

Research Data Management

Sonja Bezjak, Irena Vipavc Brvar, ADP

***Seminar for PhD students, 5 January 2017,
Faculty of Social Sciences, University of Ljubljana, Ljubljana***

Content

1. About Social Science Data Archives (ADP)
2. Open Data
3. Research Data Life Cycle
4. Data Repositories
5. Data Management Plan
6. Data Management Planning (DMP) themes (DCC & UC3)
7. Research Data Life Cycle (steps)
8. Further readings



Ravnanje z raziskovalnimi podatki in odprti dostop

sreda, 18. 1. 2017, od 9.00 – 12.00

Kje: Centralna tehniška knjižnica Univerze v Ljubljani, čitalnica

Predavateljci: mag. Irena Vipavc Brvar in dr. Sonja Bezjak

Delavnica je namenjena raziskovalcem in študentom s področja družboslovja in humanistike.

Udeleženci se bodo seznanili z dobrimi načini ravnanja z raziskovalnimi podatki in vprašanji, povezanimi z deljenjem podatkov. Pri tem bomo posebno pozornost namenili aktualnim nacionalnim in evropskim politikam, ki spodbujajo odprto znanost.

Na delavnici bomo obravnavali kritična področja ravnanja z raziskovalnimi podatki na poti v odprti dostop: priprava načrta za ravnanja z raziskovalnimi podatki, etični in pravni vidiki deljenja podatkov (licenciranje in intelektualna lastnina, občutljivih ali zaupnih podatkov), priprava raziskovalnih gradiv za potrebe druge rabe, arhiviranje podatkov in dolgoročno skrbništvo, varnost podatkov in hramba.

Delo bo potekalo kot kombinacija predavanj, razprav in skupinskega dela.

V branje priporočamo [Priročnik za raziskovalce: Priprava raziskovalnih podatkov za odprti dostop](#)

Udeležba na delavnici je brezplačna. Potrebno se je [registrirati](#).

1 Social Science Data Archives, University of Ljubljana



Photo by: <http://www.fdv.uni-lj.si/>

- 1997
- national data repository for social sciences
- depositors from all 4 (3 public) universities, private research centres, Statistical Office of Slovenia (8-10 research centres per year)
- 600 social science surveys
- cca. 700 users yearly (90 % education, 10 % scientific/research purpose)
- ADP is the only research data repository in Slovenia

1 Data Services

Data + metadata +
accompanying material

DATA DEPOSITORS:

- Formats
- Standards
- Consent
- Licenses
- Bibliography
- Cobiss
- ...

ADP

- Selection
- Added value
- Curation
- Access
- ...

DATA USERS:

- Search
- Use
- Tools
- Citation
- ...

1 About ADP

From its start ADP follows developments in the field of data archiving

CESSDA

FDC



Involvement on national level – cooperation with libraries (National University Library, Central Technological Library at the UL, The University of Maribor Library)

2010 - 2014



Open access to the achievements of Slovenian scientists

Joint Conference of the Special Libraries Section and the Academic Libraries Section of the Slovenian Library Association

Social Science Data Archives and open data services in the Republic of Slovenia
(roles, responsibilities, trusted repositories) (15 p. article published in 2010)

Slovene conference on long-term preservation of digital material

Analysis of current situation and perspective of digital preservation in the Social Science Data Archives (ADP). (30 p. article published in 2011)



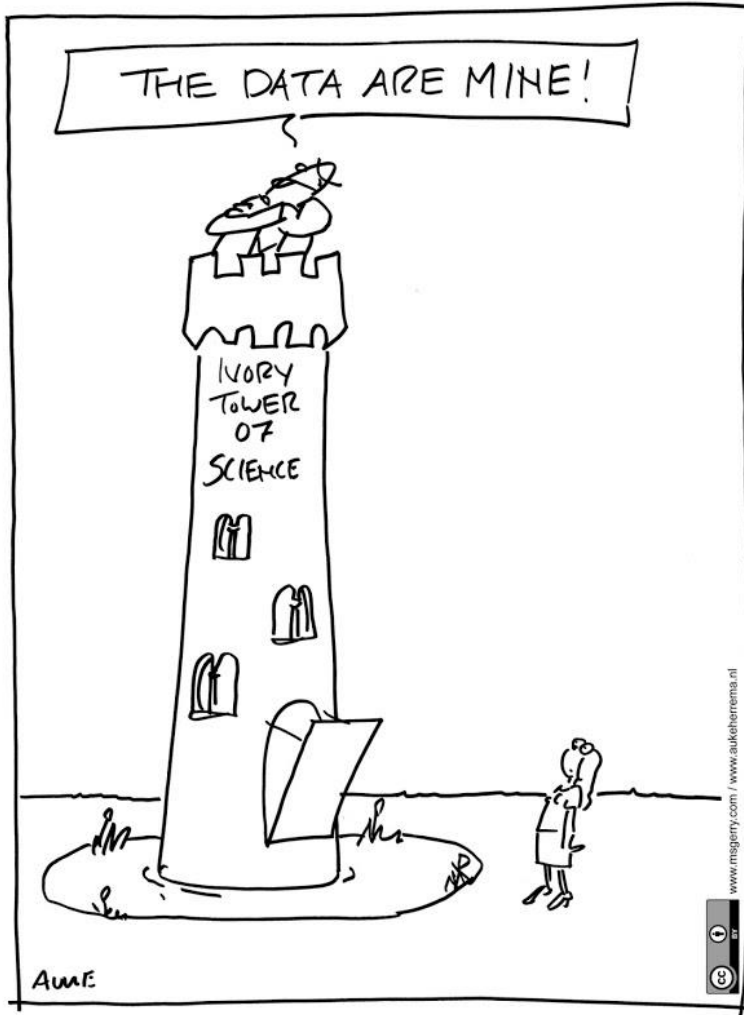
What is CESSDA?

- **Consortium of European Social Science Data Archives**
- The vision of CESSDA is to provide **full scale sustainable research infrastructures** that enables the research community to conduct high-quality research which in turn leads to effective solutions to the major challenges facing society today.

CESSDA's main tasks

- Coordinate the network of European data service providers
- Promote the results of social sciences
- **Facilitate researcher access to important resources of relevance to the European social science research agenda regardless of the location of either researcher or data**
- Include further data sources
- Provide training
- Promote wider participation in CESSDA
- Develop and coordinate standards, protocols and professional best practices

2 Open Data



SCENE FROM THE PAST ?

