



ADP - SOCIAL SCIENCE DATA ARCHIVES

Analyze data! Deposit study! Promote science!

Research Data Management following FAIR principles

Bezjak Sonja & Vipavc Brvar Irena

Doctoral Seminar II, Faculty of Social Sciences, 30. 3. 2022



Schedule

17.00-18.15 Lecture

About ADP

Introduction to DMP

- what is data
- FAIR principles
- what is RD life-cycle
- research data management
- DMEG chapters (interactive)

18.15-18.30 Break

18.30-19.30 Discussion & Wrap up

Aims of this lecture

Context:

Doctoral School University of Ljubljana

Rules and Regulations for Doctoral Studies at the University of Ljubljana (PDF) in force from 1 October 2021

(Official Rules can be found here), https://www.uni-lj.si/doctoral_school/rules/

- 1) the PhD students understand the following concepts of open science: "open data", "FAIR principles", "research data lifecycle", "research data management", "data publication", "data citation"
- 2) students are able to prepare RDMP adopted for social sciences

Slovenian Social Science Data Archives (*ADP-Arhiv družboslovnih podatkov*)



- Founded in 1997
- Slovenian **national research data centre** for social sciences
- **Member of CESSDA** ERIC since 2017
- Status of a **trust-worthy archive** (CoreTrustSeal since 2018)
- involved in EU and national projects



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...)

Slovenian national research data centre *for social sciences*



ADP - SOCIAL SCIENCE DATA ARCHIVES

Analyze data! Deposit study! Promote science!



Subscribe to eNews

eNews Blog Contact FAQ SLO

Home | USE DATA | DEPOSIT STUDY | LEARN ABOUT | DISCOVER ADP | SEARCH

HOW TO GET DATA?



FIND →



REGISTER



→ ANALYZE

HOW TO DEPOSIT DATA?



RECORD →



PREPARE



→ DEPOSIT

BLOG

NEWS

10. February 2022

Nekaj praktičnih napotkov pri analizi obsežnejših zbirk kvalitativnih podatkov

12. January 2022

Uporaba fotoaparata in drugih aplikacij na pametnih telefonih na potovanjih za ohranjanje kulturne dediščine

20. December 2021

Odnos do cepljenja proti SARS-CoV-2: dve vzporedni anketi

26. November 2021

Spletni seminar CESSDA Roadshow: Rak in kronične bolezni

24. November 2021

Družboslovni podatki o Covid-19 v obliki tematske zbirke

More »

EXPOSED



CURRENT PROJECTS



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

QUICK FACTS

- **600 social science studies** research data accessible in a data catalogue + 150 metadata only
- **1000 users registered per year** (90 % education, 10 % scientific/research purpose)
- **100-200 units of research data reused** for detailed secondary-analysis per year



CESSDA - Consortium of European Social Science Data Archives

*“ Member countries seek to **increase the scientific excellence and efficacy of European research in the social sciences**, as well as to expand easy access to data and metadata regardless of borders. They want to provide a research infrastructure for their researchers, and join forces among their (national) data service providers.”*

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](#).

