

Research data management when working with children and youth



DAY 1

Workshop
Ljubljana, Slovenia
27 – 28 March 2023



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101008589



SPEAKERS

- IRENA VIPAVC BRVAR, *Slovenian Social Science Data Archives*
- SONJA BEZJAK, *Slovenian Social Science Data Archives*
- MARIANNE HØGETVEIT MYHREN, *Sikt*
- MATEJA SEDMAK, *Science and Research Centre Koper*
- TONI BABAROVIĆ, *Institute of Social Sciences Ivo Pilar, Zagreb*

Event organised by ADP - Slovenian Social Science Data Archives

(<https://www.adp.fdv.uni-lj.si/>)





Introduction

*Sonja Bezjak and Irena Vipavc Brvar,
Slovenian Social Science Data Archives*



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101008589



Workshop
Ljubljana, Slovenia
27 – 28 March 2023



COhort cOmmunity Research and Development Infrastructure Network for Access Throughout Europe → COORDINATE



The aim of the **COORDINATE** project is **to mobilise the community** of researchers and organisations that will drive forwards the coordinated development of **comparative birth cohort panel and associated survey research** in Europe which **focus on children's well-being**.

The infrastructural community network brought together by **COORDINATE** will **promote the harmonisation of and improve access to international survey data**, in particular panel survey data, in the study of children and young people's well-being as they grow up.

The research that **COORDINATE** will complete, using a child-centric approach, continues the research initiated in MyWEB and ECDP projects, which will support elements of the **preparatory phase of Europe's first cross-national accelerated birth cohort survey of child well-being: EuroCohort - Growing Up in Digital Europe (GUIDE/EuroCohort)**.

<https://www.coordinate-network.eu/>



Project partners

- Alma Mater Studiorum – University of Bologna
- Centerdata
- **Consortium of European Social Science Data Archives and ERIC, European Research Infrastructure Consortium → ADP, Slovenia**
- European Centre for Social Welfare Policy and Research
- GESIS
- Geary Institute, University College Dublin
- Institute of Social Sciences Ivo Pilar
- Institut national d'études démographiques
- Ipsos
- Kantar Public
- Manchester Metropolitan University
- Pompeu Fabra University, Research and Expertise Centre for Survey Methodology (RECSM)
- Royal Netherlands Academy of Arts and Sciences
- **Science and Research Centre Koper**
- University College London
- University Institute of Lisbon
- University of Essex
- University of Helsinki
- cApStAn

Slovenian Social Science Data Archives

ADP - Arhiv družboslovnih podatkov



- Founded in 1997 → 25th anniversary
- Slovenian **national research data centre** for social sciences
- **Member of CESSDA** ERIC
- Status of a **trust-worthy archive** (CoreTrustSeal since 2018)
- involved in EU and national projects

<https://www.adp.fdv.uni-lj.si/eng/>

ADP's mission

To ensure and promote *sustainable services* of **ingest, storage and access** to *quality research data from the field of Slovenian social sciences* and broader, with *potential for secondary analysis*.

Main services:

- **Acquiring** important research data from a wide range of social sciences
- **Appraisal** of submitted research data and their **selection** for deposit **Ingesting and processing** research data and other documentation, together with the creation of metadata
- Long-term digital **preservation** (AIP), **access** and **re-use** for scientific, educational and other purposes (DIP)
- **Training** researchers on:
 - research data management
 - re-use of research data
 - **Promotion** of open data and open science (students, librarians, journals, citizens...)

QUICK FACTS ABOUT ADP

HOW TO GET DATA?



FIND →



REGISTER



→ ANALYZE

HOW TO DEPOSIT DATA?