„Open data is data that is free to:

- access,
- reuse,
- repurpose,
- and redistribute.“

(<https://www.openaire.eu/opendatapilot>)

2 Open Data Policy



DATA SHARING

1) EU level:

→ Horizon 2020 ([Open Research Data Pilot](#))

2) National level: Slovenia

→ [National strategy of open access to scientific publications and research data in Slovenia 2015-2020](#)

2 Open Data Policy

3) Scientific Journals

- [PLOS journals Data Availability](#)
- [Oxford Journals Editor's Note](#)
- [APSA Data Access, Production Transparency, and Analytic Transparency](#)

Acceptable Data-Sharing Methods

Data deposition (strongly recommended)

All data and related metadata underlying the findings reported in a submitted manuscript should be deposited in an appropriate public repository, unless already provided as part of the submitted article. Repositories may be either subject-specific (where these exist) and accept specific types of structured data, or generalist repositories that accept multiple data types, such as [Dryad](#). Guidance on acceptable repositories is included below.

journals.plos.org/plosone/s/data-availability

[Oxford Journals](#) > [Social Sciences](#) > [Public Opinion Quarterly](#) > [Volume 69, Issue 4](#) > Pp. 507.

To permit competent professionals to confirm the results and analyses, authors are expected to retain raw data for a minimum of 5 years after publication of the research. Other information related to the research (e.g., instructions, treatment manuals, software, details of procedures) should be kept for the same period. This information is necessary if others are to attempt replication. Authors are expected to comply promptly and in a spirit of cooperation with such requests. . . Sometimes special concerns must be addressed, such as confidentiality of the participants and proprietary or other concerns of the sponsor of the research. . . Generally, the costs of complying with the request should be borne by the requester.

poq.oxfordjournals.org/content/69/4/507.full

Data Citation

Data Sets

Levy, Jack S. and T. Clifton Morgan. *Great Power Wars, 1495-1815*. [Computer file]. ICPSR09955.v1. 1989. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1994. doi:10.3886/ICPSR09955.v1

www.apsanet.org/APSR-Submission-Guidelines-August-2016



2 Sharing research data - WHY

- encourages scientific enquiry and debate
- promotes innovation and potential new data uses
- leads to new collaborations between data users and data creators
- maximises transparency and accountability
- enables scrutiny of research findings
- encourages the improvement and validation of research methods
- reduces the cost of duplicating data collection
- increases the impact and visibility of research
- promotes the research that created the data and its outcomes
- can provide a direct credit to the researcher as a research output in its own right
- provides important resources for education and training

3 Research Data Life Cycle



3 Support in RD Life Cycle

Research institution:

- Assures and manages internal open access policies and designs procedures to **help** research projects **with the preparation and implementation of DMPs**.
- Manages expert **training for researchers** and support staff.
- Provides infrastructure in the form of **technological and advisory services** to provide digital preservation for data throughout the lifecycle (research offices, libraries, CSIC, SICRIS, operating data centre and service networks).
- Provides common **services and tools** to support research groups.

Štebe, Janez et al (2015): Preparing Research Data for Open Access. Guide for Data Producers. Faculty of Social Sciences. Ljubljana. P. 7.

3 Support in RD Life Cycle

Library:

- Provides information about the **availability of existing data** sources.
- Provides information about **options to deposit data** in a data centre or data archive.
- Helps select an appropriate or recommended data centre or data archive.
- Provides information about open access conditions and advantages.
- Supports **preparation of formal DMPs**.
- Provides support with preparation of basic study metadata and documentation, author's rights, and explains other deposition requirements.

Štebe, Janez et al (2015): Preparing Research Data for Open Access. Guide for Data Producers. Faculty of Social Sciences. Ljubljana. P. 7.

3 Support in RD Life Cycle

Funder:

- Manages **national/disciplinary policies** which require the preparation of DMPs as part of the research project application process, based on the principles of open access to research data financed by public funds, perceived as being a public good.
- **Acknowledges costs** for open access data and metadata preparation and **provides funds** to cover them.
- **Monitors** implementation of open access obligations.

Štebe, Janez et al (2015): Preparing Research Data for Open Access. Guide for Data Producers. Faculty of Social Sciences. Ljubljana. P. 7.

4 Data Repositories

Depositing:

Only a (short time) back up

VS.

Long-term preservation and access of trusted and evaluated data

- Own or organizational web page
- Institutional repository (at the University)
- Journal's repository
- General repository (as [Zenodo](#))
- Disciplinary repository
- **Trusted repository**

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

4 Why ADP

1. ADP **evaluates the importance** of research data for science and their long-term usability.
2. Adopts an **acknowledged** digital curation **approach**.
3. Provides **access to data** and enables searching and browsing through standard data descriptions for the purposes of discovery.