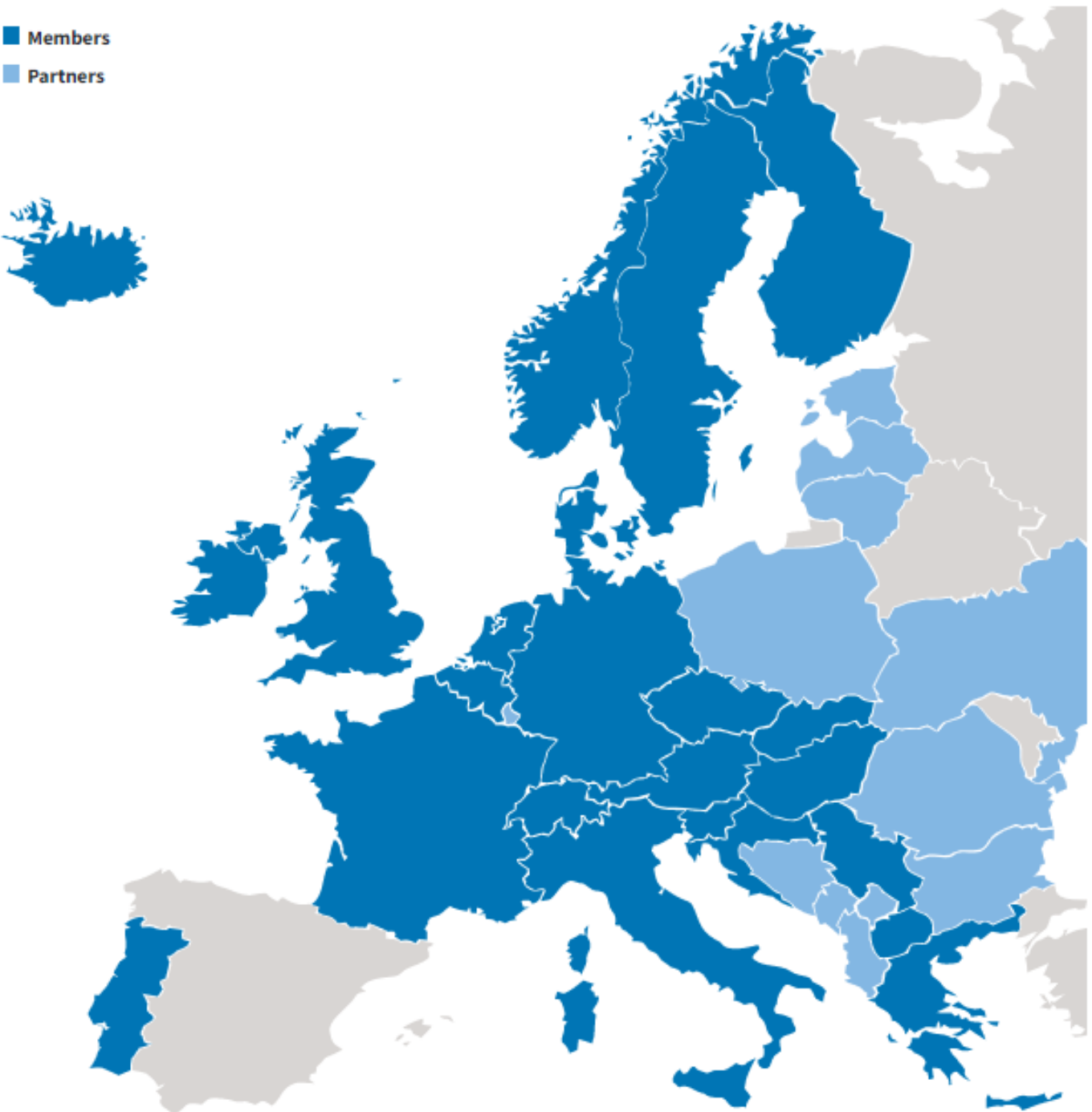


<https://www.cessda.eu/About/Mission-Vision>

CESSDA - Consortium of European Social Science Data Archives



■ Members
■ Partners



What is research data...



... **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)

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).



INFORMATION TYPES



Rich diversity of research data...

- Types
- Formats
- Size
- Sensitive data (personal data, state secret)
- Long term / Short term value



Research data in social sciences

Type of data

Quantitative data

Qualitative data



General description

In quantitative research, the gathered information is in numerical form. Quantitative research is used to quantify behaviour, attitudes or opinions. The goal of quantitative research is often to test ideas stated at the start of the research, to formulate facts and uncover patterns.

Qualitative research is primarily exploratory research. It gathers information that is not in numerical form. The goal of qualitative research is often to develop (new) ideas and a deeper understanding not achievable by numerical scores.

FAIR principles

The ultimate goal of **FAIR** is to optimise the reuse of data.



Findable

To aid automatic discovery of relevant datasets, (meta)data should be easy to find by both humans and machines and be assigned a persistent identifier.

Accessible

Limitations on the use of data, and protocols for querying or copying data are made explicit for both humans and machines.

Interoperable

(Meta)data should use standardised terms (controlled vocabularies), have references to other (meta)data and be machine actionable.

Reusable

(Meta)data are sufficiently well described for both humans and computers to be able to understand them and have a clear and accessible data usage license.

Research data management

... 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.

- A good data management strategy takes into account technical, *organisational, structural, legal, ethical and sustainability aspects*.
- Makes your research *time-efficient*, reproducible and safe as possible, if your data management is well thought through, *structured*, and *documented*.



Research data management



<https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/1.-Plan/Benefits-of-data-management>

Research Data Management

- **Increasingly demanded by** research funders, universities and other academic organisations in the context of Open Science
- *Many researchers view it as yet **another bureaucratic hurdle***
- Varying requirements by different organisations drive researchers into **despair**

Recommendation: stick to the disciplinary/field standards

Open Science will become the modus operandi of Horizon Europe. It will go beyond the open access policy of Horizon 2020 and require open access to publications, data, and to research data management plans.



2021-2027



Reinforce openness

How to write a DMP

Guide developed by CESSDA Archives

Training / Training Resources / Data Management Expert Guide



Data Management Expert Guide

This guide is designed by European experts to help social science researchers make their research data Findable, Accessible, Interoperable and Reusable (**FAIR**).

You will be guided by different European experts who are - on a daily basis - busy ensuring long-term access to valuable social science datasets, available for discovery and reuse at one of the [CESSDA social science data archives](#).

**Self-study for researchers
(15 hours of online content)**

www.cessda.eu/DMEG

DMP through chapters



CESSDA
Consortium of European
Social Science Data Archives

Adapt your Data Management Plan

A list of Data Management Questions based on the Expert Tour Guide on Data Management



This CESSDA list of Data Management Questions (2017) is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



The CESSDA Expert Tour Guide on Data Management is available at <https://www.cessda.eu/DMGuide>

PLAN

Overview

Title of the project

Date of this plan

Description of the project

- What is the nature of the project?
- What is the research question?
- What is the project time line?

Origin of Data

- What kind of data will be used during the project?
- If you are reusing existing data: What is the scope, volume and format? How are different data sources integrated?
- If you are collecting new data can you clarify why this is necessary?

Principal researchers

- Who are the main researchers involved?
- What are their contact details?

Collaborating researchers (if applicable)

- What are their contact details and their roles in the project?

Funder (if applicable)

- If funding is granted, what is the reference number of the funding granted?

Data producer

- Which organisation has the administrative responsibility for the data?

Project data contact

- Who can be contacted about the project after it has finished?

Data owner(s)

- Which organisation(s) own(s) the data?
- If several organisations are involved, which organisation owns what data?

Roles

- Who is responsible for updating the DMP and making sure that it's followed?
- Do project participants have any specific roles?
- What is the project time line?

Costs

- Are there costs you need to consider to buy specific software or hardware?
- Are there costs you need to consider for storage and backup?
- Are potential expenses for (preparing the data for) archiving covered?

Adapt your DMP: Part 1

« Previous

Next »

Search this guide

Search

The Data Management Plan (DMP) is an important tool to structure the research data management of your project. After working on each chapter you should be able to answer part of the questions which make up a DMP.

This is the first of six 'Adapt your DMP' sections in this tour guide. When you have finished the chapter on data management planning, you can start filling in the 'Overview of your research project' section. Below you can see what elements and corresponding questions are generally included in that section. You can select appropriate questions and answer them to adapt your own DMP.