RECORD →



PREPARE



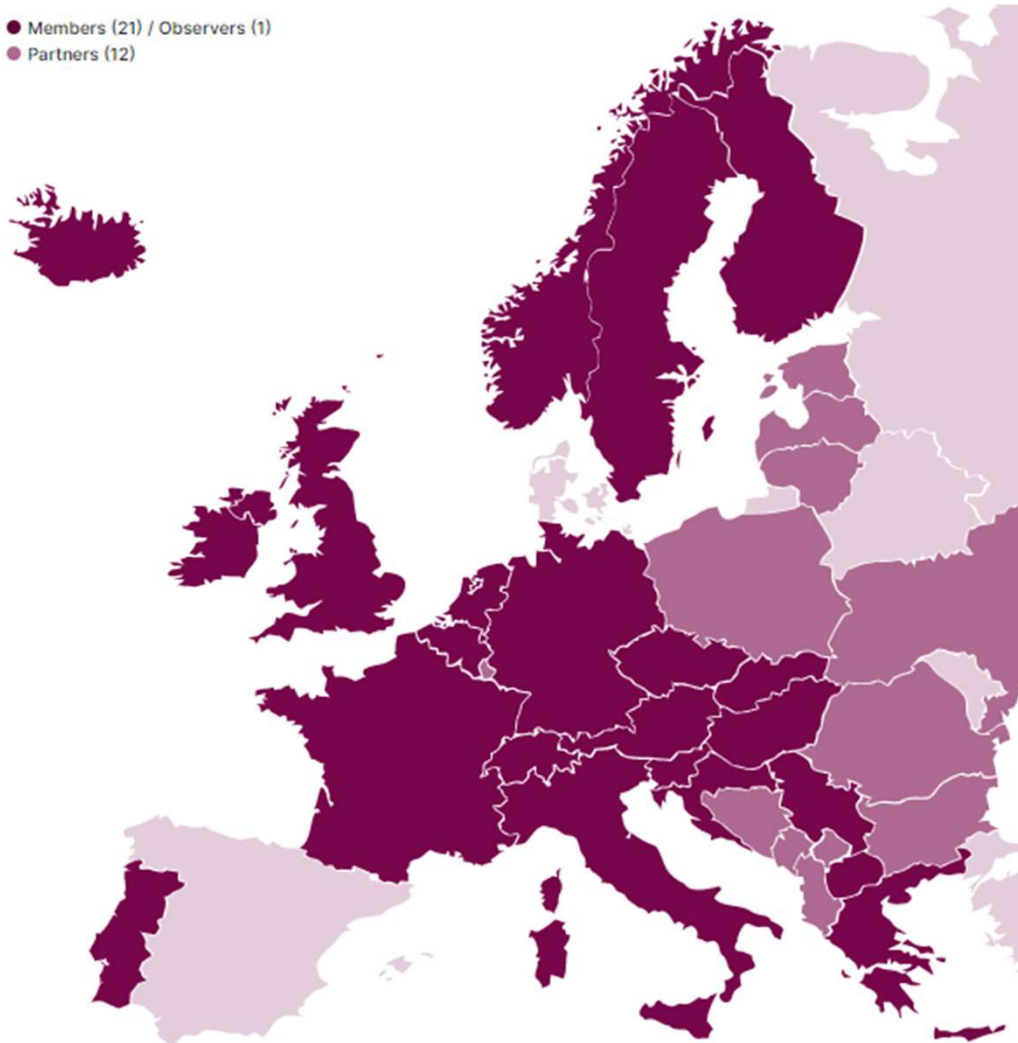
→ DEPOSIT

<https://www.adp.fdv.uni-lj.si/eng/>

- **For 775 social science studies** research data accessible in a data catalogue
- **1000 users registered per year** (90 % education, 10 % scientific/research purpose)
- **Cca. 500 units of research data reused** for detailed secondary-analysis per year

CESSDA - Consortium of European Social Science Data Archives

● Members (21) / Observers (1)
● Partners (12)



“Member countries seek to increase the scientific excellence and efficacy of European research in the social sciences”

Key tasks:

Developing **standards and best practices** around the management and archiving of social science data.

Facilitating access to important data resources

Work done by **developing tools, training and co-ordinating network.**

[CESSDA data catalogue.](https://datacatalogue.CESSDA.eu/)

(<https://datacatalogue.CESSDA.eu/>)



AGENDA FOR MONDAY, 27th March 2023

9:00 - 9:10	Welcome and introduction Sonja Bezjak and Irena Vipavc Brvar
9:10 - 10:00	Getting to know each other All participants
10:00 - 11:00	Data Management Planning in general Sonja Bezjak and Irena Vipavc Brvar, Slovenian Social Science Data Archives
11:00 -11:30	<i>Coffee break</i>
11:30 - 13:00	Legal grounds for processing personal data Marianne Høgetveit Myhren, Sikt
13:00 - 14:00	<i>Lunch</i>

AGENDA FOR MONDAY, 27th March 2023

14:00 - 15:00	Processing personal data: Hands-on Marianne Høgetveit Myhren
15:00 - 15:30	<i>Coffee break</i>
15:30 - 17:00	Challenges in doing research with migrant children Mateja Sedmak, Science and Research Centre Koper GUIDE pilot survey: Example of the data management plan and content of child and parental consent Toni Babarović, Institute of Social Sciences Ivo Pilar, Zagreb
19:00 - 21:00	<i>Dinner</i>

