This is the publication of the document “Research Data Management Plan of the Modelling Human-Environment Interactions Research Group”.

This document is part of a group of three documents that are used by the Modelling Human Environmental Interactions Group at the Institute of Geographical Sciences, Freie Universität Berlin (Germany) for complying with the Freie Universität´s “Statutes for Safeguarding Good Research Practice”*, Section 9 Usage rights, paragraph (1): “Regulations on rights to use research results are to be produced at the earliest possible date and documented in writing.”

Related documents are:

- Researcher Data Agreement (http://dx.doi.org/10.17169/refubium-46356)
- Student Data Agreement (http://dx.doi.org/10.17169/refubium-46355)

* https://www.fu-berlin.de/en/sites/gwp/informationen/satzung/index.html

Authors:

Piquer-Rodriguez, Maria[1+], https://orcid.org/0000-0001-8181-5049

Reitzug, Sophia[1], https://orcid.org/0009-0008-4785-9863

Hübner, Andreas[2], https://orcid.org/0000-0001-7342-9789

Asef, Esther[2], https://orcid.org/0000-0003-2411-4953

[1] Institute of Geographical Sciences, Freie Universität Berlin, Germany

[2] University Library, Freie Universität Berlin, Germany

[+] Corresponding author: maria.piquer-rodriguez@fu-berlin.de

Please cite the document as:

Piquer-Rodriguez, M., Reitzug, S., Hübner, A., Asef, E. (2025) Research Data Management Plan of the Modelling Human-Environment Interactions Research Group. http://dx.doi.org/10.17169/refubium-46357

This work is published under a CC0 Public Domain Dedication, https://creativecommons.org/publicdomain/zero/1.0/

Research Data Management Plan (DMP) of the Modelling Human-Environment Interactions (MHEI) Research Group, Freie Universität Berlin

Version 1.0 – November 2024

Content

1. MHEI Storage Facilities 2
2. MHEI folder structure and access rights 3
3. MHEI Backup Solutions 4
4. MHEI Metadata Documentation 5
5. MHEI Archiving Solutions 5
6. MHEI Publishing Solutions 6

This document is supported and approved by the Freie Universität legal department and the University Library Research Data Management team.

The Modelling Human-Environment Interactions (MHEI) Research Group´s Research Data Management Plan (MHEI-DMP) provides students and researchers of the MHEI Group with guidelines about data management in the research group. It addresses storage facilities, folder structures, backup solutions, metadata documentation, and archiving facilities used by the MHEI Group. The MHEI DMP is based on the Freie Universität´s Research Data Policy and the Statutes for Safeguarding Good Research Practice.

Additionally, all students and researchers of the MHEI Group will have to agree on an individual, (project) specific data management agreement which complements the MHEI-DMP.

1. MHEI Storage Facilities

Every member of the MHEI group will have several storage facilities on their laptops additionally to the hard disk. We advise to only store data and working progresses on the hard disk if no internet connection is available and to transfer the data to the personal or group network as soon as possible. In addition to personal laptops, the MHEI Group also has access to a work station. The work stations' purpose is to facilitate the running of large script and large data processing. It is accessible for all member of the MHEI Group upon reservation.

1.1 Personal network drive

Each researcher is provided with their own, personal network drive provided by the university. The file and folder structure of this drive is the choice and responsibility of each researcher but MHEI Group advises to follow the structure specified in section 2. This personal network drive is intended to be the researcher´s main working location.

The storge capacity of this driver is 25 GB by default, which is automatically assigned to every Freie Universität (FU) employee. Upon request to FUB IT, this can be expanded to 50 GB (more GB are possible, but have to be paid for).

- The drive is automatically backed up every 3-4 h by FUB IT.
- If any data is lost, data recovery is possible up to 3 months by FUB IT.
- This network drive is by default provided on the researcher´s personal FU computer, usually it is drive H.
- Information on how to add this drive to your computer (in case it is not already present): *tip4u://157 Nutzung von Home- und Gruppenlaufwerken* (https://www.zedat.fu-berlin.de/tip4u_157.pdf) or have a look at the MHEI Group Wiki (https://wikis.fu-berlin.de/display/mhei/Welcome+Document).

1.2 MHEI Group network drive

Additional to the personal network drive, there is the MHEI Group network (I:) AG-Modellierung-MU (\\campus.fu-berlin.de\daten\GEOWISS\Geographie)). Here, mainly documents for organizational and administrative purposes of the MHEI Group are stored.

1.3 MHEI Group Trove drive

"Trove" is a large central data storage space provided by FU especially for research data. The MHEI Group has access to storage capacity of this FU central resource. The MHEI Group trove drive is connected to the MHEI Group work station and is accessed via (Z:) mmui (\\trove.storage.fu-Berlin.de). The MHEI Group trove drive is MHEI Group´s main storage for big geographical data and also the main data backup device. The structure of this locations is divided into **backup** and the **personal user's folders**.

- ***Backup*** *(raw data)*: PDFs, scripts and processed data do not belong here. The idea is to keep this space for large, (usually unprocessed) data that everybody of the MHEI Group can access and use. The backup-folder also contains the groups data archive as a subfolder.
- ***Personal user´s Folders*** *(processed data):* All members of the MHEI Group also have a personal user's folder on the MHEI Group Trove drive where they can save data additionally to their personal network drive. The MHEI Group Trove drives not intended for processing data, but rather to save and backup your already processed data.