For easy reference, we have put together a list of DMP-questions for all chapters in this tour guide. You can [view and download it](#) (CESSDA, 2017) and keep it as a reference while you are studying the contents of this guide.

- + Title of the project
- + Date and version of this plan
- + Description of the project
- + Origin of the data
- + Principal and collaborating researchers
- + Funder (if applicable)
- + Data producer
- + Project data contact
- + Data owner(s)
- + Roles
- + Costs



CESSDA Training Team (2017 - 2020). *CESSDA Data Management Expert Guide*. Bergen, Norway: CESSDA ERIC. Retrieved from <https://www.cessda.eu/DMEG>



Data life cycle

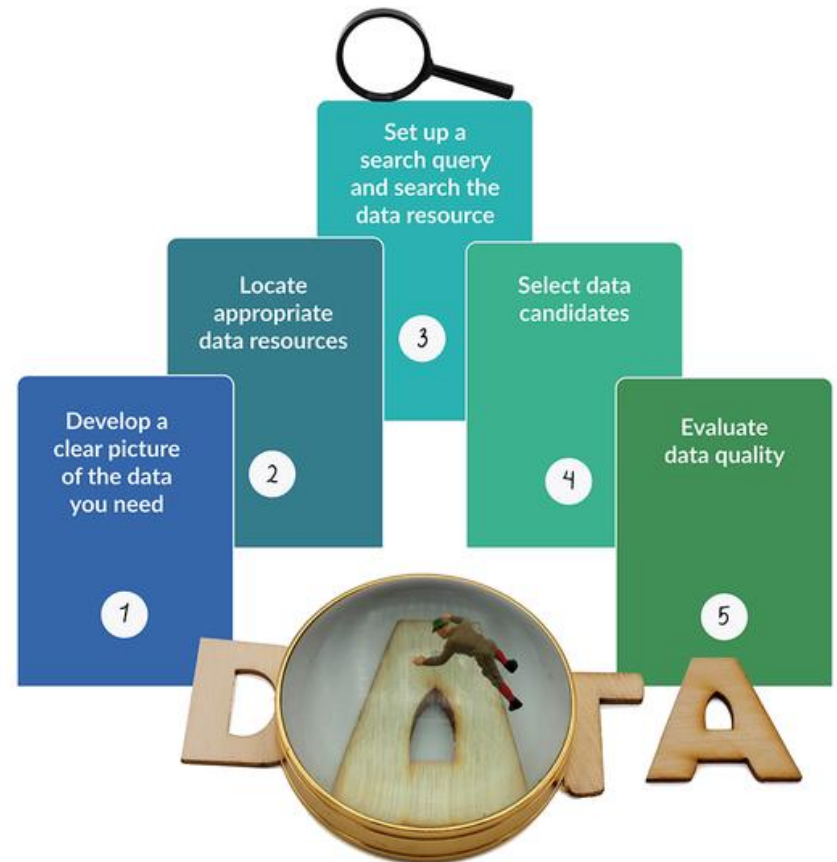


CESSDA Training Team (2017 - 2020). *CESSDA Data Management Expert Guide*. Bergen, Norway: CESSDA ERIC. Retrieved from <https://www.cessda.eu/DMGuide>

7. DISCOVER

- Why
 - ❖ Reuse data and save costs and time
 - ❖ Compare results or make replication studies
 - ❖ Reuse verified elements of research design
 - ❖ Enhance data quality and foster innovation

Steps in data discovery



[CESSDA Training Team](#) (2017 - 2020).



Discovering data

Identification of needs

- Do you plan to use existing data for your research?
 - What is the purpose for which you need the data?
 - What do you want to learn from the data?
 - What type of data do you need?
-

Search for data

- Do you know where the data may be located?
 - How do you plan to search for the data?
-

Evaluation of data quality

- What is the minimal required quality of the data (in terms of origin, contents, scope, size, methods, etc.)?
 - How do you plan to evaluate data quality (evaluation of metadata, tests, analysis, comparisons)?
-

Gaining access to data

- What are the (expected) terms and conditions for data access and use?
- What is the (expected) process for gaining access to the data?
- What is the (expected) time-span of the process for gaining access to the data?
- What are the (expected) costs for data access and use?



Overview

Title of the project/study

Date of this plan

Description of the project

- What is the nature of the project?
 - What is the research question?
 - What is the project time line?
-

Origin of Data

- What kind of data will be used during the project?
 - If you are reusing existing data: What is the scope, volume and format? How are different data sources integrated?
 - If you are collecting new data can you clarify why this is necessary?
-

Principal researchers

- Who are the main researchers involved?
 - What are their contact details?
-

Collaborating researchers (if applicable)

- What are their contact details and their roles in the project?
-

Funder (if applicable)

- If funding is granted, what is the reference number of the funding granted?
 - What is the project's title in the funding contract?
-



Overview

Data producer

- Which organisation has the administrative responsibility for the data?

Project data contact

- Who can be contacted about the project during and after it has finished?

Data owner(s)

- Which organisation(s) own(s) the data?
- If several organisations are involved, which organisation owns what data?

Roles

- Who is responsible for updating the DMP and making sure that it's followed?
- Do project participants have any specific roles?
- What is the project time line?

Costs and Resources

- Are there costs you need to consider to buy specific software or hardware?
- Are there costs you need to consider for storage and backup?
- Are potential expenses and resources for (preparing the data for) archiving covered?
- What resources will be dedicated to data management ensuring that data will be FAIR?



