Universitetet i Sørøst-Norge

URGE

UNIVERSAL README FILE GENERATOR

Dag-Even Martinsen Torsøe

Ali Abdurhman Kelil

Ahmet Dogan

DOI: https://doi.org/10.7557/5.8194

What is a ReadMe file?

A file containing summary information about the dataset

Contains information about context, processing methodology and processing tools

Archived with the data files to ensure that others can read, understand and reuse the data (the R in FAIR)

URGE
USN README FILE GENERATOR

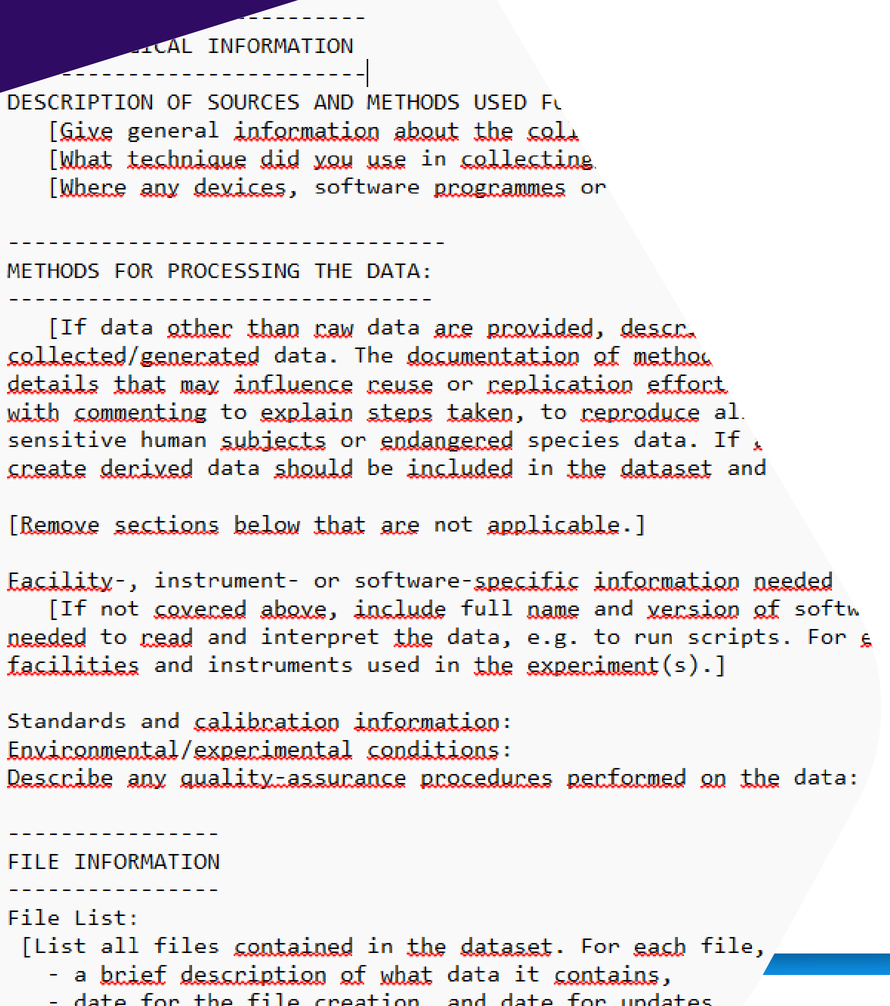
The “Traditional" README file template (. txt). Ex: DataverseNo templae

```
GENERAL INFORMATION
------------------
// Title of Dataset:
// DOI:
// Contact Information
<The person to be contacted for questions about the dataset
     // Name:
     // Institution:
     // Email:
     // ORCID:

<Whenever applicable, the following information should be registered in the metadata schema of DataverseNO. In the text below, remove fields/lines that are not applicable,
and leave the rest unchanged. >
// Contributors: See metadata field Contributor.
// Data Type: See metadata field Data Type.
// Date of Collection: See metadata field Date of Collection.
// Geographic location: See metadata section Geospatial Metadata.
// Funding sources: See metadata section Funding Information.

// Description of dataset:
<(Short) description of what the dataset is about, including reference to related project(s) and publication(s), if applicable. Should correspond to the information entered
in the metadata fields Description and Related Publication.>

--------------------------
METHODOLOGICAL INFORMATION
--------------------------
<Note! If the documentation referred to is not openly available through a persistent URL, it must be added here or uploaded as file(s) to the dataset.>

<Note! It may generally be considered appropriate to have overlap in the methods section of a research data README file with citation of the original article. See Committee
on Publication Ethics (COPE) guidance on text recycling: https://publicationethics.org/resources/guidelines-new/text-recycling-guidelines-editors-0.>

// Description of sources and methods used for collection/generation of data:
<Include links or references to sources, publications, reports or other documentation (e.g. survey questionnaires, interview protocols, Preregistrations or Registered
Reports) containing (experimental) study design or protocols, or other collection techniques used, as well as personnel involved in data collection/generation.>

// Methods for processing the data:
<If data other than raw data are provided, describe how the submitted data were processed from the raw or collected/generated data. The documentation of methods used for dat
processing should include (if applicable): details that may influence reuse or replication efforts; data cleaning and analysis syntax; code or algorithms, with commenting to
explain steps taken, to reproduce all reported findings; de-identification procedures for sensitive human subjects or endangered species data. If applicable, code, algorithm
or command files used to create derived data should be included in the dataset and referred to in this section.>

<Remove sections below that are not applicable.>
```