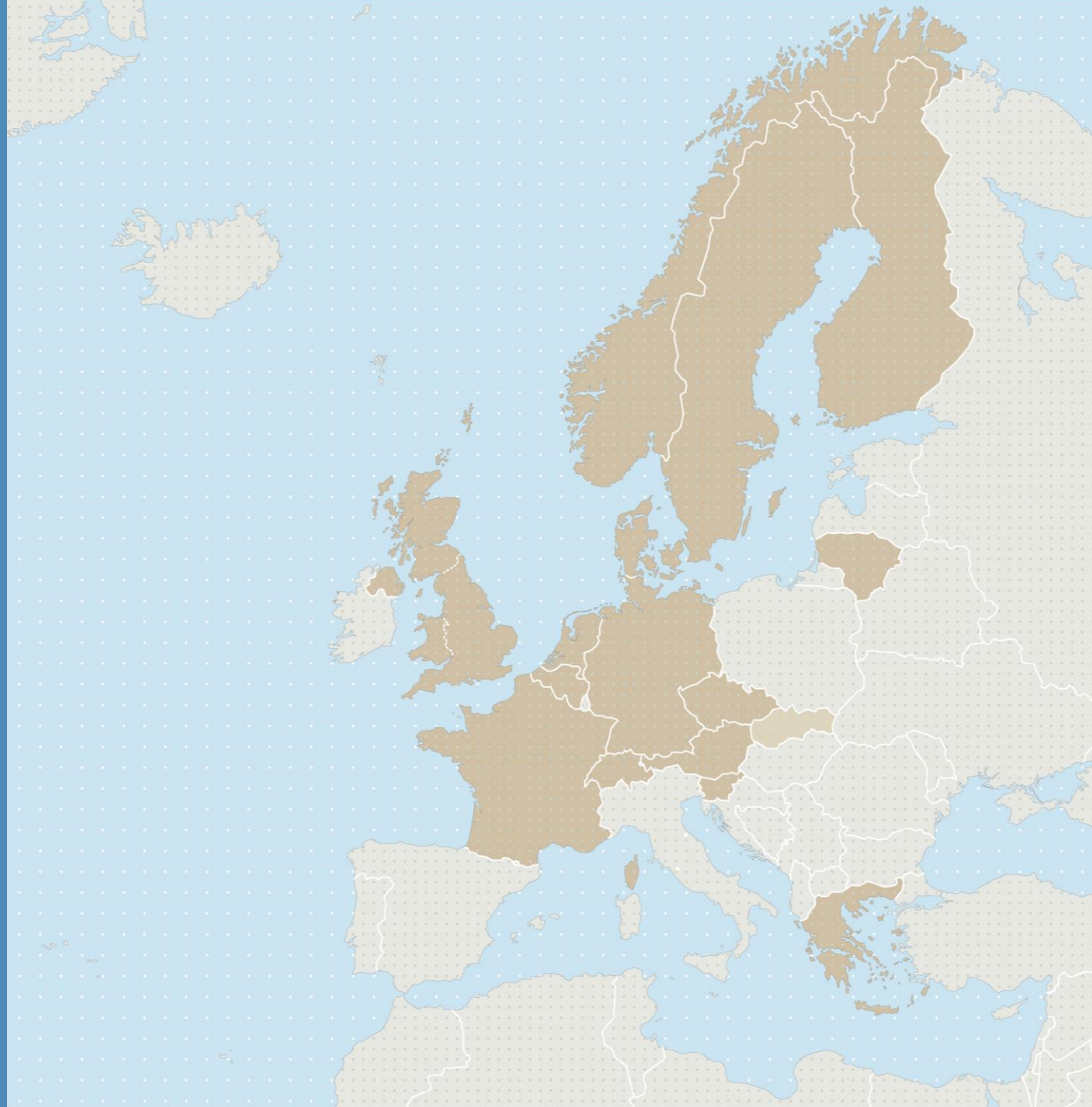
cessda



4. **Offers support** with data management planning and preparation of data for open access.

Štebe, Janez et al (2015): Preparing Research Data for Open Access. Guide for Data Producers. Faculty of Social Sciences. Ljubljana. P. 9.

- 15 CESSDA members
 - Austria
 - Belgium - SOHDA
 - Czech Republic - CSDA
 - Denmark - DDA
 - Finland - FSD
 - France - PROGEDO/Réseau Quetelet
 - Germany - GESIS
 - Greece - So.Da.Net
 - Lithuania - LiDA
 - Netherlands - DANS
 - Norway - NSD
 - Slovenia - ADP
 - Sweden - SND
 - Switzerland - FORS
 - UK - UKDS
-
- 1 observer
 - Slovakia - SASD (observer)



CESSDA Service Providers

- Operational bodies in CESSDA
- Provide services in their country and contribute to the development of the consortium
- Important instrument for dissemination of data, tools and knowledge in Europe
- Must fulfil obligations as stated in the Statutes
- CESSDA partners or linked third parties in H2020 projects aiming to strengthen and widening its services
- Link between the national and the international services

5 Data Management Plan

DMP is the process of planning, describing and communicating the activities carried out during the research lifecycle in order to...

- Keep sensitive data safe
- Maximise data's reuse potential
- Support longer-term preservation

Donnelly, M. (2014): Data Management Planning. HORIZON 2020 and the DMPONLINE TOOL. Seminar: Preparing research data for open access. Faculty of Social Sciences. Ljubljana.

5 Data Management Plan

A research data management plan describes strategies for the collection, storage, validation, security, preservation, and sharing of data where possible, throughout the data lifecycle.

(<http://cessda.net/CESSDA-Training/Research-Data-Management-Plans>)

5 Data Management Plan in H2020

General definition of DMP:

„Data Management Plans (DMPs) are a *key element* of good data management. A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. As part of making research data findable, accessible, interoperable and re-usable (FAIR)“. (Guidelines, 2016: 4)

5 Data Management Plan in H2020

A DMP should include information on:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).

(Guidelines, 2016: 4)

5 Data Management Plan in H2020: DMP 1

Your application should address the following issues:

- What standards will be applied?
- How will data be exploited and/or shared/made accessible for verification and reuse? If data cannot be made available, why?
- How will data be curated and preserved?

Your policy should also:

- reflect the current state of consortium agreements on data management
- be consistent with exploitation and Intellectual Property Rights (IPR) requirements

[Guidelines on FAIR Data Management in Horizon 2020](#), 26. 7. 2016



5 Data Management Plan in H2020: DMP 1

You should also ensure **resource** and **budgetary planning** for data management and include a deliverable for an initial DMP at month 6 at the latest into your proposal.

(Guidelines, 2016: 5)

5 Data Management Plan in H2020: DMP 2 (deliverable)

As a minimum, the DMP should be updated in the context of the periodic evaluation/assessment of the project. If there are no other periodic reviews envisaged within the grant agreement, an update needs to be made in time for the final review at the latest.

(Guidelines, 2016: 6)

5 Data Management Plan in H2020: DMP 2 (deliverable)

As a minimum, the DMP should be updated in the context of the periodic evaluation/assessment of the project. If there are no other periodic reviews envisaged within the grant agreement, an update needs to be made in time for the final review at the latest.

(Guidelines, 2016: 6)

5 Issues to cover in your Horizon 2020 DMP

1. Data summary
2. FAIR Data
 - 2.1. Making data **findable**, including provisions for metadata
 - 2.2. Making data openly **accessible**
 - 2.3. Making data **interoperable**
 - 2.4. Increase data **re-use** (through clarifying licences)
3. Allocation of resources
4. Data security
5. Ethical aspects
6. Other

(Guidelines, 2016: 10-12)

Before we continue...

What kind of data are you planning to create / collect / use?



INFORMATION TYPES

To achieve good quality...

One should follow good practice of research data management through the whole research project.