AGENDA FOR TUESDAY, 28th March 2023

9:00 - 10:30	Data discovery, Longitudinal and secondary data: lecture and hands-on Sonja Bezjak and Irena Vipavc Brvar, Slovenian Social Science Data Archives
10:30 - 11:00	<i>Coffee break</i>
11:00 - 12:00	Update your DMP Sonja Bezjak and Irena Vipavc Brvar, Slovenian Social Science Data Archives
12:00 - 12:30	Wrapping up the workshop
12:30 - 13:30	<i>Lunch</i>



Getting to know each other

All participants



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101008589



COORDINATE
COhort cOMmunity Research and Development
Infrastructure Network for Access Throughout Europe



Getting to know each other

Name, country, research interests



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101008589



PARTICIPANTS' RESEARCH INTERESTS

1 CHILDREN

2 ADOLESCENTS

3 CULTURE

4 SCHOOL & EDUCATION

5 SPORT

6 SOCIAL INTEGRATION

7 POLITICAL AND CIVIL
PARTICIPATION

8 CHILD AND FAMILY LAW

9 HEALTH

10 COVID-19

11 POVERTY

12 MIGRATIONS

13 DIGITAL DEVICES



Data Management Planning in general

Sonja Bezjak and Irena Vipavc Brvar,

Slovenian Social Science Data Archives



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101008589



Workshop
Ljubljana, Slovenia

27 – 28 March 2023

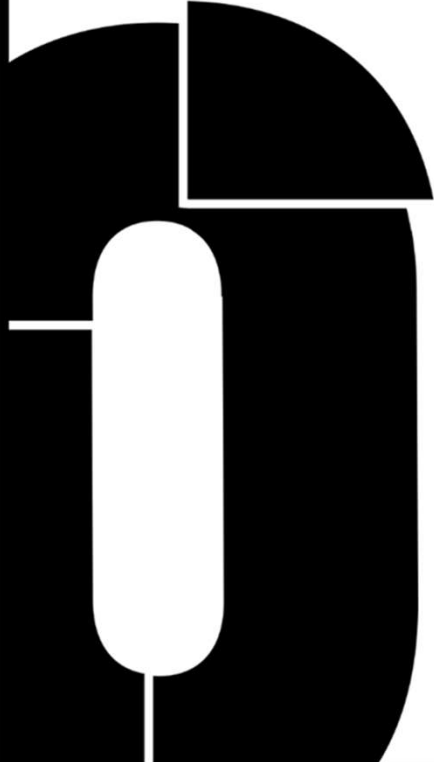


WHAT WILL WE BE TALKING ABOUT

- Open science & Open data
- FAIR principles
- Research data & methods
- Research Data Lifecycle
- Research Data Management Planning
- Various DMPs
- Chapters from CESSDA DMP checklist

INTRODUCING EMA →

Open Science Game: Open Up Your Research



OPEN UP YOUR RESEARCH

With this game, you follow Emma on her way to her PhD and decide for her to either practice science the traditional way or to follow a more open approach. While this game is intended to make researchers aware of the Open Science practices that could be applied in one's research workflow, not all of these practices might be equally suitable for all disciplines. What is more, it is not always easy to decide which parts of the research workflow should be open as there are many other factors at play that influence one's decision, such as funder requirements. Nevertheless, the game will give you an (albeit sometimes simplified) overview of the kind of open science practices that exist.

START

<https://www.openscience.uzh.ch/en/moreopenscience/game.html>

Open Science Definition

Open Science is the practice of science in such a way that others can **collaborate and contribute**, where research data, lab notes and other research processes are **freely available**, under **terms that enable reuse, redistribution and reproduction** of the **research** and its underlying **data** and **methods**.

(FOSTER Open Science)

Open Data Definition

Open Data are **online, free of cost, accessible** data that can be used, reused and distributed provided that the data source is attributed.

(FOSTER Open Science)

<https://www.fosteropenscience.eu/taxonomy/term/6>



Open science & Open data

Career benefits

- Data publication may lead to **increased visibility, reuse and citation** and therefore **recognition of scholarly work**.
- Be aware that *whenever you use the published data you are obliged to cite them*. For more information see the paragraph on data citation.



Open science & Open data

Scientific progress

- Benefits for the research itself (more robust), for the discipline and for science in general by **enabling new collaborations, new data uses** and establishing **links to the next** generation of researchers.



Norms

- Norms of the project, research group, and/or discipline **may determine whether a researcher is prone** to publish his/her data. Overall, the **openness of research data is at the heart of scientific ethics...**



Open science & Open data

External drivers

Funders

- Some funders consider costs related to data archiving and publication eligible and require a DMP.