Organising and documenting your data

Data collection

- How will the data be collected?
 - Is specific software or hardware or staff required?
 - Who will be responsible for the data collection?
 - During which period will the data be collected?
 - Where will the data be collected?
-

Data organisation

- How will you organise your data?
 - Will the data be organised in simple files or more complex databases?
 - How will the data quality during the project be ensured?
 - If data consists of many different file types (e.g. videos, text, photos), is it possible to structure the data in a logical way?
-

Data type and size

- What type(s) of data will be collected?
 - What is the scope, quantity and format of the material?
 - After the project: What is the total amount of data collected (in MB/GB)?
-

File format

- In what format will your data be?
 - Does the format change from the original to the processed/final data?
 - Will your (final) data be available in an open format?
-

Folder structure and names

- How will you structure and name your folders?
-

File structure and names

- How will you structure and name your files?



Organising and documenting your data

Documentation

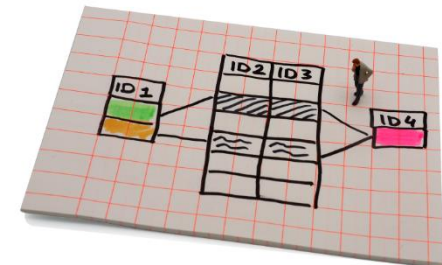
- What documentation will be created during the different phases of the project?
- How will the documentation be structured?

Metadata

- What metadata will be provided with the collected/ generated/ reused data?
- How will metadata for each object be created?
- Is there any program that can be used to document the data?
- Can metadata be added directly into the files or will the metadata be produced in another program or document?

Metadata standard (if applicable)

- What metadata standard(s) will you use?





DEPOSIT STUDY

Why Deposit?

Research Data

Management Plan

Criteria for Acquisition

Procedure of Ingest

1. Record Study

2. Sign License
Agreement

3. Prepare Data

Recommended
Formats

4. Prepare Study
Description

5. Check Submitted
Materials

6. Deposit to ADP

Publication at ADP

Deposit Data on COVID-
19

Self-archiving

After Publication

[Procedure of Ingest](#) / / [3. Prepare Data](#)

3. PREPARE DATA

Publication Year: 08. 04. 2017

Date of last inspection: 08. 04. 2017

[Data files](#)

[Quantitative data](#)

[Qualitative data](#)

It is recommended to start preparing data already in **the phase of study planning**. A good practice of preparing data for long-term storage and access includes the preparation of a **Data Management Plan**. By constantly documenting and taking into consideration good practices and standards you can expand the quality of your data.

The instructions on how to deposit data and documentation:

- **Data:** should be in a digital form, accompanied by appropriate descriptions (size of data file, format; in case of quantitative data: number of variables and units, meanings of codes – for example, the data file ASCII, together with an SPSS program for reading the data file and with the labels of questions and codes, etc.; in case of qualitative data: description of content, context, process of creation of data file etc.).
- **Questionnaire** (if used in the study): in case your study used a questionnaire to collect the data, you need to include it in the submission package in its original format.
- **Accompanying materials:** it is advisable to include all other possible materials that would help in understanding the content of the submitted data files, such as a codebook, frequency lists, instructions for interviewers, information on how the study was conducted etc.
- **Publications, based on the data:** links to the publications (for example [COBISS](#) or [URN / DOI](#)).
- All of the files should be **sent to the ADP through Cloud** or given to us on an appropriate data file exchange media (for example CD-ROM, USB stick etc.).





Processing your data

Versioning

- What is your strategy concerning versioning your data files (and scripts) during the project?
- Will you create and/or follow a convention for versioning your data?
- Who will be responsible for securing that a "Masterfile" will be maintained, documented and versioned according to the project guidelines?
- How can different versions of a data file be distinguished?

Interoperability

- Will you make use of established software and hardware? If not, how does the software and hardware you use relate to other research?

If applicable:

- Will you make use of established terminologies/ontologies (i.e. structured controlled vocabularies) in the project? If not, how do your terminologies relate to established ones?
- Which coding is used (if any)? Will you build on established coding schemes? If not, how does your coding relate to other research?

Data Quality

- How will data quality be evaluated?
- What data quality control measures will be used?



5 Service and sales workers
51 Personal service workers
511 Travel attendants, conductors and guides
512 Cooks
513 Waiters and bartenders
514 Hairdressers, beauticians and related
515 Building and housekeeping supervisors
516 Other personal services workers

File formats and data conversion

- Short-term data processing: file formats for operability
 - Proprietary vs. open formats
 - Export / portable formats
- Long-term data preservation
- Link to the table of Recommended file formats



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) ASCII (*.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.)
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

<https://www.adp.fdv.uni-lj.si/eng/deli/postopek/priprava/formati/>

STORE



Storing your data and metadata

Storage

- How and where will the (meta)data be stored during the project?
- For how long will the (meta)data be stored?

Backup

- How, where and at what intervals will the (meta)data be backed-up?
- How will data be recovered in the case of a (meta)data loss incident?

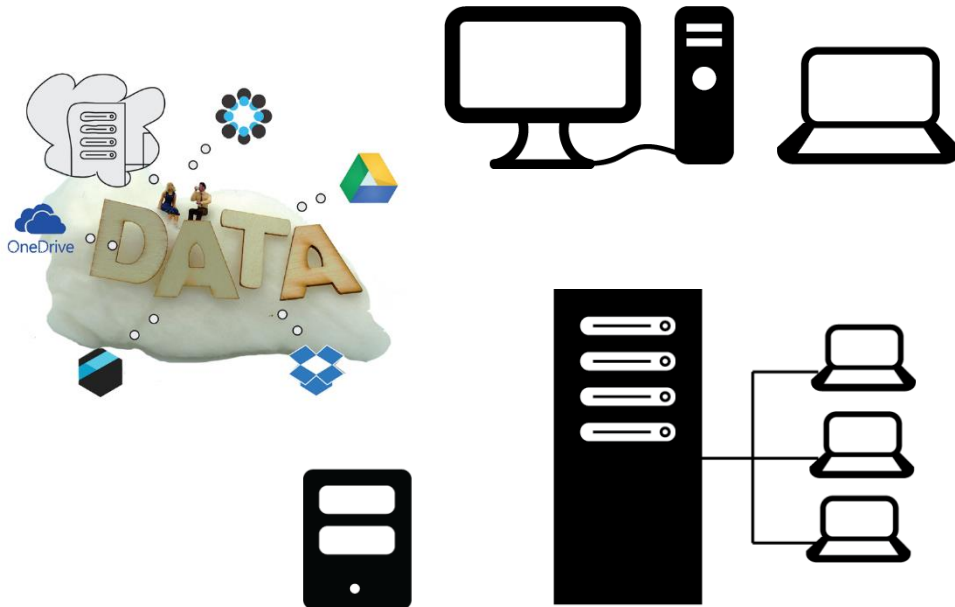
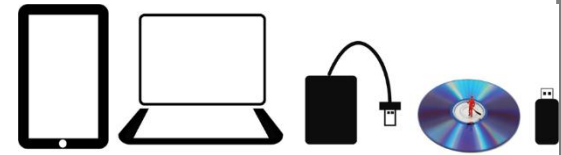
Security