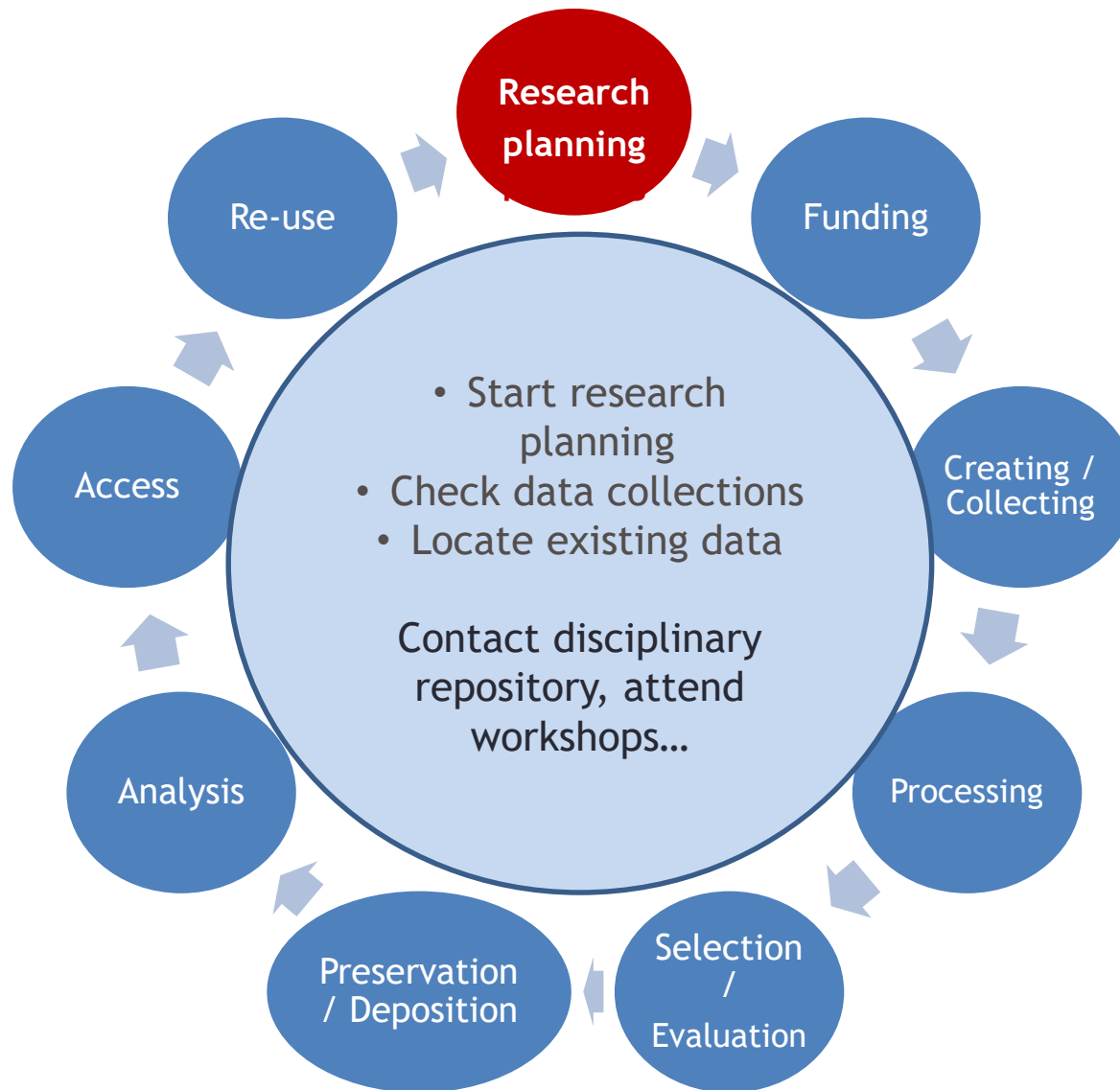
6 Data Management Planning (DMP) themes

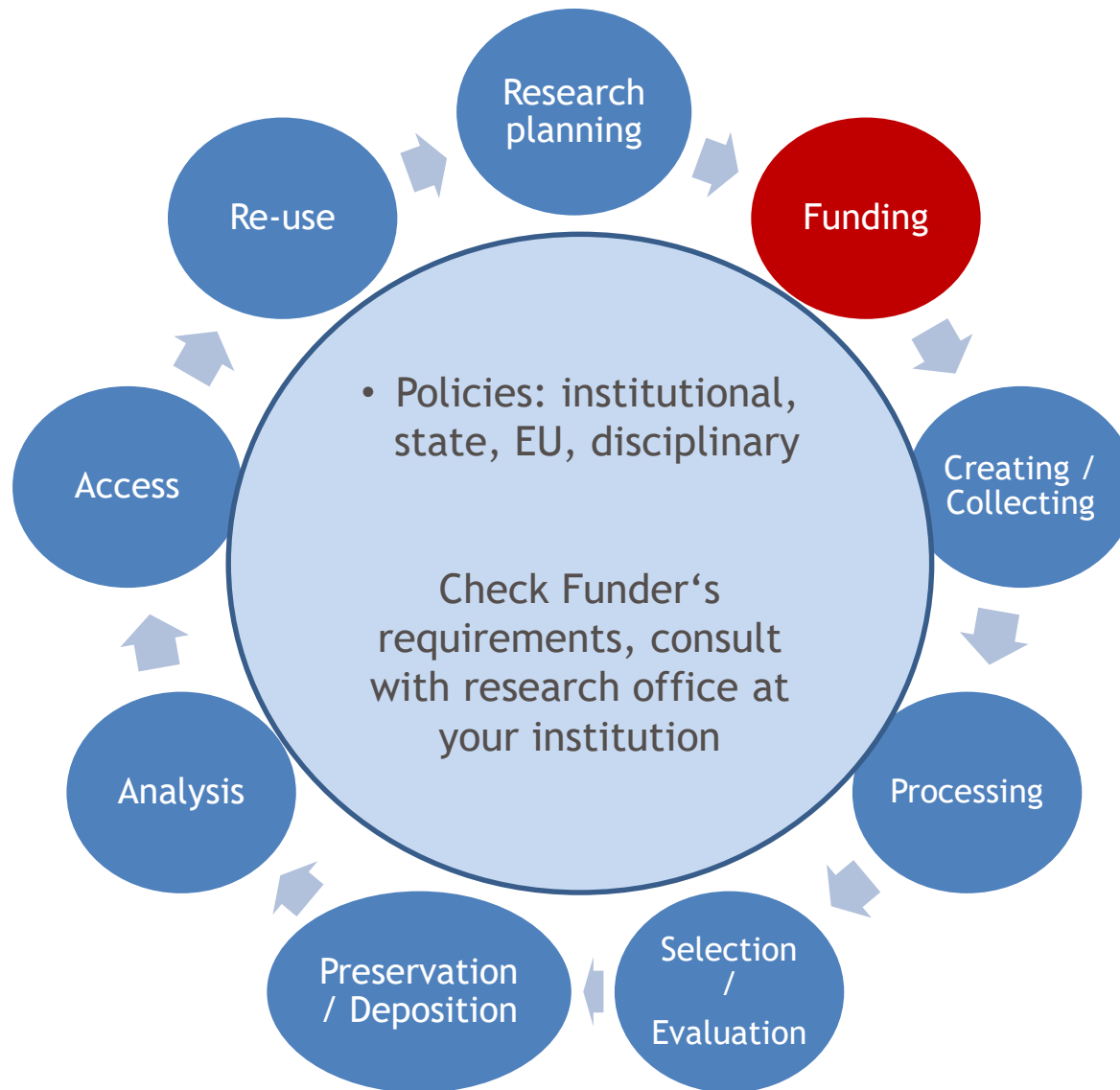


Data Management Planning (DMP) themes

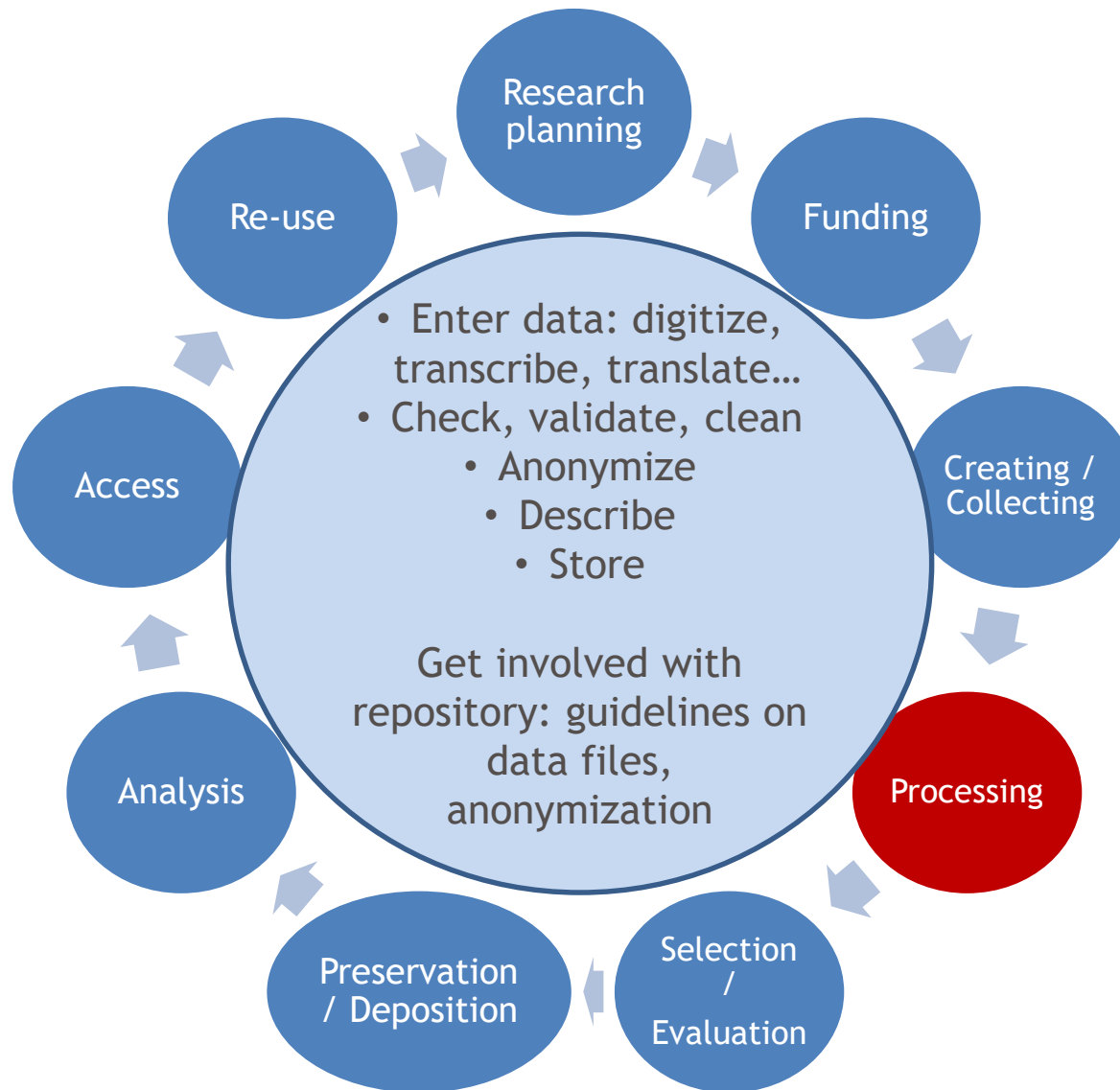
This is a consolidated set of Data Management Planning themes and guidance, issued by the Digital Curation Centre (DCC) and the University of California Curation Centre (UC3) in December 2016. The revisions are based on community feedback and helped us to consolidate an original, longer set of themes used in DMPonline. Details on the consultation process and future plans are available in the associated blog post: <http://www.dcc.ac.uk/blog/dmp-themes-and-then-there-were-14>