Open science & Open data

External drivers

Publishers

- Scientific journals are increasingly adopting data availability policies that **advise or even request authors of manuscripts to make the research data, on which a manuscript is based, available.**



Horizon Europe mandate for DMP

Proper Research Data Management (RDM) is mandatory for any Horizon Europe project generating or reusing research data. It is a key part of Horizon Europe's open science requirements.

In Horizon Europe, *beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles, and should at least do the following:*

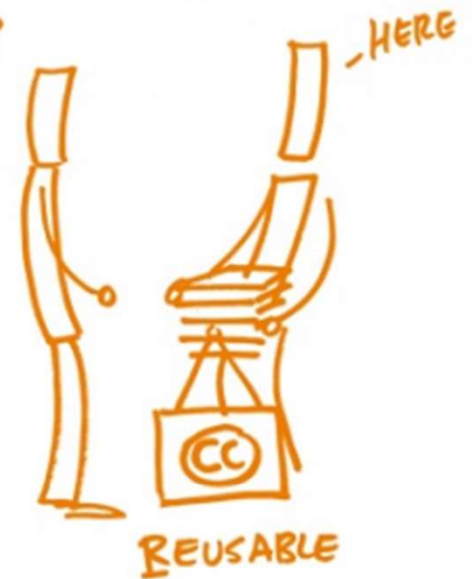
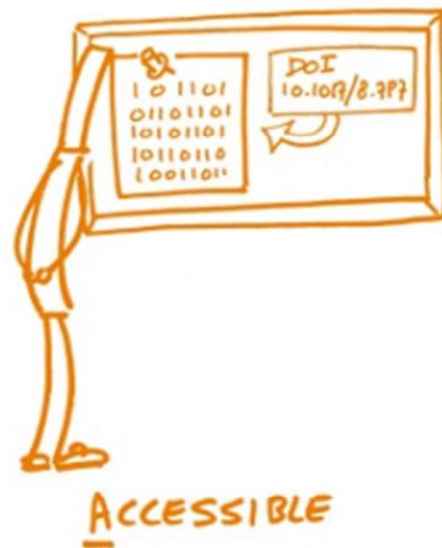
- Prepare a Data Management Plan (DMP) and keep it updated throughout the course of the project
- Deposit data in a trusted repository and provide open access to it ('as open as possible, as closed as necessary')
- Provide information (via the same repository) about any research output or any other tools and instruments needed to re-use or validate the data

Keep in mind that 'research data' is a very broad concept and certainly not limited to numerical/tabular data.



Open Science Training Handbook

FAIR DATA PRINCIPLES



FAIR principles



F → **FINDABLE**

It should be *easy to find the data and the metadata* for both humans and computers. Automatic and reliable discovery of datasets and services depends on machine-readable persistent identifiers (PIDs) and metadata.

FAIR principles



A → **Accessible**

The (meta)data should be *retrievable by their identifier using a standardized and open communications protocol*, possibly including authentication and authorisation. Also, metadata should be available even when the data are no longer available.

FAIR principles



I → **Interoperable**

The data should be able to be combined with and used with other data or tools. ***The format of the data should therefore be open and interpretable for various tools,*** including other data records. The concept of interoperability applies both at the data and metadata level. For instance, the (meta)data should use vocabularies that follow FAIR principles.

FAIR principles



R → **Re-usable**

Ultimately, FAIR aims at optimizing the reuse of data. To achieve this, *metadata and data should be well-described so that they can be replicated and/or combined in different settings*. Also, the reuse of the (meta)data should be stated with (a) clear and accessible license(s).

Real life experience from ADP

1) I FORGOT TO ASK THE RESEARCH PARTICIPANTS FOR THEIR
CONSENT TO SHARE DATA

1) I PROMISED THE PARTICIPANTS THAT I WOULD USE THE DATA
EXCLUSIVELY FOR THIS PROJECT.

1) I NEED A DOI „ASAP“, BUT I DON'T HAVE TIME TO TRANSCRIBE ALL
50 INTERVIEWS AND HAND THEM OVER TO THE ARCHIVES.

Real life experience from ADP

Common & Challenging situations



GOOD AND TIMELY DATA
MANAGEMENT PLANNING
CAN BE A GUARANTEE OF
DATA QUALITY.