- How will sensitive (meta)data be protected? (if applicable)
- How will (meta)data access be managed?



Towards a Storage Strategy

- A storage strategy contains
 - storage solutions and media
 - backup strategy and disaster recovery
 - data protection
- systematically implemented in a data management plan



Passwords

Encryption

Encryption is the process of encoding digital information in such a way that only authorised parties can view it. It's especially useful when you are transmitting personal or confidential data.

When you encrypt a file, the information it contains is “translated” to meaningless code. To translate this code back into meaningful information a key is required. Attacks with ransomware such as the [Locky virus](#) (“Locky”, 2017) have demonstrated that recovering information from encrypted files without the key is near impossible. It is therefore extremely important that you do not lose the key to decrypt your files.

Do: encrypt confidential data, especially before transmitting it online, uploading it to the cloud, or transporting it on portable devices. When working in a team, make sure that the key can be accessed by everyone who needs to access it (but only those people).

Do: ensure that you do not lose the key to decrypt your files, e.g. by keeping it in a sealed envelope in a secure location such as a safe

Encryption software

The UK Data Service (2017c) [has compiled information on encryption](#) and offers short video tutorials demonstrating the use of different software tools to encrypt data.

Commonly used encryption software includes:

- [BitLocker](#) (2017)

Standard on selected editions of Windows. For the encryption of disk volumes and USB

Research Ethics

- Disciplinary Code of Ethics (ASA)
- National Code of Ethics – Soc. Assoc.
- [European Code of Research Integrity](#)
- University ([UNI-LJ](#))
- Institute
- Funder – Horizon Europe / other EC projects / grants
- Scientific Journal <-ethical committee approval before publishing

Ethics are an **integral part of a research project**, from the conceptual stage of the research proposal to the end of a research project.



Short definition “personal data” by GDPR

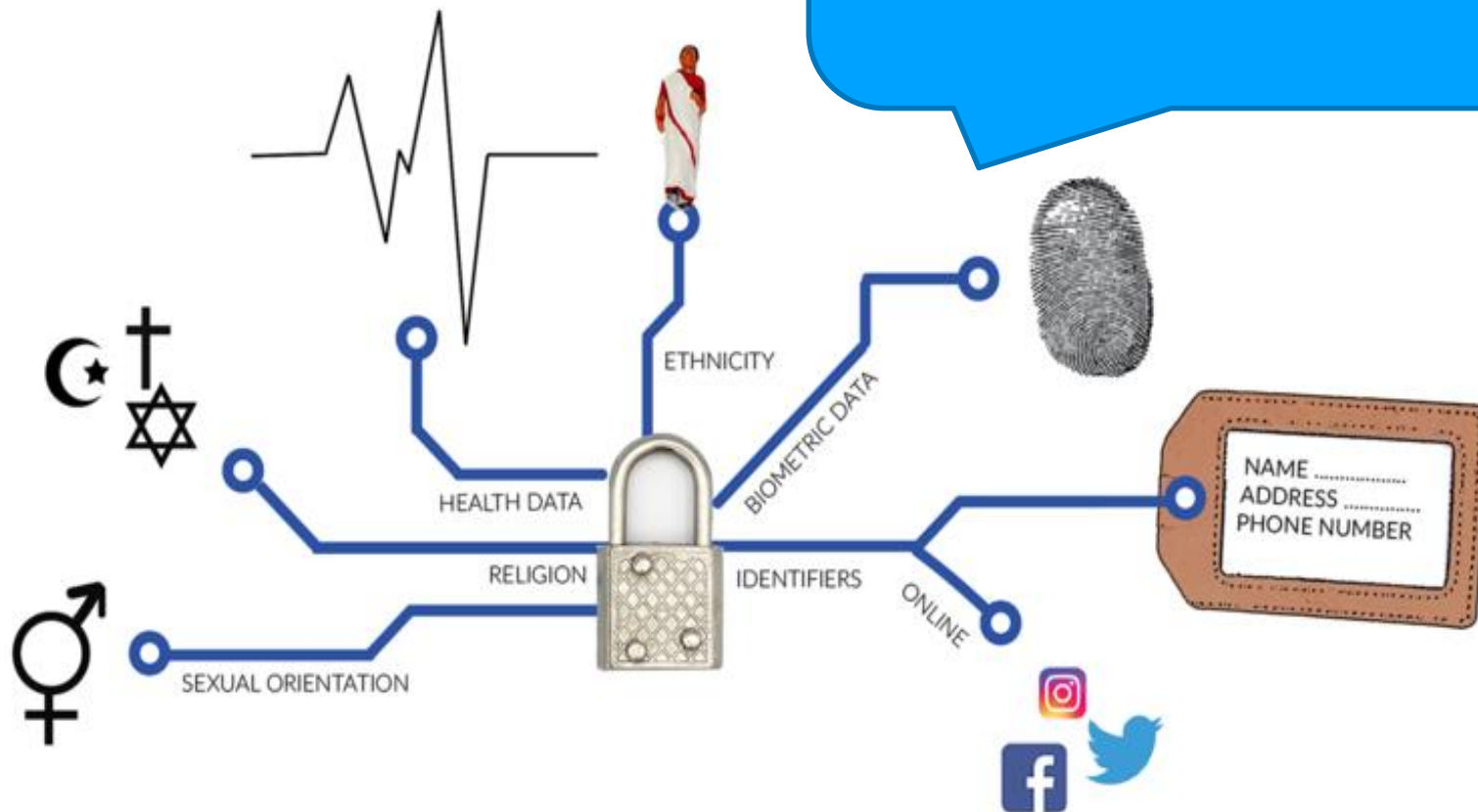
Personal data is any information that may be used to identify a person directly or indirectly