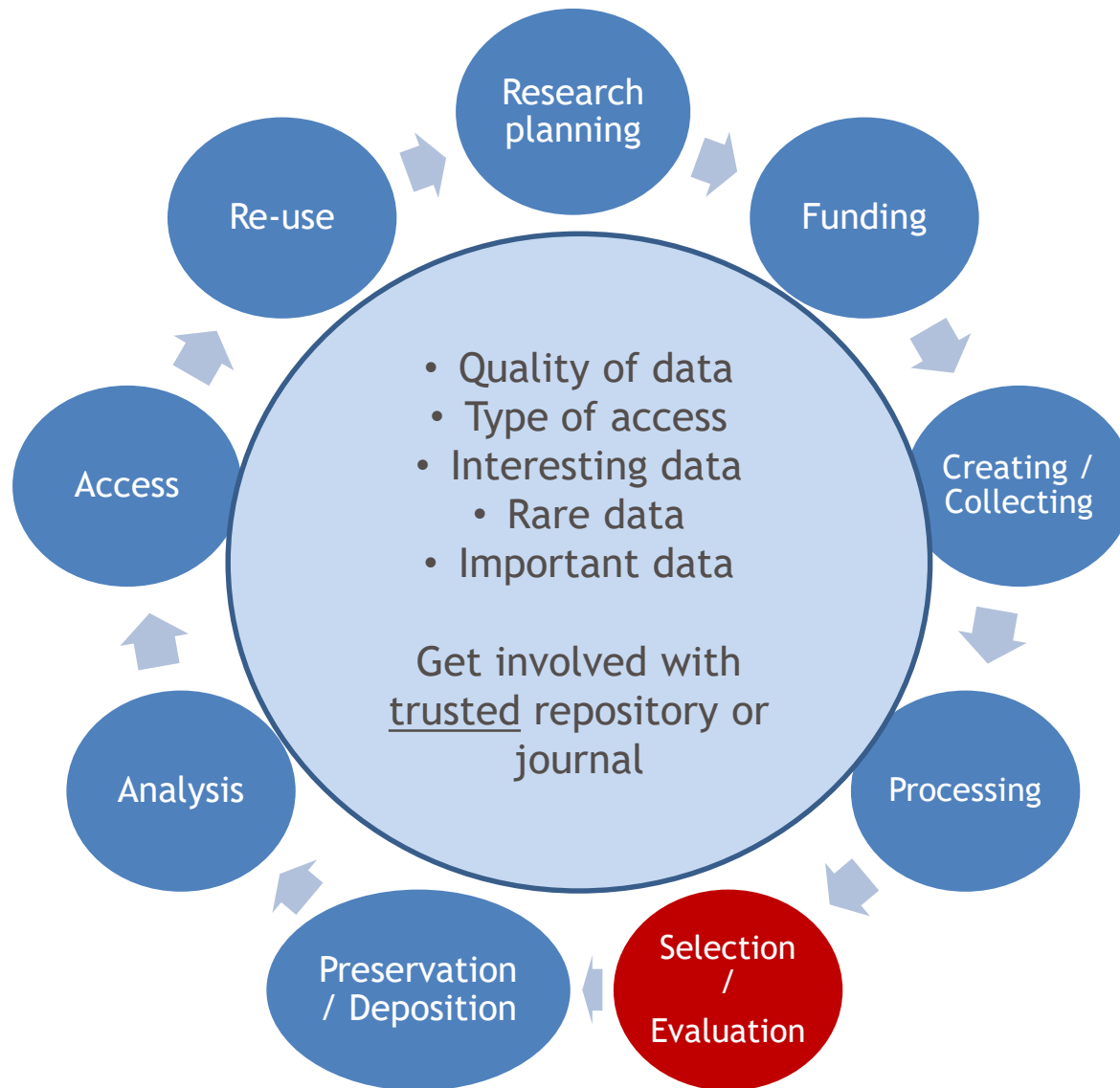
Theme	DCC & UC3 guidance
DATA DESCRIPTION	<ul style="list-style-type: none">• Give a summary of the data you will collect or create, noting the content, coverage and data type, e.g., tabular data, survey data, experimental measurements, models, software, audiovisual data, physical samples, etc.• Consider how your data could complement and integrate with existing data, or whether there are any existing data or methods that you could reuse.• If purchasing or reusing existing data, explain how issues such as copyright and IPR have been addressed. You should aim to minimise any restrictions on the reuse (and subsequent sharing) of third-party data.
DATA FORMAT	<ul style="list-style-type: none">• Clearly note what format(s) your data will be in, e.g., plain text (.txt), comma-separated values (.csv), geo-referenced TIFF (.tif, .tiff).• Explain why you have chosen certain formats. Decisions may be based on staff expertise, a preference for open formats, the standards accepted by data centres or widespread usage within a given community.• Using standardised, interchangeable or open formats ensures the long-term usability of data; these are recommended for sharing and archiving.• See UK Data Service guidance on recommended formats or DataONE Best Practices for file formats

