

KiltHub Data Deposit Requirements Guide

Welcome to KiltHub

Carnegie Mellon University is now providing a comprehensive repository for research data and other forms of scholarly outputs. Powered by figshare, KiltHub is an online platform for citing, sharing, and discovering research data and other forms of scholarly output. This user guide will walk you through the process of depositing, publishing, and sharing your research and scholarship on KiltHub. For more information on creating collections and projects, please see our collections and projects guides.

Contents

[What is KiltHub?](#)

[Required Metadata Form Fields](#)

[Optional Metadata Form Fields](#)

[Required README File Records](#)

[Additional information and resources for completing the README file](#)

What is KiltHub?

KiltHub is a place to make all your research data and other scholarly outputs citable, shareable, and discoverable. It's a place to share your supplementary research to make them first class outputs. Share both positive and negative results and get credit for all your research.

The most important things about KiltHub:

- Your research is assigned a **DOI**. This means you can cite it as a research output alongside your paper or as a freestanding piece of data.
- KiltHub complies with **funder mandates** around **making data openly accessible** and **stored in perpetuity**.

Please see the [KiltHub Deposit Guide](#) for guidance on how to deposit, publish, and share your research and scholarship.

For more information, please contact:

website: <https://kiltHub.cmu.edu>

twitter: [@KiltHub](#)

email: kiltHub@andrew.cmu.edu

support site: library.cmu.edu/kiltHub

KiltHub Data Deposit Requirements Guide

Required Metadata Form Fields

- **Title:** Descriptive title that includes the name of the research project.
- **Authors:** Include the names of those involved in creating the data, either by name, full email, or ORCID.
- **Categories:** Please select the subject or subjects that best represent your data. These entries are from the Australian and New Zealand Standard Classification (ANZSRC) Fields of Research (FOR) system. You can choose more than one and either select from the drop-down menu or search for your subject area. Click [here](#) for the full category list. If you can't find exactly what you're looking for, we recommend finding the best option and getting more specific in the Keywords field.
- **Item Type:** Please select the item type that best describes your data (e.g. dataset, fileset, code). A *fileset* includes many different files types in one deposit, however you should select *dataset* if the primary component of the deposit is data. While KiltHub accepts all forms of data for deposit, only certain file and item types are viewable in the browser. Explore [Item Types](#).
- **Keywords:** Add keywords (or phrases) that describe the dataset. These may include words describing the topic or methods and you may add as many as you like to make the dataset discoverable. Some keywords will auto populate but you may also add new keywords. These should be more specific than the Category entries and help others find your data. Add as many relevant keywords as you need - just hit enter after each keyword. KiltHub remembers ones you've entered previously and will suggest keywords as you begin typing.
- **Description:** Provide an abstract describing your project including the context of the dataset and methods used. If you'd like to include a link to a publication or other related research items (data, code, images, etc.) you may include it here or in the references section below.
- **License:** The default license is CC-BY (Creative Commons Attribution). It is the most permissive (and preferred) license. There are specific licenses for software and code (e.g. MIT, Apache). The full list of available licenses can be found [here](#).

KiltHub Data Deposit Requirements Guide

Optional Metadata Form Fields

- **References:** References provide context through related content. Add links (full URLs or hyperlinked DOIs) to references, such as published articles or software code.
- **Funding:** If applicable, include the Funding Agency and grant number.
- **Publisher Statement:** This is a field for inserting text and the full citation to the published version of your material(s) if required by the publisher to comply with their Green Open Access standards. If a set Publisher Statement is not required, please use this KiltHub Standard Entry: “This is the [insert name of Version of Publication] of, “Full APA-Style Citation of Publication with hyperlinked DOI.” Copyright [Year of Publication] [Name of Publisher].”
- **Date:** Enter the date that the item was published or presented. If left blank, the items will inherit the date of deposit as the publication date.

Required README File

The README file provides context to your dataset, outside and independent of any publication, presentation, or web-based resource. All deposited data and software code must include a README file.