The “Traditional" README file template (. txt). Ex: DataverseNo templae, continued

```
// File List:
<List all files (or folders, as appropriate for dataset orga
conventions, instead of listing individual files. For each
not obvious from file extension (e.g. .txt). If necessary,

// Relationship between files, if important:

// Is this an updated version of a dataset published on Data
<This question is relevant if you make changes to the origi
the information below and repeat for each file that was upd
future>
     Version number of dataset:
     File name:
     Why was the file updated?
     When was the file updated (YYYY-MM-DD)?:
     What was changed?


---------------------------------------
DATA-SPECIFIC INFORMATION FOR: [FILENAME]
---------------------------------------
<If file-level documentation is provided elsewhere (e.g., as integrated file metadata e.g. ), it is sufficient to refer to this documentation below.>

<Repeat this section for each dataset, folder or file, as appropriate. Recurring items may also be explained in a common initial section.>

<For TABULAR data, provide a data dictionary/code book containing the following information:>
// Variable/Column List:
<List variable/column name(s), description(s), unit(s) of measurement, decimal separator (comma or point), value labels, and source(s) as appropriate for each.>
// Missing data codes:
<Define codes or symbols used to indicate missing data.>
// Specialized formats or other abbreviations used:

<For QUALITATIVE data (e.g. interviews, images), provide a data list:>
// Data List:
<The data list should contain the following information, as appropriate: interview identifier (ID); age; gender; occupation, organisation; location; place of interview; date
of interview; transcript file name. Information included should be detailed enough to enable sub-setting and filtering, but not detailed enough to enable identification of
participants where confidentiality has been promised.>

// Contextual Information:
<Where applicable, provide biographical or other contextual information relating to the interview or interviewee, either embedded at the beginning of the transcript files, as
a summary page, or added below, as appropriate.>


---------------------------
```

The "Traditional" README file template (. txt)



```
----------
...ICAL INFORMATION
--------------------------
DESCRIPTION OF SOURCES AND METHODS USED F
   [Give general information about the col
   [What technique did you use in collecting
   [Where any devices, software programmes or

---------------------------------
METHODS FOR PROCESSING THE DATA:
---------------------------------
   [If data other than raw data are provided, descr
collected/generated data. The documentation of metho
details that may influence reuse or replication effort
with commenting to explain steps taken, to reproduce al
sensitive human subjects or endangered species data. If
create derived data should be included in the dataset and

[Remove sections below that are not applicable.]

Facility-, instrument- or software-specific information needed
   [If not covered above, include full name and version of softw
needed to read and interpret the data, e.g. to run scripts. For
facilities and instruments used in the experiment(s).]

Standards and calibration information:
Environmental/experimental conditions:
Describe any quality-assurance procedures performed on the data:

----------------
FILE INFORMATION
----------------
File List:
 [List all files contained in the dataset. For each file,
   - a brief description of what data it contains,
   - date for the file creation, and date for updates
```

- Less intuitive, and not user-friendly
- Requires extra work and time
  - Irrelevant fields must be deleted
  - Help texts must be deleted
- Does not ensure completion of mandatory fields
- Not machine-readable (.txt)

URGE
USN README FILE GENERATOR

Let's go live!