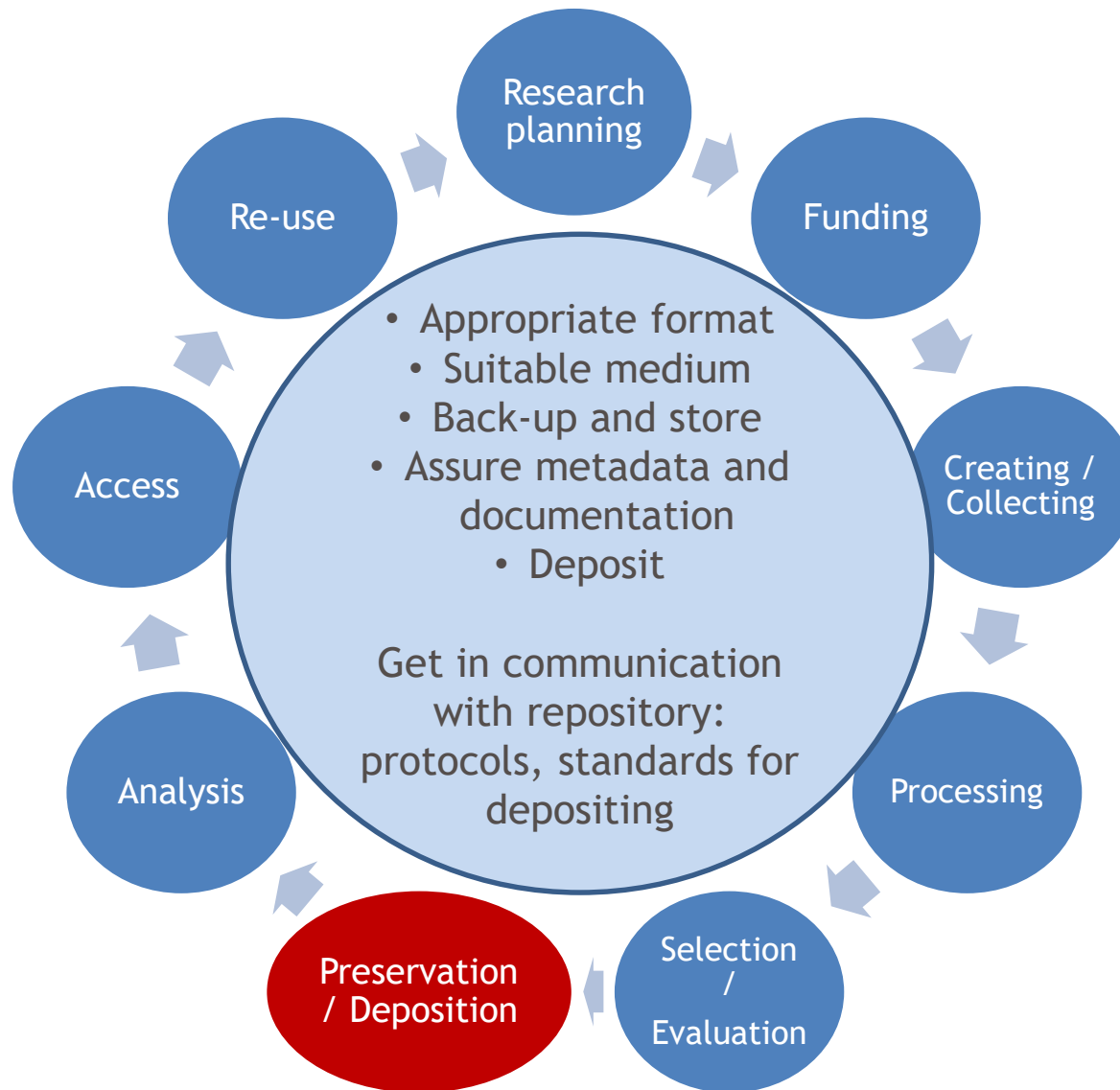


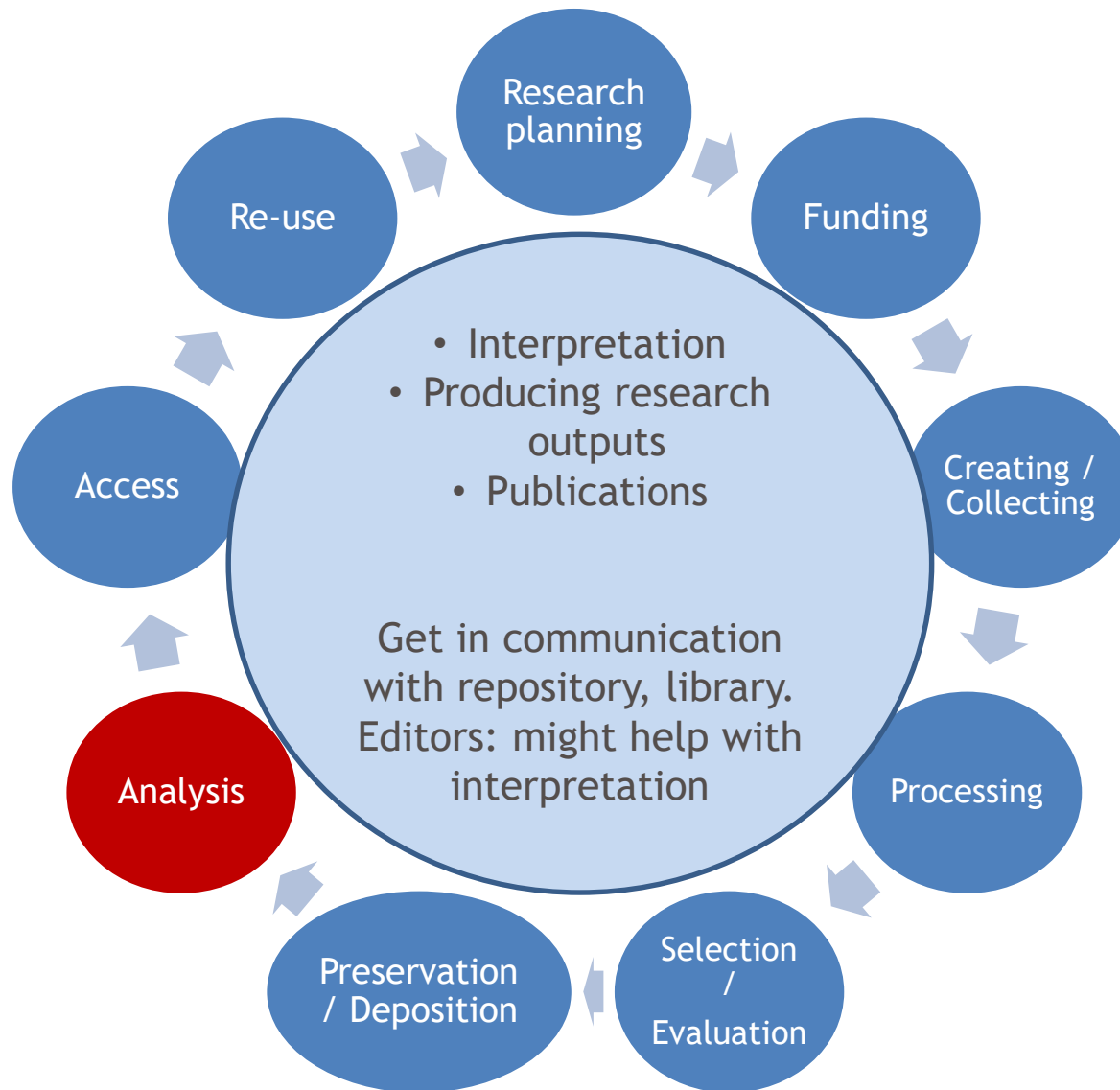




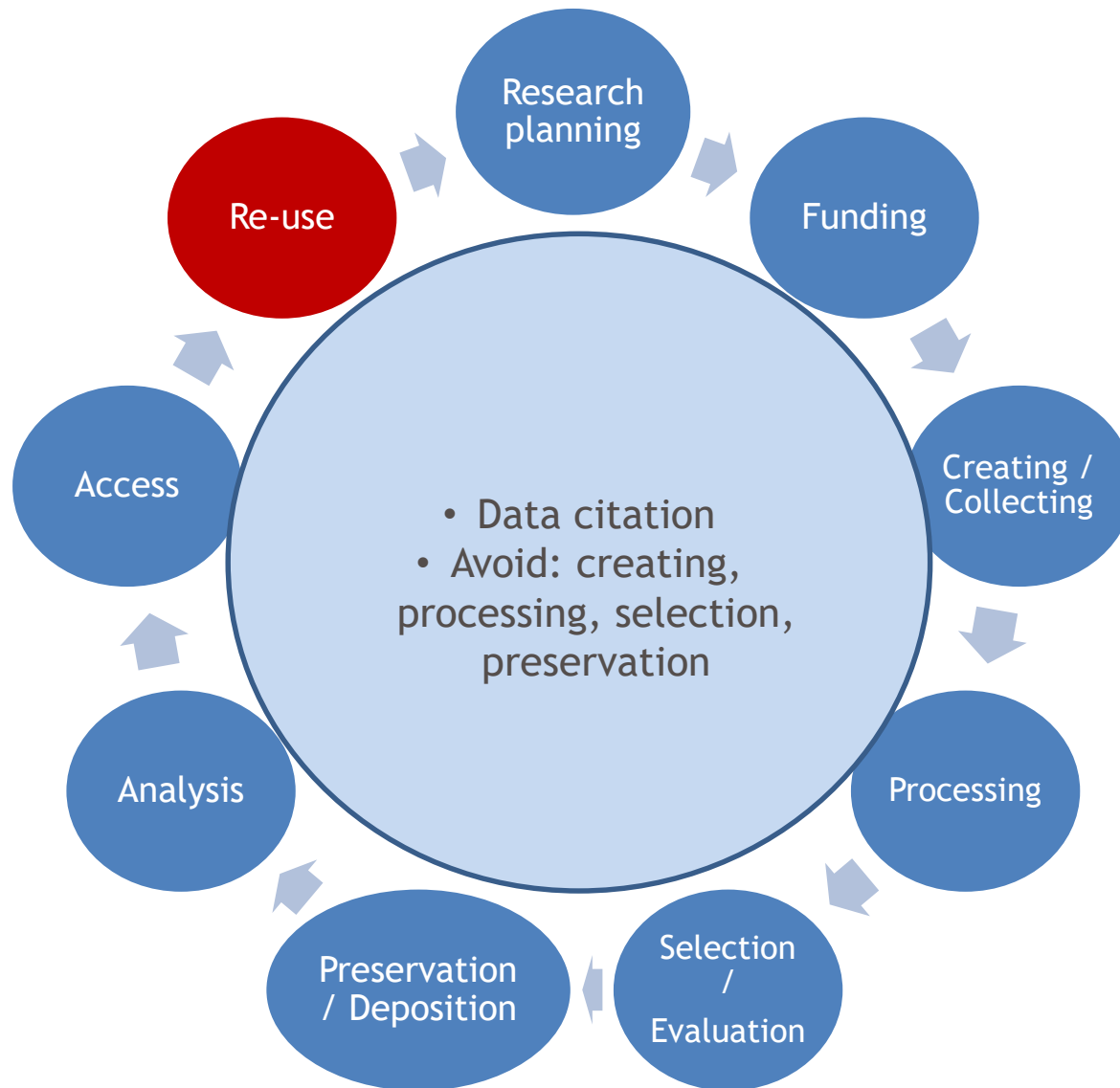








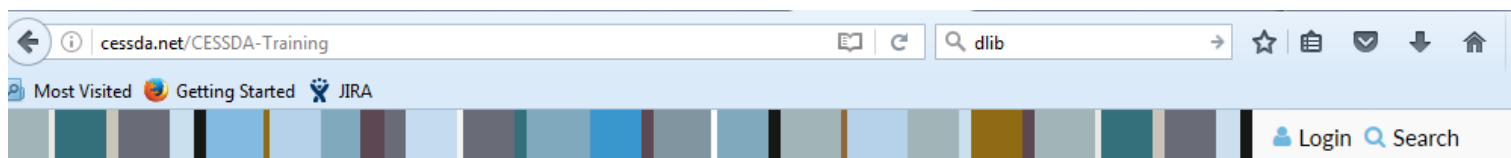




8 Further readings

- Štebe, Janez, Bezjak, Sonja, Vipavc Brvar, Irena (2015). Preparing research data for open access : guide for data producers. URN:NBN:SI:DOC-G0DPXMZ1 from <http://www.dlib.si>
- Digital Curation Centre: Resources for digital curators. [<http://www.dcc.ac.uk/resources>, 20. 1. 2015].
- Digital Curation Centre: DMPonline. [<https://dmponline.dcc.ac.uk/>, 20. 1. 2015].
- OpenMetadata.org: Data Management Plan (DMP) Editor. OpenMetadata.org. [http://www.openmetadata.org/site/?page_id=373, 21. 1. 2015].
- EDINA (2014): MANTRA – Research Data Management Training. The University of Edinburgh. [<http://datalib.edina.ac.uk/mantra/>].
- ADP: About archiving. [http://www.adp.fdv.uni-lj.si/za_uporabnike/o_arhiviranje/, 21. 1. 2015].
- UK Data Service: Prepare and manage data. [<http://ukdataservice.ac.uk/manage-data.aspx>, 21. 1. 2015].
- ICPSR: Data Management & Curation. [<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/index.html>, 21. 1. 2015].
- Tjalsma, Heiko and Jeroen Rombouts (2011): Selection of research data, Guidelines for appraising and selecting research data. Haag in Delft: Stichting SURF, Data Archiving and Networked Services (DANS), 3TU.Datacentrum. [<http://act.dans.knaw.nl/nl/over/organisatie-beleid/Publicaties/DANSselectionofresearchdata.pdf>].
- ICPSR and DANS (2010): Preparing data for sharing; Guide to Social Science data archiving. Amsterdam: Pallas Publications. [<http://act.dans.knaw.nl/nl/over/organisatie-beleid/Publicaties/DANSpreparingdataforsharing.pdf>].

