How to Manage Research Data : Handbook for Researchers

Poljak, Ljiljana; Celjak, Draženko; Dorotić Malič, Ivana; Matijević, Marta; Posavec, Kristina; Turk, Ivana

Conference presentation / Izlaganje na skupu

How to Manage Research Data

Handbook for Researchers

Pubmet, 17.9.2020.

Background

- growing need for research collaboration and research reproducibility
- the infrastructure for research data is becoming increasingly available
- national RDA node was established in Srce in 2019
- University of Rijeka Library, the University of Split Library, the City and University Library Osijek, and the National and University Library in Zagreb

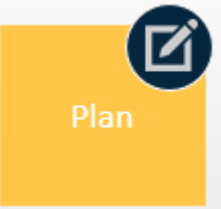
Research data – How to Manage it?

Introduction

- What is research data?
- Why share research data?
- How to share data – how to make data FAIR?

Pubmet, 17. 9. 2020.

Before the research



Data Management Plan (DPM)

the process of organizing the research: planning, collection, processing, storage and distribution of data in accordance with FAIR principles

Recommendation… DPM

- Data and documentation
- Ethical and legal aspect, safety
- Storing and preserving
- Sharing and reuse

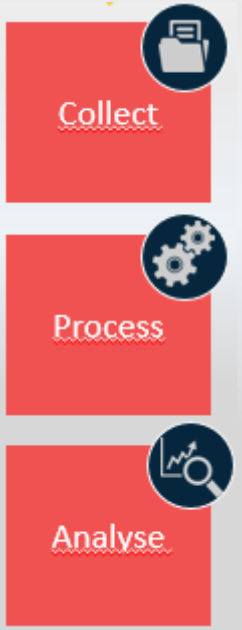
Store and preserve

- How the data will be stored and shared during the project (on a computer, USB memory or cloud computing, data encryption, backup, etc.)
- Where the data will be permanently stored and published after the completion of the research / project (selection of the appropriate repository for permanent data storage)

Pubmet, 17. 9. 2020.

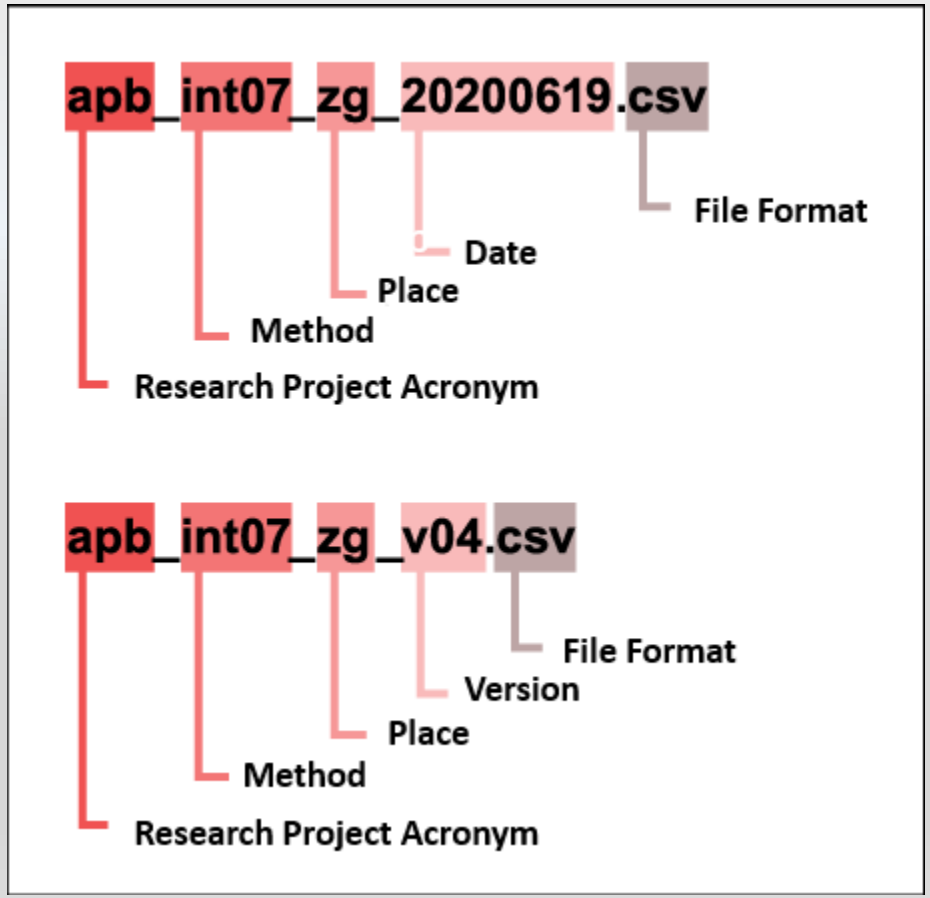
During the research

- Naming and organizing research data
- Research data versioning
- Guidelines for selecting a format for permanent data storage
- Documentation and metadata
- Reliable data storage during research