- **Directly identifying personal data**
 - *through full name, personal identification number*
- **Indirectly identifying personal data**
 - *through a combination of background information*



Sensitive personal data

... „by reference to an identifier such as a **name, an identification number, location data, an online identifier** or to one or more factors specific to the **physical, physiological, genetic, mental, economic, cultural or social identity** of that natural person“ (GDPR, Article 4)



Grounds for Processing Personal Data

The General Data Protection Regulation (GDPR)

There are 6 grounds for the processing of personal data, and one of these must be present in order to process a data subject's personal data:

1. Consent of the data subject

2. Necessary for the performance of a contract
3. Legal obligation placed upon controller
4. Necessary to protect the vital interests of the data subject
5. Carried out in the public interest or is in the exercise of official authority
6. Legitimate interest pursued by controller

Informed consent



Informed consent is the process by which a researcher discloses appropriate information about the research so that a participant may make a voluntary, informed choice to accept or refuse to cooperate.

Consent needs to be freely given, informed, unambiguous, specific and by a clear affirmative action that signifies agreement to the processing of personal data.

- ⊖ Click to see examples of consent forms
 - ⊕ UK Data Archive
 - ⊕ MRC Cognition and Brain Sciences Unit - University of Cambridge
 - ⊕ FORS (Swiss Centre of Expertise in the Social Sciences)

Strategy for Sharing Data

- Obtain **informed consent**, also for data sharing and preservation or curation
- Protect identities where needed e.g. **anonymisation** and not collecting personal data if not necessary
- **Regulate access** where needed (all or part of data) e.g. by group, use or time period
- **Securely store** and protect personal and sensitive data

DPIA - Data Protection Impact Assessment

The DPIA is a written document to be formally approved by the University and DPO.

- Sensitive data
- *Consent not possible*
- Long term processing / archiving
- *Vulnerable group*
- Very identifiable data
- *Combination of the above*

Protecting your data

Ethical review (if applicable)

- Does your project require approval by a local ethics committee?
 - How will possible ethical issues be taken into account, and codes of conduct followed?
-

Informed consent (if applicable)

- Do you require informed consent for your project?
 - If so, how will permission be obtained?
 - How are consent files organised and stored?
-

(sensitive) Personal data /confidential information (if applicable)

- How will access to (sensitive) personal data during the project be controlled?
 - How will collaborators be granted access to the data in a secure way?
 - If the research project is going to have data that includes confidential information or information that requires informed consent, is there a requirement to notify a privacy officer?
 - Is there any confidential information within the material that requires special treatment and/or limits the access to it during/after the project?
 - How will the material be protected during/after the project?
 - How will permissions and restrictions be enforced?
-

Intellectual property rights (IPR)/Copyrights

- Are there IPR or copyright issues to consider?
 - Will permission be needed to collect/reuse the data?
 - Will these rights be transferred to another organisation for data distribution and archiving?
-

Agreements (if applicable)

- What are the agreements with other stakeholders?
-

Restrictions (if applicable)

- Are there any other restrictions that need to be considered?



DATA Publication (P!!!)



PUBLICATIONS AND DATA

It is expected that a Data Publication will ensure that data will potentially 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.

Where to publish?