8 Further readings



cesdda

Consortium of European Social Science Data Archives

About us

CESSDA Services

CESSDA Training

National Data Services

Home / CESSDA Training

About us

CESSDA Services

CESSDA Training

Research Data Management

Research Data Management
Plans

Data Archives and Digital
Preservation

Customised Training

Online Materials

CESSDA Training Staff

National Data Services

CESSDA Training

CESSDA Training promotes research data management and data curation in the social sciences. We address the needs of two communities, that of research data management and digital preservation.

Our purpose is to ensure excellence in the creation, management, and long-term preservation of research data. By promoting the adoption of standards in research data management and archiving, we support data availability, re-use, and the repurposing of archived data.

Our team offers support at every stage of the [research data lifecycle](#), from conception to archiving and re-use.

Select one of the following topics to learn more:

- [Are you interested in making sure that your data is of the best possible quality and can be reused?](#)
- [Do you need a data management plan?](#)
- [Are you involved in digital preservation?](#)
- [Do you need a specialised training?](#)

See upcoming trainings:

The FOSTER-CESSDA training project, which is part of the FOSTER initiative, supports European



Thank you!

University of Ljubljana
Faculty of Social Sciences
Social Science Data Archives
Kardeljeva ploščad 5
1000 Ljubljana
SI - Slovenia



www.adp.fdv.uni-lj.si



arhiv.podatkov@fdv.uni-lj.si



Arhiv.Druzboslovnih.Podatkov



@ArhivPodatkov