Research data & methods



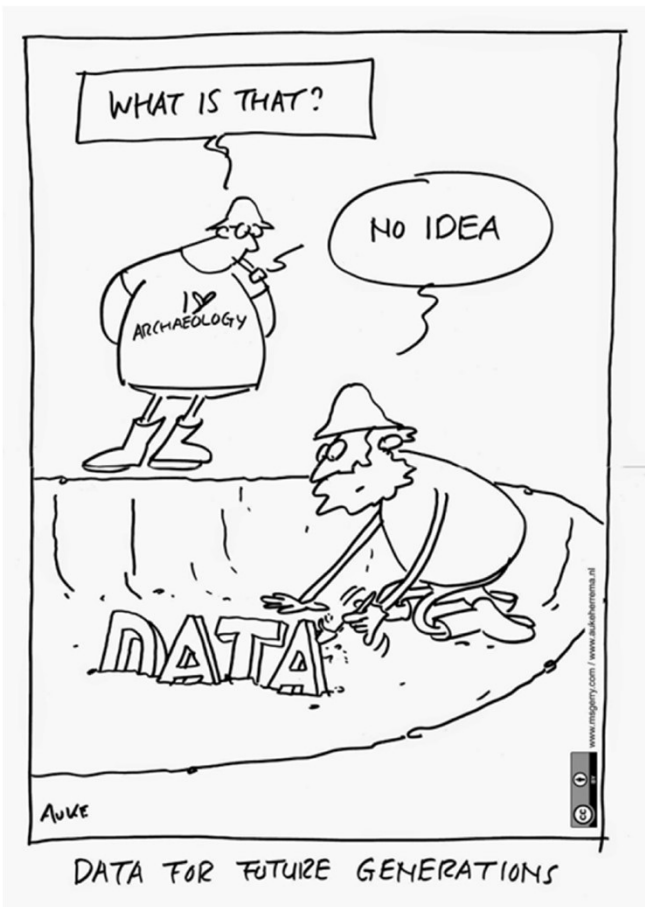
INFORMATION TYPES

Research data is ...
... primary sources that underpin scientific research and enable derivation of theoretical or applied findings.

(Preparing research data for open access :
guide for data producers, 2015)

https://www.adp.fdv.uni-lj.si/publikacije_adp/publikacija/177/

Research data & methods



The tangible forms this “material” may take are e.g. *facts, observations, interviews, recordings, measurements, experiments, simulations, and software*; numerical, descriptive and visual; raw, cleaned up and processed (Van Berchum & Grootveld, 2017).

Research data & methods

SOCIAL SCIENCES

Methods

- Opinion polls
- *Surveys*
- Interviews
- *Mass media, social media*
- Laboratory experiments
- *Field experiments*
- Fieldwork notes
- *Demographic records*
- Census records
- *Voting records*
- Economic indicators

Sources

- ❖ Generate your own data
- ❖ Obtain it from other researchers
- ❖ Data repositories
- ❖ Existing records

Research data & methods

ARTS & HUMANITIES

Methods

- Newspapers
- *Photographs, video material*
- Letters
- *Diaries*
- Literature: books, articles
- *Church records*
- Court records
- *Maps*
- Art artefacts
- *Historic artefacts*

Sources

- ❖ Libraries
- ❖ Archives
- ❖ Museums
- Public/corporate/government records
- ❖ Data repositories

CESSDA Vocabulary Service

CVs search > ModeOfCollection v.4.0.3



DDI Alliance Controlled Vocabulary for Mode Of Collection

CV name Mode Of Collection
CV short name ModeOfCollection
CV definition The procedure, technique, or mode of inquiry used to attain the data.
CV notes This vocabulary was first published by the DDI Alliance. Please see: <https://ddialliance.org/controlled-vocabularies/all>.
Language English (en) Version 4.0.3

Date of publication

EN DA DE FI FR IT

Details Versions Identity and general Usage License and Citation Export/Download

Code value	Code descriptive term (en)	Code definition(en)
▼Interview	Interview	A pre-planned communication between two (or more) people - the interviewer(s) and the interviewee(s) - in which information is action is part of the method, use "Focus group".
▼Interview.FaceToFace	Face-to-face interview	Data collection method in which a live interviewer conducts a personal interview, presenting questions and entering the responses. CAPI/PAPI or not.
Interview.FaceToFace.CAPIorCAMI	Face-to-face interview: Computer-assisted (CAPI/CAMI)	Computer-assisted personal interviewing (CAPI), or computer-assisted mobile interviewing (CAMI). Data collection method in which screen of a computer, laptop, or a mobile device like tablet or smartphone, and enters the answers in the same device. The adminis program/application.
Interview.FaceToFace.PAPI	Face-to-face interview: Paper-and-pencil (PAPI)	Paper-and-pencil interviewing (PAPI). The interviewer uses a traditional paper questionnaire to read the questions and enter the an
▼Interview.Telephone	Telephone interview	Interview administered on the telephone. Use this broader term if not CATI, or if not known whether CATI or not.
Interview.Telephone.CATI	Telephone interview: Computer-assisted (CATI)	Computer-assisted telephone interviewing (CATI). The interviewer asks questions as directed by a computer, responses are keyed d managed by a specifically designed program.
Interview.Email	E-mail interview	Interviews conducted via e-mail, usually consisting of several e-mail messages that allow the discussion to continue beyond the firs
Interview.WebBased	Web-based interview	An interview conducted via the Internet. For example, interviews conducted within online forums or using web-based audio-visual t communicate in real time.