⊕ Journal supplementary material service

⊕ Institutional data repository

Repository of the
University of Ljubljana

⊕ General purpose repository

zenodo

⊕ Domain specific data repository

DARIAH-SI

⊕ Trusted domain specific data repository

CLARIN.SI

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES



Data publication with domain spec. repository



PUBLICATIONS AND DATA

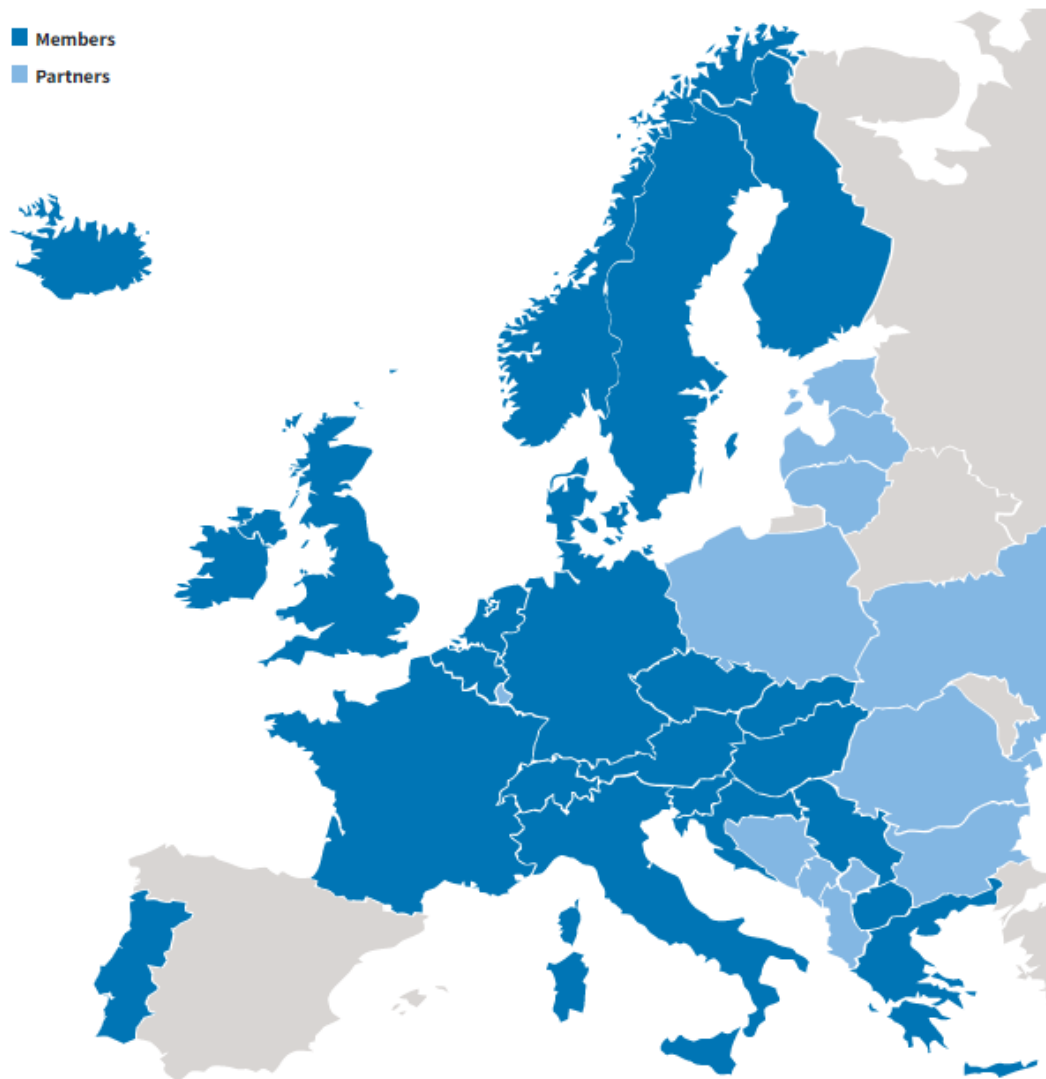
Advantages

- Offers **specialist domain knowledge** and data management expertise, e.g. to create a catalogue record and documentation;
- More likely to **accept complete datasets**;
- Provides preservation and curation to **community standards**, e.g. file formats migration;
- **Ability to control access** of (sensitive) personal data;
- May handle data re-use queries;
- May make your data visible via dissemination and promotion.

Disadvantages

- Most likely to be selective about what kind of data they accept;
- Requires advance planning of the effort needed to meet high standards for metadata and documentation.

Consortium of European Social Science Data Archives



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



Archive and publish with ADP

- 1) Slovenian national data archive for social sciences
- 2) Trust-worthy
- 3) Get credits for publication
- 4) Get advise and support from data experts
- 5) Get training
- 6) Get involved with ADP's partnering data archives



Archiving and publishing your data

Archiving

- How and where will the data be stored after the project's completion?
- Will you archive your data in a trusted data repository?
- Will the application of a persistent identifier to your data be ensured?

Data formats

- What formats will you provide your data in for archiving (and sharing)?
- Will specific software be required to process your data? Can this software be deposited with the data?

Access (if applicable)

- Will your data be available (Open Access)?
- Will all data or only parts of it be published?
- What licenses do you need for your data?
- How should your data be cited when reused?
- Will there be an embargo period for (all or some of) the data?
- Are there other agreements or restrictions (see above) that need to be considered?
- Are there any legal/ethical restrictions that prevents the publication of all the material?
- Will these restrictions mean that action must be taken before the material can be made available?
- Is there a risk of delayed publication/making data available (all or parts of)?
If so what might be needed to do to avoid this?



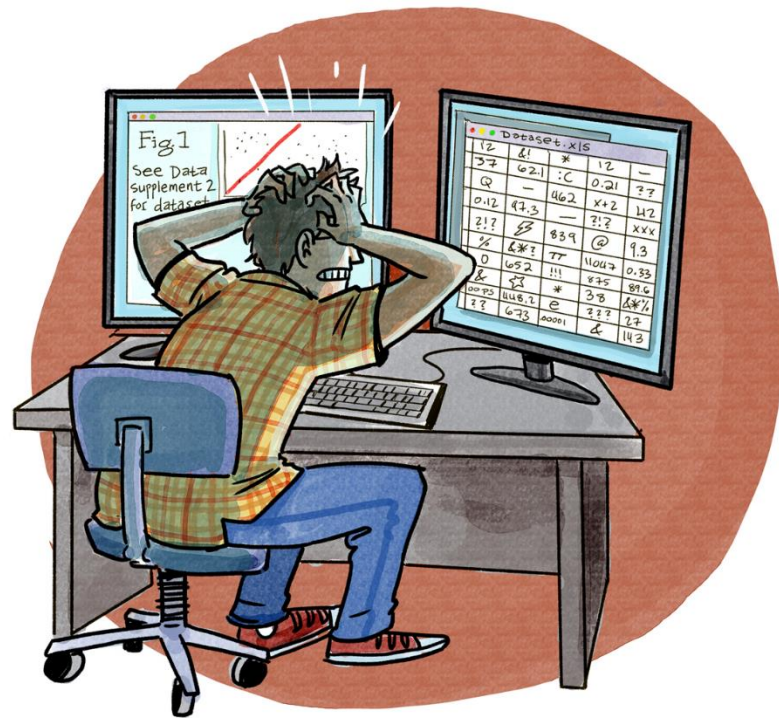
Data life cycle



CESSDA Training Team (2017 - 2020). *CESSDA Data Management Expert Guide*. Bergen, Norway: CESSDA ERIC. Retrieved from <https://www.cessda.eu/DMGuide>

Time to...

