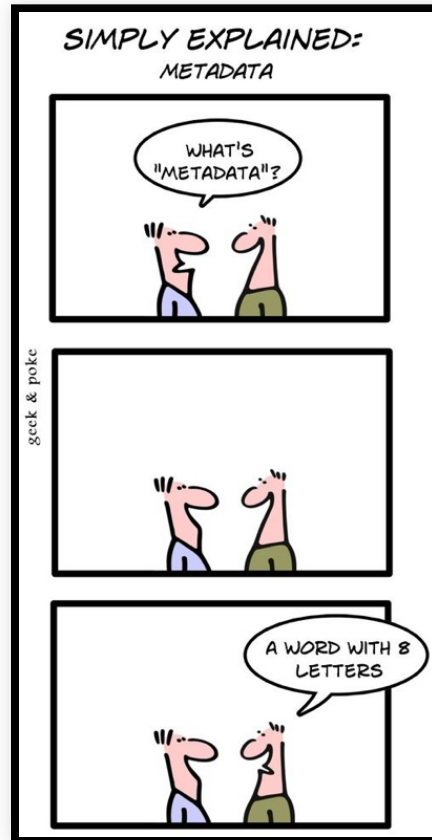
Prepare the documentation

- ✓ Research project description
- ✓ underlying hypotheses
- ✓ information on the collection of data
- ✓ measures for data cleansing or data weighting

Prepare the documentation

- ✓ Research project description
- ✓ underlying hypotheses
- ✓ information on the collection of data
- ✓ measures for data cleansing or data weighting
- ✓ explanations for codes, labels and data versions

What is good metadata?



<http://geekandpoke.typepad.com/geekandpoke/2010/04/meta.html>

Metadata types

Metadata type	Objective	Example
Descriptive metadata	Identification and traceability	Title, Author, Abstract, Keywords...
Administrative metadata	Creation and storage	Rights management and access rights, storage...
Structural metadata	Structure of the Resource	Chapters, sections, structure, files...
Technical metadata	technical characteristics	Scope, coding, technical structure, formats

Exercise “Metadata”

- Get into two groups
- Choose an object
- Describe the metadata (10 min)
- Each group then briefly presents the results

Why is metadata important?

- Metadata ensures long-term data usage and interpretation
- Well-documented data are also understandable by people who were not involved in the research
- Structured metadata can be read and processed by machine for fast, easy and reliable searches

Overview of discipline-specific metadata standards

*<http://www.dcc.ac.uk/resources/metadata-standards>
<http://rd-alliance.github.io/metadata-directory/subjects/>*

- Keywords
 - Data Documentation Initiative (Social Sciences)
 - GESIS; Datorium or SSOAR (Social Science Open Access Repository)
 - Dublin Core (bibliographic data format)
 - Marc21 (bibliographic data format)

Required Metadata

DOI

enter DOI if your files have already been published; if empty a new DOI will be assigned

Publication Date

enter publication date if your data have already been published; current date as default

Authors

specify as many authors as you like, preferably with ORCID

Description

enter abstract if you publish a scientific paper; for files describe the data (or use the abstract of the data journal publication)

