gesis
Leibniz Institute for the Social Sciences

Introduction to Research Data Management for Social Scientists

Astrid Recker & Sebastian Netscher
CESSDA Training at the Data Archive for the Social Sciences
GESIS - Leibniz Institute for the Social Sciences
@CESSDA_Data

FOSTER

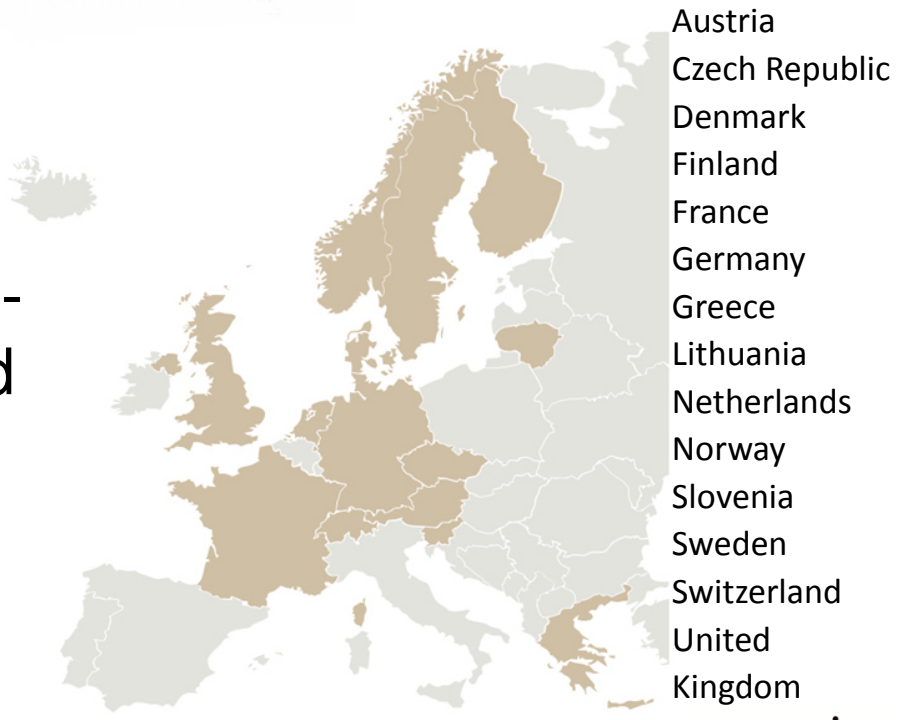
cessda

Introduction

gesis
Leibniz Institute for the Social Sciences

CESSDA, What's That?

- Consortium of European Social Science Data Archives a pan-European Research Infrastructure
  - providing large scale, integrated and sustainable data services
  - supporting research and co-operation in areas expected to be of great importance



FOSTER

cessda

gesis
Leibniz Institute for the Social Sciences

What is CESSDA Training?

- Provides training and consulting in Research Data Management and Digital Preservation
- Promotes strategies and procedures to ensure data quality and long-term availability
- Get in touch:
  - visit our webpage: http://cessda.net/CESSDA-Training
  - follow us on Twitter: @CESSDA_Data
  - follow our blog: http://www.cessdatraining.wordpress.com
  - join our workshops

FOSTER

cessda

gesis

Leibniz Institute for the Social Sciences

Image by See-ming Lee (CC-by-sa)

FOSTER

cessda

gesis

Leibniz Institute for the Social Sciences

gesis
Leibniz Institute for the Social Sciences

Outline of the Workshop

Introduction to Research Data Management for Social Scientists	
09.15-10.30	Research Data Management, Data Discovery and Collection *(Astrid & Sebastian)*
10.30-10.45	*Coffee Break*
10.45-12.30	Data Handling: Documentation, Organizaton and Storage *(Sebastian)*
12.30-13.30	*Lunch break*
13.30-15.15	Research Ethics and Legal Compliance *(Sebastian)*
15.15-15.30	*Coffee Break*
15.30-16.30	Data Sharing and Long-term Preservation *(Astrid)*
16.30-17.00	Wrap-Up *(Astrid & Sebastian)*

FOSTER

cessda

gesis
Leibniz Institute for the Social Sciences

Research Data Management

FOSTER

cessda

gesis
Leibniz Institute for the Social Sciences

What is Research Data Management (RDM)?