<https://vocabularies.cessda.eu/>

COORDINATE

What data repositories usually want to know



- **Types** (qualitative, quantitative)
- **Formats** (*.rtf, *.doc, *.txt, html, *.raw, *.png, etc.)
- **Size** (big data, small data ...)
- **Sensitive data** (human participants, species of plants or animals, commercially sensitive data, state secret)
- **Long term / Short term value**

Recommended formats in ADP

Type of Materials	Recommended Formats	Other Formats
Structured text files (Study Description Form, Questionnaire, Codebook etc.)	Structured metadata description of the questionnaire (*.xml), according to the DDI or CAI software (*.bmi) *.rtf or other textual format (*.doc, *.txt, etc.)	Printed version of the material *.pdf or other graphic format
Structured numeric data (Data file)	SPSS (*.por, *.sav) ASCI (*.txt metric or a data file, equipped with labels + computer-readable description of the data file with the names and categories of variables)	Other statistical packages (e.g. STATA, R, Microsoft Excel) Tables (*.xls etc.) Databases
Freely formulated textual materials for tracking original documentation (Questionnaire, Instructions for Interviewers, Address to Respondents, copies of research reports)	*.pdf or another graphic format + printed version	*.rtf or other textual format (*.doc, *.txt, etc.)

Recommended formats in ADP

... and there are some more

Textual data

Rich Text Format (.rtf)

plain text, ASCII (.txt)

eXtensible Mark-up Language (.xml) text according to an appropriate Document Type Definition (DTD) or schema

Hypertext Mark-up Language (.html)

Common formats: MS Word (.doc/.docx)

OpenDocument Text (.odt)