1.4 Hard drive

In addition to the MHEI Group Trove drive, there is a hard drive connected to the work station that is mirroring the trove. There is an automated backup implemented between these two to ensure that MHEI data is safely stored.

2. MHEI folder structure and access rights

For our backup solution, the trove drive, we want to keep data easily finable and understandable. The following folder structure should also be used for the personal network drives so the transfer to the trove is less workload.

- File Naming:
  - snake_case: that means each word in a name is in lower case and separated by underscores
  - name_date_version: the order in which we name the file so they appear in a logical order and facilitate a good overview over progress and file history
  - Version control through date: if more than one version per day, then use: _v01, _v02, _final
  - YYYYMMDD
  - Example: potapov_global_cropland_2019_20240920_v01

- Folder structure on MHEI Group Trove drive:

[archived]
[sensitive_data]/ [example_project1]
/ [example_project2]
[spatial_data] / [continent] / [latin_america] / [example_dataset]
/ [north_america] / [example_dataset]
/ [south_america] / [example_dataset]
/ [global] / [example_dataset]
/ [study_area] / [baap] / [example_dataset]
/ [chaco] / [example_dataset]
[users] / [example_user1]
/ [example_user2]

- as for personal projects this is a **best-practice structure example**:

[example_project] / [figures]
/ [manuscript]
/ [presentations]
/ [r] / [input_data]
/ [script]
/ [output_data]

- Access and management:

Users' rights are set up with FUB_IT after signing the Data Agreement of the MHEI.

- [archived]: this folder is managed and accessed by the head of the group and the administrator - student assistant (currently Sophia Reitzug)
- [sensitive_data]: managed and accessed by the head of the group, administrator and the leading researcher
- [spatial_data]: managed and accessed by all
- [users]: managed and accessed by the head of the group, administrator and each user. We encourage research data to be structured clearly with folders containing key inputs, key outputs, code, temporary data, always following our best practice example of above.

3. MHEI Backup Solutions

- Everyone is responsible for their own project data and the frequency of backing up the personal data on the MHEI Group Trove drive (personal user folder) to ensure no data is lost. We advise a period of three months. The MHEI Group Trove drive is connected to the hard drive via the work station and with the help of the automated file synchronizations these two storage locations are mirrored when running the backup (currently only Ahuvit Trumper and Sophia Reitzug can start the backup process). The backup on the hard drive

- The current backup folder structure is to be respected and kept.
- The student assistant of the MHEI Group checks compliance with agreements on data management on the trove drive and the group network drive every three months. In particular, especially the naming and correct storage location of files are monitored to ensure coherent backing up of the MHEI Group Trove drive data.

4. MHEI Metadata Documentation

To ensure the compliance with FAIR principles, all data that is stored on the trove drive should have metadata documentation.

- Create a readme file for every dataset (or compilation of data)
- Has to include:
  - Source/Producer (Example: *Potapov et al. 2021 + link* or *Ahuvit Trumper*)
  - Date of access
  - Downloader (Example: downloaded by Ahuvit Trumper from Google Earth Engine)
  - Format (Example: tif, .csv, …)
  - Description (Example: Global cropland dataset for 2019, 30x30 meters resolution)
- Use txt format and xml format for publishing. A metadata generator, for example: https://dhvlab.gwi.uni-muenchen.de/datacite-generator/ can be helpful.

5. MHEI Archiving Solutions

When data is no longer used and in the foreseeable future not needed as e.g., ancillary data for other research or researchers, the data should be archived in the trove drive archive folder. This includes data from interviews, processed spatial data, scripts and whole research projects or thesis. If the data belongs to a publication, the zipped folder should contain all data that is needed to reproduce the analysis (scripts, sources, workflow description, metadata, references, …) to ensure FAIR principles. Before archiving the data, the MHEI Group leader should be informed.

- Archive data that:
  - Underpins publications (especially peer reviewed journal articles, but also other research output like bachelor, master or PhD theses, reports, [are there more examples, specific to the group?])
  - Cannot be easily reproduced
  - Is important for further research
  - Has to be kept because of legal or contractual reasons
- Minimum of 10 years
- Location: (Z:) mmui (\\trove.storage.fu-berlin.de) _Backup\archives
- Add metadata/readme file
- Zipping as a project folder when data is no longer used: all data in the archive is zipped
- Deletion of data from personal users' space on the trove drive
- Table with contents of zipped files (optional)

- Folder Structure Archive

  [research_projects] / [publication_xxx] / [scripts]
  [spatial_data]
  [references]
  [visiting_researcher] / [project_data]
  [scripts]
  […]
  [data_bsc] / [semester] / [student_name]
  [data_msc] / [semester] / [student_name]

The MHEI Group will work out a workflow with the FUB IT to transfer data to the FU archive.

6. MHEI Publishing Solutions

Publishing data is key for ensuring FAIR principles. All MHEI Group members will publish data from publications, projects and thesis following Open Access Standards Freie Universität Berlin.

A suitable and open data repository should be selected. The FU repository refubium is one option but other, more topic specific data repositories might be also suitable.

Before publishing the data, the following criteria should be met:

- Properly documented with metadata (see *MHEI Metadata Solutions*)
- Reviewed for quality, e.g., content of the study, methodology, relevance, legal consistency and documentation of materials
- Searchable and discoverable in catalogues (or databases)
- Citable in article