- RDM is like “health care” for your data
  - keeps them safe from harm
  - makes them usable and discoverable
- All strategies, processes and measures to maintain
  - data quality
  - interpretability of research results
  - (re-)usability of research data
- Specifically addresses legal and ethical issues (e.g. informed consent, ownership, licensing)

Image: CC-0

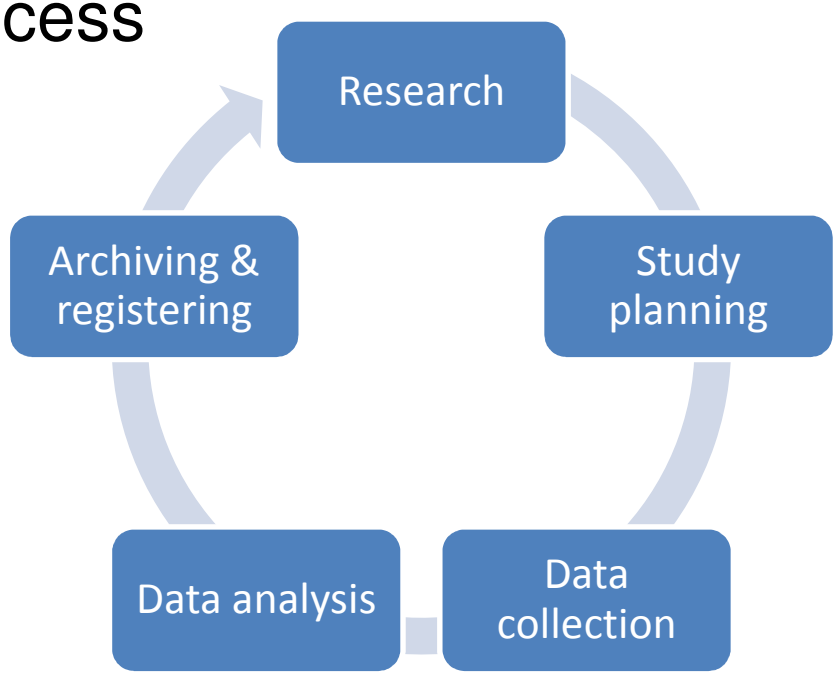
FOSTER

cessda

gesis
Leibniz Institute for the Social Sciences

RDM and the Data Lifecycle

- RDM guides the research process along the data life cycle
- Entails strategies to
  - process and validate
  - store and protect
  - document and describe
  - preserve and share data.



FOSTER
cessda

Why Do We Need RDM?

- It is an integral part of the research process
  - increases the quality of research
  - supports planning and guides research
  - creates transparency and replicability of findings
  - boosts reputation
- It may be required by others

Image: CC-0

RDM may be required by…

- Your institute, e.g. by your working contract or project agreements
- Funding agencies, e.g. to ensure reusability
- Journals, claiming your data before publishing your article
- Your supervisor…

⇒ check for such conditions

Image by A. Herrema & H. Bouwteam (CC-by)

Without RDM …

… your data remain an undiscovered mystery
  – neither understandable for you in the future
  – nor for anyone else who tries to reuse them

… your data is lost for yourself as well as for the research community, in disadvantage of research

