Shut Up & Write

Documentation: README.txt

Welcome

Welcome: What you can expect

This slide set accompanies a short presentation on documentation and readme-files for an online "Shut Up & Write" session. Here, researchers, students, or other faculty come together for one hour to one and a half hours to write – silently but in company. The session is usually started with a brief introduction and finishes with a short feedback round.

To ensure that all participants feel welcome and safe, we agree to be nice to each other.

You are welcome to reuse, remix, and re-publish these slides. For license information and citation, see the final slide (page 13).

Schedule

10 min   Setup & Introduction

Download a template for a README.txt from
https://data.research.cornell.edu/content/readme

Make a copy of the file, save the template somewhere smart

Why do we write documentation for our data?

45 min   Shut Up & Write

5 min   Short feedback round on what worked & what didn't

Good Research Practice

Good Research Practice: Why Documentation?

- all **you** need to know about your data
- all **your collaborators** need to know about your data
- all you need to know about your collaborators' data
- all **anyone else** who wants to work with your data needs to know about it
- more information is better than less
- structured information is better than unstructured information
- no documentation is the worst
- **start early, save time & have less stress**
- re-use **templates** for good documentation, READMEs, etc.
- you **must** have documentation when you archive your data

Anatomy of a README-file for datasets   #1

- This **DATSETNAMEreadme.txt** file was generated on **YYYY-MM-DD** by **NAME**
- General Information
  - Title of Dataset
  - Author Information
  - Date/Timespan of data collection
  - Geographic location of data collection
  - Information about funding sources
- Sharing/Access Information
  - Licenses/restrictions
  - Links to publications that cite the data
  - Links to other publicly accessible locations of data (repositories, etc.)
  - Links to ancillary data sets
  - Was the data derived from another source
  - Recommended citation for the dataset

- Data & File Overview
  - File List
  - Relationship between files
  - Additional related data collected that wasn't included in current data package
  - Are there multiple versions of the dataset?
- Methodological Information
  - Description of methods used for collection/generation of data
  - Methods for processing the data
  - Instrument- or software-specific information needed to interpret the data
  - Standards and calibration
  - Environmental/experimental conditions
  - Describe quality-assurance procedures performed on data
  - People involved with sample collection, processing, analysis, submission

8

- Data-specific Information for: [FILENAME]
  - Number of variables
  - Number of cases/rows
  - Variable list
  - Missing data codes
  - Specialized formats or other abbreviations used

- Repeat data-specific information for each dataset, folder or file – as appropriate

- Update your README file whenever there is a significant change, datasets are added, modified, removed, etc.

- Yes, you can – and should! – customize the template to fit your data

Shut Up & Write

45min

Turn off your microphone. If you don't want to listen to anything, turn off your headphones, too. If you wish to listen to the ambient sound of someone else typing, leave the headphones on, your instructor is writing their README.txt, too! You may switch off your video, but it feels more enjoyable writing in good company, even if it's virtual only!

Debriefing & Feedback

Thank you for participating!

Questions or Comments about Research Data?
research-data@uio.no