Software specific formats: NUD*IST, Nvivo, ATLAS.ti in MAXQDA

Still image

TIFF 6.0 uncompressed (*.tif),

PEG (*.jpeg, *.jpg, *.jp2),

GIF (*.gif),

TIFF other versions (*.tif, *.tiff),

RAW image format (*.raw),

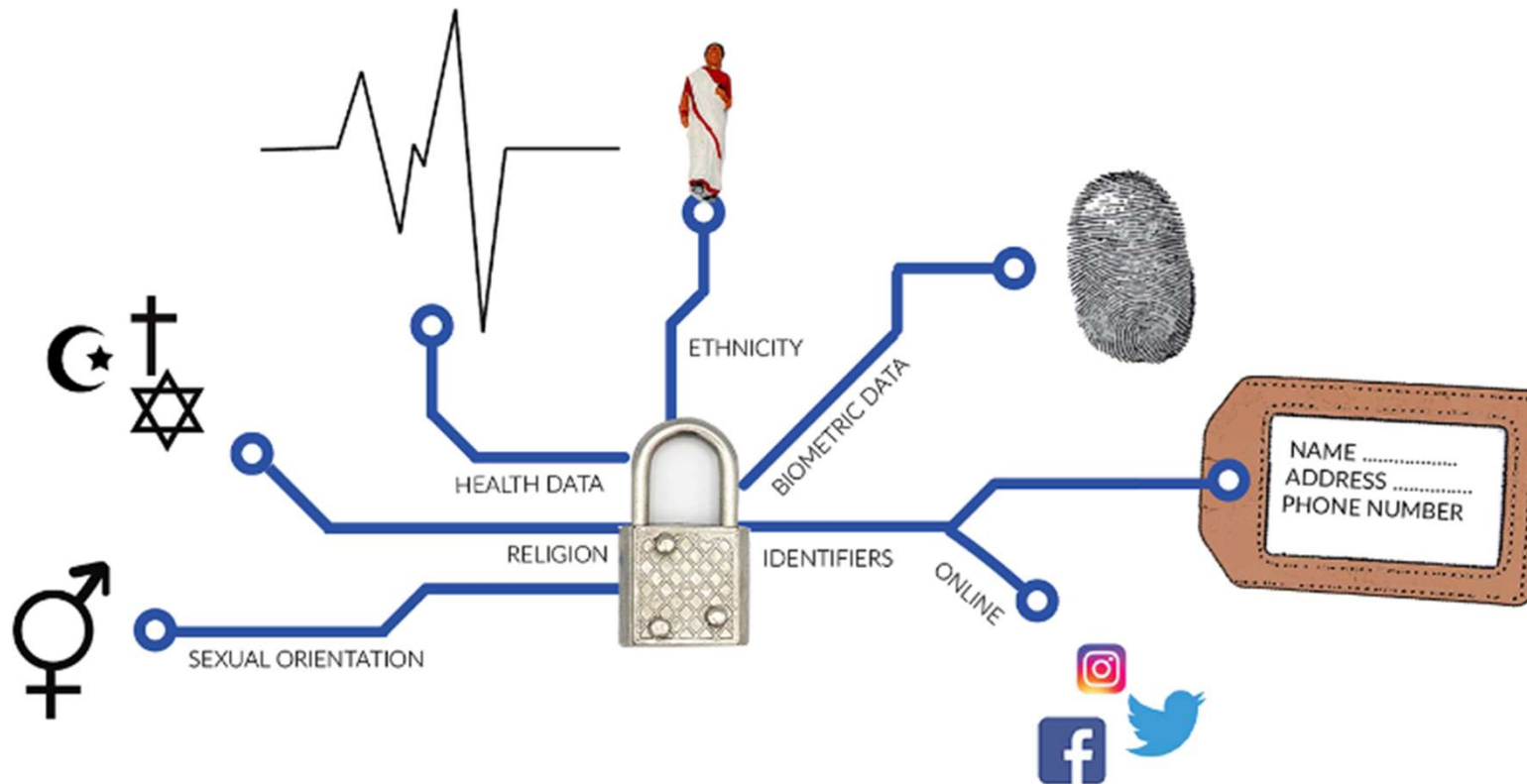
Photoshop files (*.psd),

BMP (*.bmp),

PNG (*.png),

Adobe Portable Document Format (PDF/A, PDF) (*.pdf)

Sensitive data



Other materials needed

Types of research materials

- ★ *Materials of the study*
- ★ Research results
- ★ *Related publications*

Types of research materials



Materials of the study

- Questionnaire
- Codebook
- Data processing program (syntax)
- Instructions for interviews
- Information for respondents
- Informed consent form
- ...