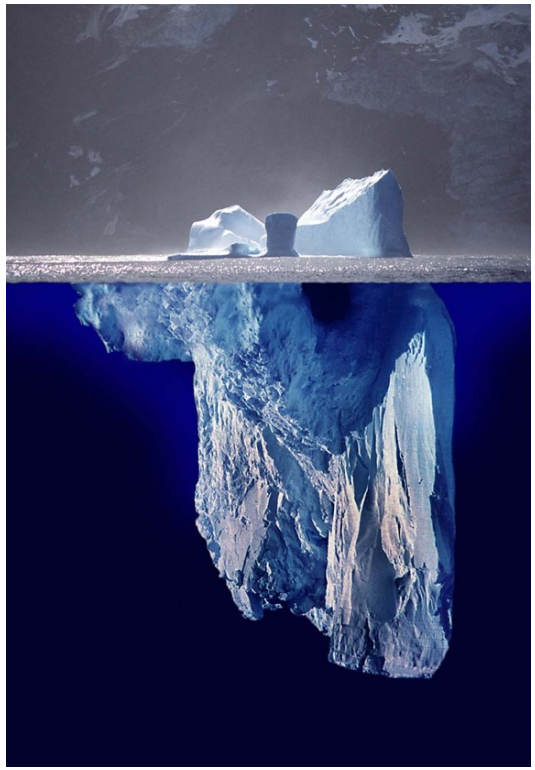
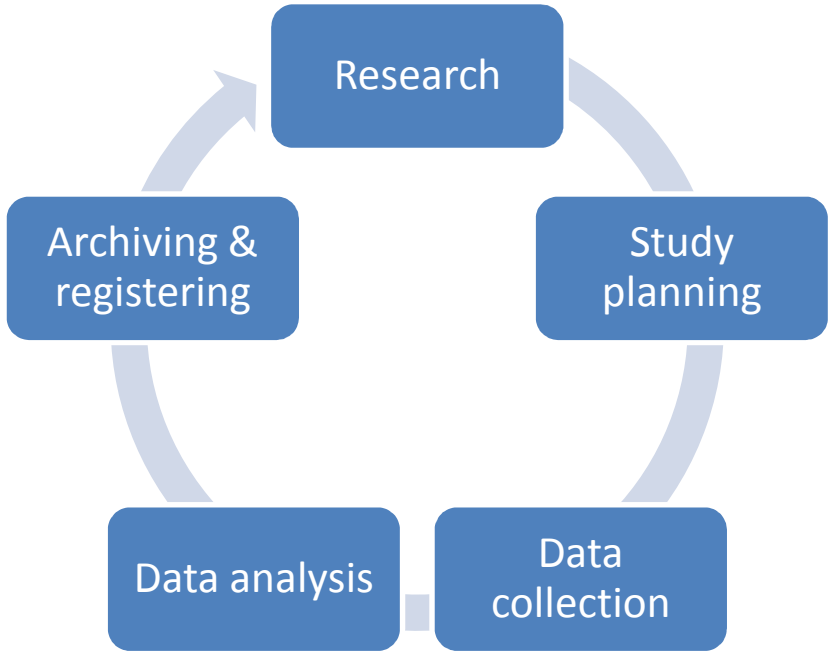


Image: Iceberg by Uwe Kils (CC-by-sa)

gesis
Leibniz Institute for the Social Sciences

Data Management Planning



FOSTER

cessda

The Data Management Plan (DMP)

- A DMP is a systematic documentation of RDM
- It desribes your strategies to
  - process and validate
  - store and protect
  - preserve and share your data

throughout the data lifecycle

Image: CC-0

How to Write a DMP?

- A DMP uniquely relates to your research project
  - start right at the beginning of your research project
  - document what you did and why
  - frequently up-date and adapt it
- A DMP is not just a plan, it is the implementation of a (research) plan