License

Open | Embargoed | Restricted | Closed Access

Recommended Metadata

Keywords

keyword that will help to find your data

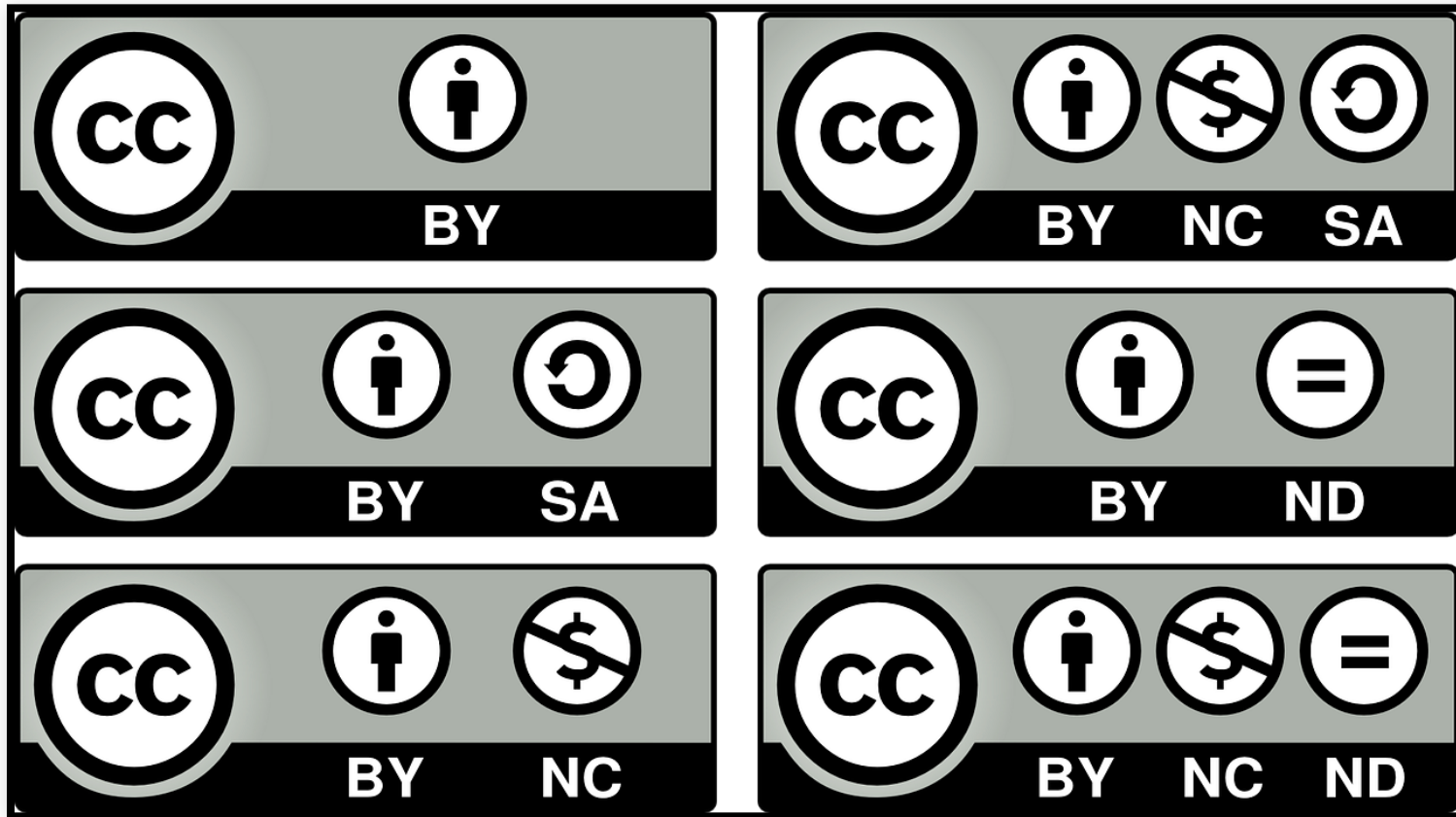
Funding

the FDR is reporting to the European Commission via OpenAIRE

Related Identifiers

specify identifiers of related publications

What license should I choose?



Can you copyright data?