Types of research materials

Research results

- Research report
- Data summary
- Variable list
- Methodological information



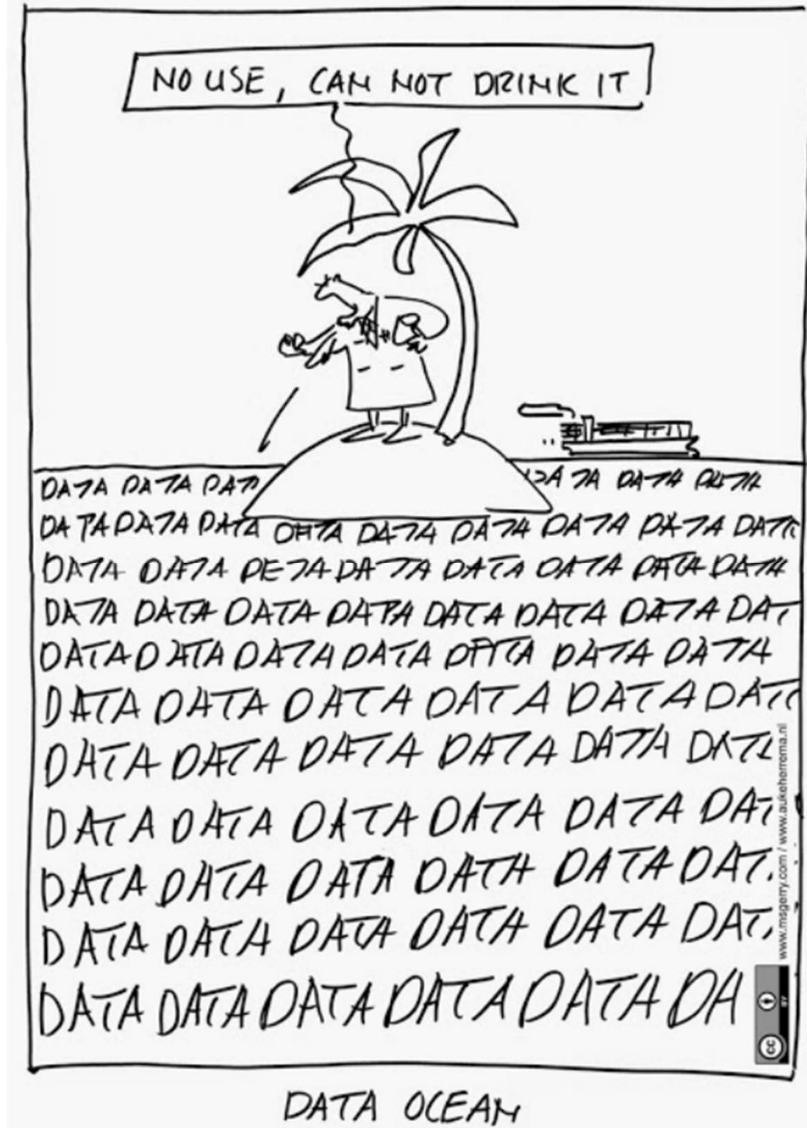
Types of research materials

Related publications

- Project's webpage
- Reports
- Scientific publications
- Related studies



Where to start?



Research Data Lifecycle



„The research data lifecycle is a model that illustrates the stages of data management and describes how data flow through a research project from start to finish.“

(Princeton Research Data Service, <https://researchdata.princeton.edu/research-lifecycle-guide/research-lifecycle-guide>)

Research Data Management Planning

... refers to how you *handle, organise, and structure* your research data throughout the research process.

... addresses also your plans for the data *after* the research is complete.

- It is a **"living" document** that **changes together with the needs** of a project and its participants.
- **It is updated** throughout the project to make sure that it tracks such changes over time and that **it reflects the current state** of your project.
- A lot of **diversity exists in DMPs** because they are always built around the particular needs of the data collected within your project.

Various DMPs



Disciplinary specific:

- DMP for social sciences developed by CESSDA
 - ❖ You can [view and download the checklist as pdf](#) (CESSDA, 2019a) or [editable form](#) (CESSDA, 2019b)

(https://static-archive.cessda.eu/content/download/4302/48656/file/TTT_DO_DMPExpertGuide_v1.3.pdf,

https://www.cessda.eu/content/download/4304/48666/file/TTT_DO_DMPExpertGuideEditVersion_v1.3.docx)

Institutional:

- DMP for PhD students at University of Ljubljana
- DMP for researchers at the Faculty of Social Sciences, UL

DOCTORAL SCHOOL
University of Ljubljana

General:

- **RDM Guidance for Researchers**
 - ❖ [Template for Data Management Plans](#)
 - ❖ [Guiding the Selection of a Trustworthy Repository](#)
- **RDM Guidance for Reviewers**
 - ❖ [Template for a Data Management Plan Evaluation Rubric](#)
(<https://scienceeurope.org/our-priorities/research-data/research-data-management/>)



Research Data Management Planning



Goal of Data Management Planning



PUBLICATIONS AND DATA

Data Publication should be considered as a **first-class research output** (Knowledge Exchange, 2013).

For a dataset to »count« as a publication should be:

- Properly **documented with metadata**,
- Reviewed for **quality**;
- Searchable and discoverable **in catalogues** (or databases);
- **Citable** in articles.

CESSDA DMP Checklist



Data Management Expert Guide

<https://dmeg.CESSDA.eu/>

Adapt your DMP

European diversity

Expert tips

Tour operators

As the data management plan (DMP) is an important tool to structure the research data management of your project, it plays a central role in this guide. Each chapter ends with a section with questions that are generally to be answered in a DMP. In the chapter's paragraphs you will be presented with the information you need to answer the proposed questions.



We have designed a list of DMP-questions especially for this Data Management Expert Guide. You can [view and download the checklist as pdf](#) (CESSDA, 2018a) or [editable form](#) (CESSDA, 2018b), and keep them as a reference while you are studying the contents of this guide.

Further readings and relevant sources

- CESSDA Training Team (2017 - 2022). *CESSDA Data Management Expert Guide*. Bergen, Norway: CESSDA ERIC. Retrieved from <https://dmeg.CESSDA.eu/>
- List of national data service providers at CESSDA: <https://www.CESSDA.eu/About/Consortium-and-Partners/List-of-Service-Providers>
- European Language Social Science Thesaurus (ELSST): <https://www.CESSDA.eu/Tools/ELSST-Thesaurus>
- CESSDA Vocabulary Service: <https://vocabularies.CESSDA.eu/>
- [Social media and research: 10 legal and ethical issues to consider](#), [SERISS](#) project
- [Using Administrative Data for Research: 10 legal and ethical issues to consider](#), [SERISS](#) project
- Data management costing tool and checklist, Created by UK Data Archive, UK Data Service: <https://ukdataservice.ac.uk/app/uploads/costingtool.pdf>