NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN

SciLifeLab

README and Markdown

Introduction to Data Management Practices course

NBIS DM Team

data-management@scilifelab.se

https://nbisweden.github.io/module-readme-markdown-dm-practices/

Learning objectives

1. Understand the purpose and importance of README files in research projects.
2. Learn the main sections to include in various types of READMEs.
3. Apply basic Markdown syntax to create and format README files.

What is a README file?

A README explains what this thing is and how to use it.

- A README is a **plain-text documentation file**
- Usually named README.md
- Sits at the **top level** of a project
- Answers the question: **“What am I looking at?”**

Why use a README?

- Reduces onboarding time for collaborators
- Supports reproducibility
- Helps with FAIR principles
- Makes datasets usable later on in the project, for yourself and others

When use a README?

- A dataset you plan to:
  - Share internally
  - Upload to a repository
  - Attach to a paper
- A code project (even small scripts)
- A folder with:
  - Multiple files
  - Non-obvious filenames
  - Generated outputs
- A project lasting more than a few weeks

"If it's worth saving, it's worth explaining."

markdown

```
# Urban Heat Island Dataset – Copenhagen

## Description
This dataset contains temperature measurements collected from
June–August 2024 as part of the UHI project.

## Data
- `raw/` – unprocessed sensor data
- `processed/` – cleaned datasets used in analysis
- `metadata.csv` – sensor locations and calibration info

## Methods
Temperature recorded every 10 minutes using HOBO sensors.
Processing performed in R (scripts in `scripts/`).

## Contact
Jane Doe (jane.doe@university.edu)
```

Different types of README files

1. Study level (root folder)
    - General information, project organisation, usage
2. Data level (subfolders)
    - General information, data organisation, data reuse or restrictions
3. Code level (e.g., on GitHub)
    - Purpose, dependencies, examples demonstrating the usage

Study level README

This README file was generated on [YYYY-MM-DD] by [NAME]

GENERAL INFORMATION
- Study title
- Study description
- Principal investigator
- Link to DMP

ORGANIZATION
- Folder structure
- File naming conventions
- File formats

OpenAI. (2025, March 21). AI-generated image of a modern workspace with a README file. DALL·E.

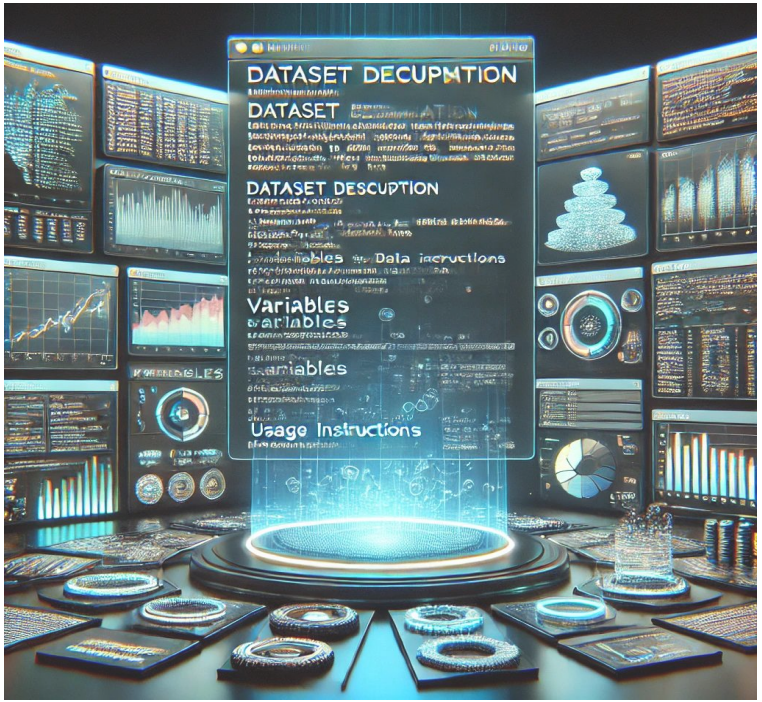
Data level README

This README file was generated on [YYYY-MM-DD] by [NAME]

GENERAL INFORMATION
- Dataset description
- Data delivery report

ORGANIZATION
- File naming conventions

DATA REUSE
- Accession numbers
- License
- Use restrictions



OpenAI. (2025, March 21). AI-generated image of a futuristic data lab with a README file at the data level. DALL·E.

What is Markdown?

Lightweight markup language

Plain-text formatting syntax

What are the benefits of Markdown?

- **Versatile:** usable for documents, websites, README files, and emails
- **Portable:** plain text, not tied to proprietary software (e.g. MS Word)
- **Platform-independent:** works across operating systems and tools
- **Future-proof:** readable even if the original software becomes obsolete

NBIS

How do you write Markdown?

```
# Project overview
---
## **Data description**
---
### _File organization_
---
#### _**Data processing**_
---
##### `Usage instructions`
---
## ~~Contact information~~
```

Project overview

Data description

File organization

Data processing

`Usage instructions`

~~**Contact information**~~

Exercises

Link to exercises in the collaborative notes

Try on your own, but ask for help if you get stuck

Have fun!