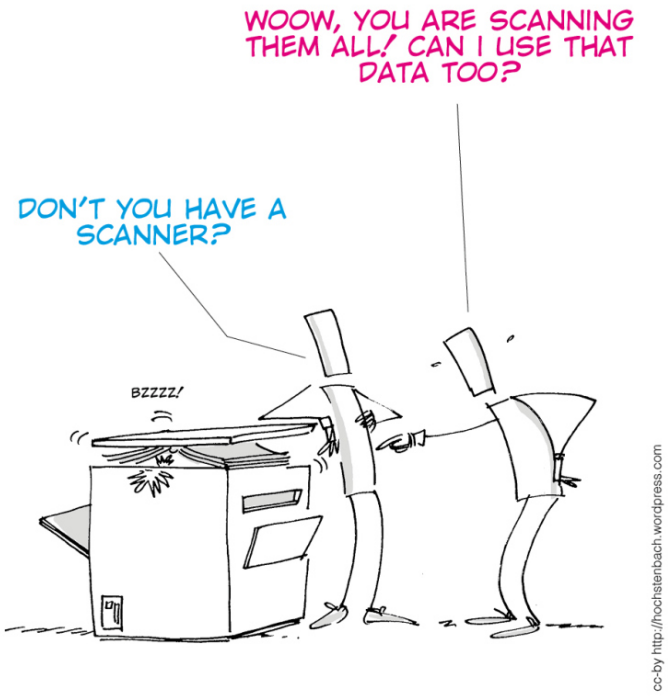


Image by P. Hochstenbach (CC-by)

Where to Get a DMP?

- Rely on the data life cycle
- Various online templates and tools, e.g.
  - DCC. (2013): *Checklist for a Data Management Plan*. http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf
  - DMPTool (2015): *Data Management Planning Tool*. https://dmptool.org//
- Use the template provided in this workshop

gesis
Leibniz Institute for the Social Sciences

A First Look at the DMP

- Have a look at the DMP template in your folder
- Familiarize yourself with its sections and subsections
- Consider …
  - … parts you would be able to fill in off the top of your head
  - … parts you would have to look for further information
  - … obstacles or problems you encountered in the past


Image: CC-0

individual work - time: about 15 minutes

FOSTER

cessda

The Structure of the DMP Template

- Consists of seven parts
  - Cover page and general information on the project
  - Six sections (with subsections):
    1. data collection ⇒ *session 1*
    2. documentation and metadata ⇒ *session 2*
    3. storage, organization and security ⇒ *session 2*
    4. ethics, legal compliance and Intellectual Property Rights ⇒ *session 3*
    5. preservation and sharing ⇒ *session 4*
    6. responsibilities and resources