Present your DMP ☺



<https://instr.iastate.libguides.com/dmp/writingDMP>



Additionally

CESSDA Quiz: www.cessda.eu/dmeg

Take the quiz below and find out which chapters of DMEG will be most useful for you.

Take the DMEG quiz

LoD Data Management Challenge game: <https://lod.sshopencloud.eu>



IF YOU HAVE ANY FURTHER QUESTIONS...



... CONTACT ADP

University of Ljubljana
Faculty of Social Sciences
Social Science Data Archive
Kardeljeva ploščad 5
1000 Ljubljana
Slovenia

 www.adp.fdv.uni-lj.si

 arhiv.podatkov@fdv.uni-lj.si

 [Arhiv.Druzboslovnih.Podatkov](https://www.facebook.com/Arhiv.Druzboslovnih.Podatkov)

 [@ArhivPodatkov](https://twitter.com/ArhivPodatkov)

QUESTIONS AND ANSWERS

- 1) PROTECT - CONSENT FORM
- 2) COSTING TOOL
(<https://ukdataservice.ac.uk/media/622368/costingtool.pdf>)
- 3) PROTECT AND STORE YOUR DATA
(<https://study.sagepub.com/corti2e/student-resources/data-collection/answers-to-in-chapter-exercises/61-book-answers>)
- 4) ANONYMIZATION OF QUALITATIVE DATA
(<https://study.sagepub.com/corti2e/student-resources/data-collection/answers-to-in-chapter-exercises/81-anonymisation-of>)
- 5) FILE FORMATS
(https://ukdataservice.ac.uk/media/622179/exercise_open_file_for_mats.pdf)
- 6) Template for transcribing interviews, with uniform style and layout
([Model qualitative interview transcript](#))

How FAIR are your data?

Findable

It should be possible for others to discover your data. Rich metadata should be available online in a searchable resource, and the data should be assigned a persistent identifier.

- A persistent identifier is assigned to your data
- There are rich metadata, describing your data
- The metadata are online in a searchable resource e.g. a catalogue or data repository
- The metadata record specifies the persistent identifier

.....

Let's see if your data is FAIR

https://www.cessda.eu/content/download/3845/35038/file/20170707_How_FAIR_are_your_data_Jones.pdf

Read the points and add notes to elements that still need to be resolved.



COSTING TOOL

<https://www.ukdataservice.ac.uk/media/622368/costingtool.pdf>

ACTIVITY	COMMENTS AND SUGGESTIONS	√	COST
<p>Data description</p> <ul style="list-style-type: none"> • Are data in a spreadsheet or database clearly marked with variable and value labels, code descriptions, missing value descriptions, etc.? • Are labels consistent? • Do textual data like interview transcripts need description of context, e.g. included as a heading page? 	<ul style="list-style-type: none"> • if data description is carried out as part of data creation, data input or data transcription – low or no additional cost • if needed to be added afterwards – higher cost • codebooks for datasets can often be easily exported from software packages 		
<p>Data cleaning</p> <ul style="list-style-type: none"> • Do quantitative data need to be cleaned, checked or verified before sharing, e.g. check validity of codes used, check for anomalous values? • Will data match documentation, e.g. same number of variables, cases, records, files? • Does textual information in data need to be spell-checked? 	<ul style="list-style-type: none"> • if carried out as part of data entry and preparation before data analysis – low or no additional cost • if needed afterwards – higher cost 		
<p>Documentation</p> <ul style="list-style-type: none"> • Do you have documentation for the data that describes the context and methodology of how data were gathered, created, processed and quality controlled? 	<ul style="list-style-type: none"> • often essential contextual and methods documentation will be written up in publications and reports • if all data creation steps are well documented and documentation is kept well organised during research – low or no additional cost 		

Why publish research data?



Data Sharing and Management Snafu in 3 Short Acts

Karen Hanson, Alisa Surkis and Karen Yacobucci (2012) NYU Health Sciences Library:
<https://www.youtube.com/watch?v=N2zK3sAtr-4>

Sources

CESSDA Training Team (2017 - 2020). *CESSDA Data Management Expert Guide*. Bergen, Norway: CESSDA ERIC. Retrieved from <https://www.cessda.eu/DMGuide>

Astell, Mathias; Admin, Springer Nature (2018): Infographic - Practical challenges for researchers in data sharing. Figshare. Journal contribution. <https://doi.org/10.6084/m9.figshare.5996786.v4>

Some slides originate from Train the Trainers package of CESSDA DMEG. We would like to thank colleagues from CESSDA to make it possible to re-use them for events like this.

Viri v slovenskem jeziku (sources in SI)

ADP: [Življenjski krog podatkov](#)

Načrt ravnanja z raziskovalnimi podatki - [vprašalnik](#)

Informacijski pooblaščenec:

- infografika – [Podlaga za obdelavo osebnih podatkov](#)
- [Ocena učinka v zvezi z varstvom osebnih podatkov](#)

[Načrtovanje zbiranja raziskovalnih podatkov skladno z načeli FAIR](#)
(predavanje za doktorske študente, 2019)

Delavnica [Ravnanje z raziskovalnimi podatki](#), marec 2020

[\(DCC: DMPOnline](#) (posnetek predstavitve uporabe, P. Čerče – ZRS Koper)
[SPLOŠNI NRRP – Obzorje Evropa](#))

Using Administrative Data for Research



SOCIAL MEDIA AND RESEARCH