- There is no copyright on ideas, concepts, principles or discoveries
 - raw data
 - factual information
 - news

Copyright applies to original creative expression (not the result of hard work)

Databases and enriched data

- EU and German Urheberrecht (as well as US copyright) protects compilations of data (database), provided that a significant investment has been made
- the protection applies to the compilation, the relation not to the data itself
- enriched data may also be protected by intellectual property rights

Legal issues and research data

- Research data in general are free of rights, i.e. for geological data, personal data, surveys, bibliographic data etc. there is no ownership and therefore copyright does not apply here
- derivatives are protected, like texts in which the data are described, evaluated or visualized can be protected
- also descriptions, pictures or other data which go beyond pure information can be protected
- the freedom of information is an essential principle of the intellectual property rights

Contractual and other restrictions

- for personal data the personal protection law applies regardless of copyright
- contractual conditions may restrict the use of data
- In contrast to intellectual property rights, the terms and conditions of the contract apply only to the person with whom the contract has been concluded
- funding agencies usually require a statement on the handling of data in advance (Research Data Management Plan)

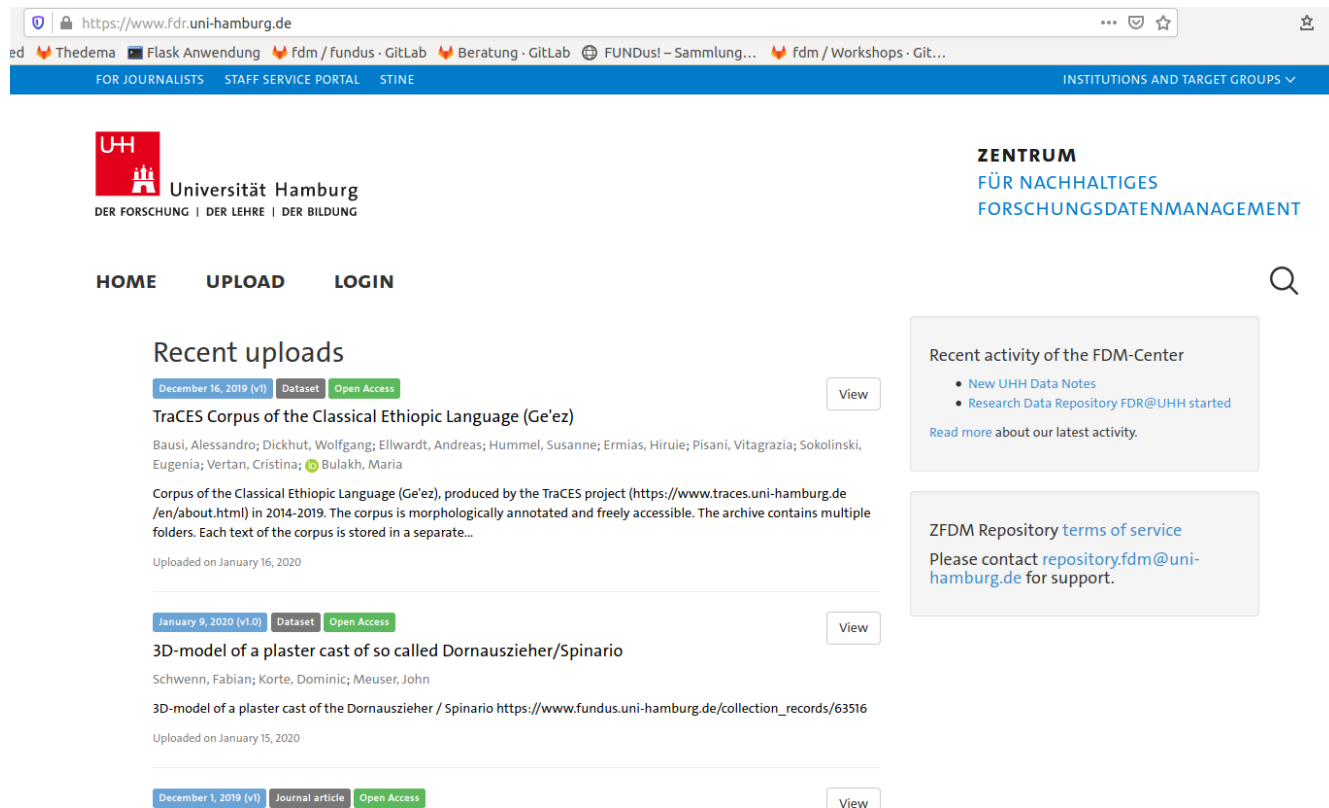
Literature



[http://hup.sub.uni-](http://hup.sub.uni-hamburg.de/volltexte/2019/195/pdf/HamburgUP_KreutzerLahmann_Rechtsfragen)