gesis
Leibniz Institute for the Social Sciences

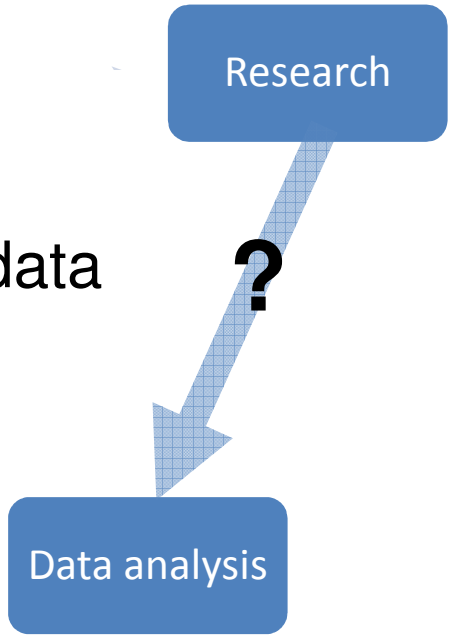
Discovering and Collecting Data

FOSTER

cessda

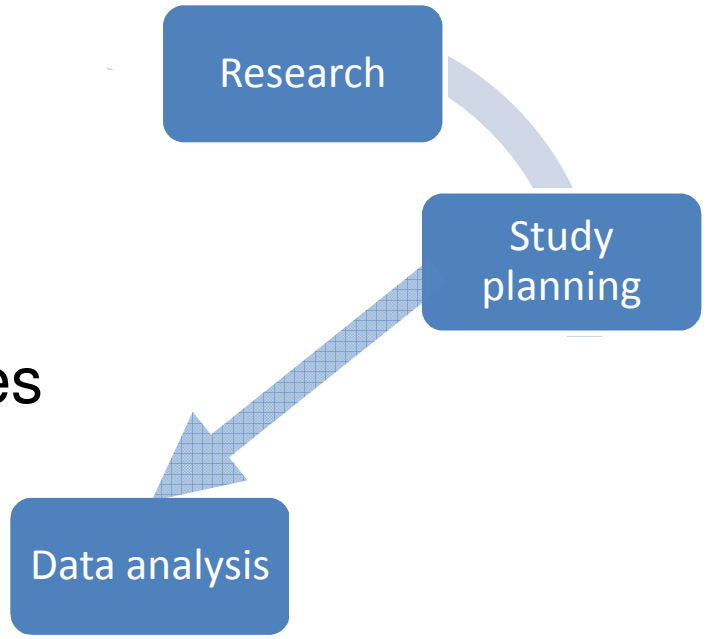
Research Idea and Research Data

- Research idea
  - ⇒ strategy to test this idea
  - ⇒ definition of data needed
    - – qualitative and/or quantitative data
    - – individual and/or macro data
    - etc.
- Decision on
  - – whether to reuse existing data
  - – or to collect new data



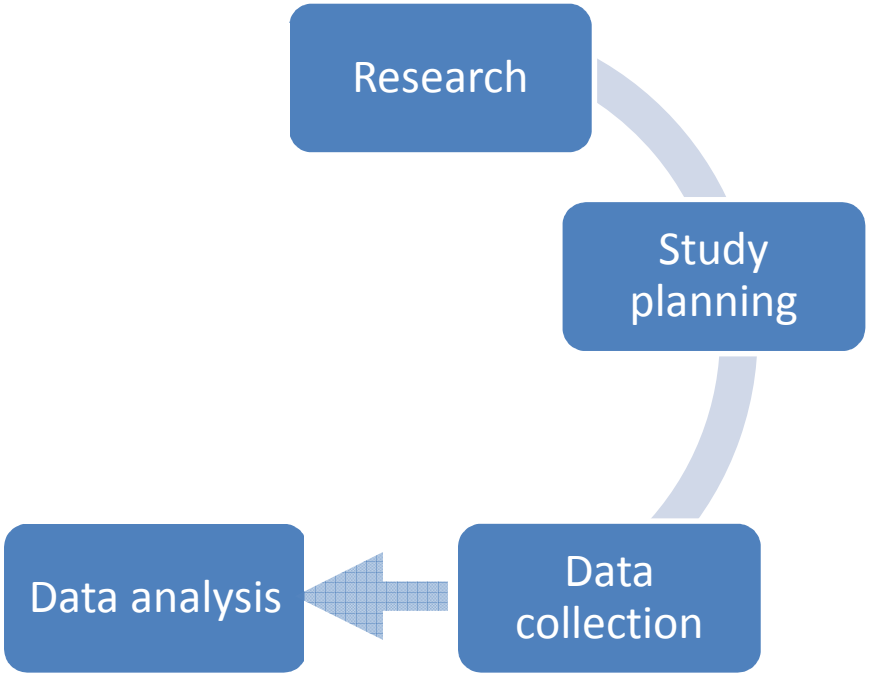
Discovering Research Data

- Reuse of existing data
  - a lot of research data available
  - saves time and effort
- Check for reuse of existing data
  - networks, papers and conferences
  - repositories,
    - ⇒ e.g. www.re3data.org
  - CESSDA Data Catalogue
    - ⇒ http://cessda.net/CESSDA-Services/Resources/Data-Catalogue



Collecting Research Data

- No existing data available for reuse
  ⇒ need to collect data
- Data collection process
  - use of standards and methodologies
  - quality assurance
  - documentation



Further Readings

- CESSDA (2015): Data Catalogue. http://cessda.net/CESSDA-Services/Resources/Data-Catalogue.
- Corti, L., van den Eynden, V., Bishop, L. & Woollard, M. (2014): Managing and Sharing Research Data. A Guide to Good Practice. London: Sage Publication Ltd.
- DANS (2010): Preparing Data For Sharing. Guide to Social Science Data Archiving. DANS Data Guide 8. The Hague: Royal Netherlands Academy of Art and Sciences & Netherlands Organization for Scientific Research.
- Horton, L., van den Eynden, V., Corti, L. & Bishop, L. (2011): Data Management Recommendations for Research Centers and Programmes. Essex: UK Data Archive.
- Jones, S. (2011): How to Develop a Data Management and Sharing Plan. Glasgow: Digital Curation Centre (DCC).