Address each of the items below (when applicable) in the README file, referencing where this information can be found if not mentioned in full in the README. You can download a sample README file template [here](#) to customize. We recommend using a plain text file.

Required information for README file:

- **Authors:** Include contact information for first at least the first author and corresponding author (if not the same), specifically email address, phone number, and institution. Contact information for all authors is preferred.
- **Directory of Files in Dataset:** List and define the different files included in the dataset, which serves as its table of contents. This information can be included in the README file, in a separate file, or as part of the data file. If it is in a separate file or in the data file,

KiltHub Data Deposit Requirements Guide

explain where this information is located and ensure that it is accessible without specialized software. We recommend using plain text files or tabular plain text CSV files exported from spreadsheet software.

- **File Naming Convention:** Define your File Naming Convention (FNC), the framework used for naming your files systematically to describe what they contain, which could be combined with the Directory of Files. For more information and examples, see <http://library.cmu.edu/datapub/dms/data/document>
- **Data Description:** A data description, dictionary, or codebook defines the variables and abbreviations used in a dataset. This information can be included in the README file, in a separate file, or as part of the data file. If it is in a separate file or in the data file, explain where this information is located and ensure that it is accessible without specialized software. (We recommend using plain text files or tabular plain text CSV files exported from spreadsheet software.)

If applicable, additional required information for README file:

- **Software:** If specialized software(s) generated your data or are necessary to interpret it, please provide for each (if applicable): software name, version, system requirements, and developer. (See below for more on basic descriptive requirements and explicit instructions for additional contextual information.)
- **Equipment:** If specialized equipment generated your data, please provide for each (if applicable): equipment name, manufacturer, model, and calibration information. Be sure to include specialized file format information in the data dictionary. (See below for further information.)
- **Dates of Data Collection:** List the dates and/or times of data collection.

KiltHub Data Deposit Requirements Guide

Additional information and resources for completing the README file

- **Data Description, Dictionary, and Codebook Examples:**
 - [Descriptors and Measurements of the Height of Runaway Slaves and Indentured Servants Codebook](#)
 - [USGS National Hydrography Dataset Data Dictionary](#)
 - [Water Chemistry Data Dictionary](#)
 - [Federal Funder Public Access Requirements Data Dictionary](#)

- **Software Code and Links:** Requirements for software description in the README file and for additional software data deposits vary based on the software's use and development context. The descriptive information below is only required for software that either generated data in the submission or that is required to interpret data in the submission.
 - The basic required descriptive fields in the README file are:
 - **Name:** A descriptive name for the software program. If the program does not have a brand, commercial or project name, please provide a descriptive title summarizing the program's function and purpose.
 - **Version:** Version name and / or number of the software. If unavailable please mark as "not applicable".
 - **System Requirements:** Minimally, system requirements should provide the operating system name and version number needed for execution. More detailed descriptive information regarding software dependencies is welcomed and will increase the potential for future data re-use.
 - *If the software is open source, please provide in the README file (if available):*
 - **Executable URL:** A valid URL linking to a compiled executable compatible with the system requirements described above.
 - **Source Repository URL:** A valid URL linking to the software's source code repository.
 - *If the software is not freely available and / or closed source, please provide in the README file (if available):*
 - **Developer:** Name of the entity responsible for the creating the software (if not self-developed).
 - **Product URL:** A valid URL linking to the software's product page or documentation.
 - *If you developed the software, please consider providing (if applicable):*
 - A copy of the software's binary executable compatible with the system requirements described above.
 - A source snapshot or distribution if the source code is not stored in a publicly available online repository.

KiltHub Data Deposit Requirements Guide

- **Equipment:** If your data was generated by a specific piece of commercial equipment, please include in the README file:
 - **Manufacturer:** Name of commercial entity responsible for creating the equipment
 - **Model:** Model number and / or name of the equipment.
 - **Embedded Software / Firmware Name and Version:** Only provide this information if the data generation is dependent on it, and you believe the above manufacturer and model number are not sufficiently descriptive.