[hamburg.de/volltexte/2019/195/pdf/HamburgUP_KreutzerLahmann_Rechtsfragen](http://hup.sub.uni-hamburg.de/volltexte/2019/195/pdf/HamburgUP_KreutzerLahmann_Rechtsfragen)

FDR @ UHH



The screenshot shows the FDR@UHH website interface. At the top, there is a navigation bar with links for 'FOR JOURNALISTS', 'STAFF SERVICE PORTAL', 'STINE', and 'INSTITUTIONS AND TARGET GROUPS'. Below this, the University of Hamburg logo and name are displayed on the left, and the ZENTRUM FÜR NACHHALTIGES FORSCHUNGSDATENMANAGEMENT logo is on the right. The main content area features a 'Recent uploads' section with three entries:

- TraCES Corpus of the Classical Ethiopic Language (Ge'ez)**
December 16, 2019 (v1) | Dataset | Open Access | View
Bausi, Alessandro; Dickhut, Wolfgang; Ellwardt, Andreas; Hummel, Susanne; Ermias, Hiruie; Pisani, Vitagrazia; Sokolinski, Eugenia; Vertan, Cristina; Bulakh, Maria
Corpus of the Classical Ethiopic Language (Ge'ez), produced by the TraCES project (<https://www.traces.uni-hamburg.de/en/about.html>) in 2014-2019. The corpus is morphologically annotated and freely accessible. The archive contains multiple folders. Each text of the corpus is stored in a separate...
Uploaded on January 16, 2020
- 3D-model of a plaster cast of so called Dornauszieher/Spinario**
January 9, 2020 (v1.0) | Dataset | Open Access | View
Schwenn, Fabian; Korte, Dominic; Meuser, John
3D-model of a plaster cast of the Dornauszieher / Spinario https://www.fundus.uni-hamburg.de/collection_records/63516
Uploaded on January 15, 2020
- December 1, 2019 (v1) | Journal article | Open Access | View**

On the right side of the interface, there are two informational boxes:

- Recent activity of the FDM-Center**
 - New UHH Data Notes
 - Research Data Repository FDR@UHH started[Read more about our latest activity.](#)
- ZFDM Repository terms of service**
Please contact repository.fdm@uni-hamburg.de for support.



Describe and publish your research data at UHH data notes

<http://uhh.de/fdm-datanotes>

- Must be a member of UHH to publish
- Data has to be stored in long term repository (e.g. FDR) which provides a DOI
- Free of charge

Zentrum für nachhaltiges Forschungsdatenmanagement

Monetastr.4

20146 Hamburg

www.fdm.uni-hamburg.de

Marie Ryan marie.ryan@uni-hamburg.de Tel.: +49 40

42838- 9598

Dr. Iris Vogel iris.vogel@uni-hamburg.de Tel.: +49 40

42838-8862

Universität Hamburg

Career Center

Hochschulübergreifende
Koordinationsstelle Frauenstudien
Frauen- und Geschlechterforschung

Zentrale Bibliothek
Frauenforschung &
Gender Studies

Monetastr.

Universität Hamburg

Helpful links

- Centre for Research Data Management (<https://www.fdm.uni-hamburg.de/>)
- FDM Blog @UHH <https://forschungsdaten.blogs.uni-hamburg.de/blog-beitraege/>
- Research Data Repository <https://fdr.uni-hamburg.de>
- DMP Online <http://dmp.fdm.uni-hamburg.de/> (coming soon)
- www.forschungsdaten.org
- www.forschungsdaten.info
- Verbund Forschungsdaten-bildung <https://www.forschungsdaten-bildung.de/index.php?la=de>