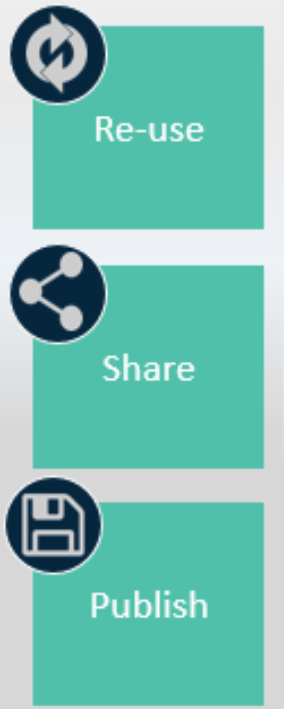
Recommendation... Naming and organizing

- Name convention must be used by all members of the research team
- Avoid using a similar name for multiple files.
- File names must not be too long, no more than 32 characters.
- Using letters and numbers of ASCII codes (a-z, A-Z and 0-9). Avoid using periods or special characters in the file name, such as &, *% #; * ()! @ $ ^ ~ ‘{} []? <>.
- Use a underscore, minus or camel case instead of a space.
- Use the date in the ISO 8601 standard (YYYYMMDD).



Pubmet, 17. 9. 2020.

After the research



- **Disclosure** of personal and sensitive data
- Defining the **right to use** research data (Licensing)
- **Publishing** research data
- **Citing** research data

Recommendation... Which licenses to use for research data?

- In general, the less restrictive the license - the greater the possibility of re-use and proper use.

- CC0, ie the transfer of a work to the public domain, enabling others to re-use this data freely. When using data published under the CC0 license, naming is not legally required, but citation is a norm imposed by scientific integrity.

- CC BY 4.0, allowing modification, processing and sharing. It is necessary to name the author, link it to the original work (dataset) and indicate all changes made. In the context of data, by using a CC BY license there is a possibility of accumulation of namings. When data is used repeatedly or in a combination of different data sets, indicating all the necessary namings can create a burden.

- The best option is to use CC0 with a naming request. There may be an author's note "Please attribute my data" or a similar phrase with the published data. The use of licenses CC0 and CC BY 4.0 is recommended by the European Commission for Horizon2020 projects and RDA.

Useful tools

- **Amnesia** (https://amnesia.openaire.eu/) online tool for anonymizing quantitative data stored in .csv and .txt format
- **RDM Advice & Tips** (https://www.lcrdm.nl/en/rdm-advice-tips) tips on privacy, anonymization and pseudonymization of research data
- **Creative Commons (CC)** (https://creativecommons.org/) licence system
- **Dabar** (https://dabar.srce.hr/repozitoriji) repository for storage and sharing of the final version of data sets and other results of the research project, mainly after the end of the research
- **Re3data** (https://www.re3data.org/) registry gathers more than 2000 repository from all fields of science
- **Zenodo** (https://zenodo.org/) multi-disciplinary open repository
- **COPTR** (Community Owned Digital Preservation Tool Registry https://coptr.digipres.org/Main_Page) online tool that helps researchers find data preservation tools
- **UK Data Service** (https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation/quantitative.aspx) tips for anonymizing quantitative and qualitative data

Glossary

- Anonymization
  - An irreversible process of exchanging personal data after which a person can no longer be identified directly or indirectly.
- Encryption
  - …
- FAIR
- DMP
- GDPR
- Licence
- …

Benefits

- provides a comprehensive set of guidelines and recommendations based on research data lifecycle
- overview of key topics in research data management
- includes a list of helpful tools to facilitate the management of research data and to ensure good data practices

Pubmet, 17. 9. 2020.

Ljiljana Poljak, University of Split Library
Draženko Celjak, SRCE- University Computing Centre
Ivana Dorotić Malič, University of Rijeka Library
Marta Matijević, National and University Library in Zagreb
Kristina Posavec, SRCE- University Computing Centre
Ivana Turk, City and University Library Osijek

Celjak et al. *Istraživački podaci - što s njima? : Priručnik o upravljanju istraživačkim podacima*. Zagreb: Srce,2020.
urn:nbn:hr:102:810090

Pubmet, 17. 9. 2020.