README File Checklist

Make your dataset easy to use and understand

A good README file is like a map that helps others navigate your data. Use this checklist to make your README clear, complete, and user-friendly, enabling others to make the most of your hard work. This checklist provides general guidance but may need to be adjusted to fit the specific requirements of your research.

The elements within the box are considered essential (core) metadata elements, while the rest are regarded as extended elements.

GENERAL INFORMATION

- [ ] **Title of Dataset:** Clearly state the name of the dataset.
- [ ] **Author(s)/Contributor(s):** Include names, affiliations, corresponding author and its contact information.
- [ ] **Date of Creation/Release:** Specify when the dataset was created or made publicly available.
- [ ] **DOI or Persistent Identifier:** Provide a unique identifier if available (depend on the repository).
- [ ] **Abstract/Objective: Provide a brief description of the dataset and its objective and research hypothesis**
- [ ] **Dataset Version:** Indicate the version number (e.g., v1.0).
- [ ] **Funding Information:** Mention any funding sources or grant numbers.
- [ ] **Acknowledgments:** List individuals or organisations that contributed.

DATASET OVERVIEW

- [ ] **Dataset Description:** Briefly describe the dataset content including basic information on the methodology.
- [ ] **File Formats:** If using uncommon formats is unavoidable, list the file formats and how to open the files.
- [ ] **Data Dictionary:** Provide a data dictionary or table describing each variable/column.
- [ ] **Directory Structure:** Describe the organisation (list) of files and folders.

METHODOLOGY

- [ ] **Data Collection Methods:** Explain how the data was collected, including tools, devices and protocols.
- [ ] **Data Processing:** Describe any cleaning, transformation, or preprocessing steps.
- [ ] **Data Analysis:** Specify the software and their versions used to analyse the datasets.
- [ ] **Ethical Approvals:** Mention any IRB or ethical board approvals if applicable.

USAGE AND ACCESS

- [ ] **Licence:** Specify the licence under which the dataset is shared (e.g., CC BY, MIT).
- [ ] **Related Resources:** Cite any papers, articles, protocols or posters related to the dataset.
- [ ] **Usage Instructions:** Provide guidance on how to open and use the files.
- [ ] **Citing the Dataset:** Include a citation format or example.

For more specific README templates covering different species and data types, please visit our GitHub repository

What would you want to know if you were encountering this dataset for the first time?

Think as a Consumer of your Data not the Producer!

Anticipate questions users might ask and address them upfront.
Remember, a well-documented dataset is more likely to be reused and cited!

Funding: MRC grant MR/X009726/1

THE UNIVERSITY of EDINBURGH

Medical Research Council

Contact us: *bio_rdm@ed.ac.uk*