# INCLUSION OF DATA ARCHIVES IN DATA MANAGEMENT PLAN

Irena **Vipavc Brvar,**
ADP, University of Ljubljana

5th Conference of the
**European Survey Research Association**
Ljubljana, Slovenia, July 15-19, 2013

## Why DA support Data Management Plan

Data DA might receive:

- Without proper documentation or it is incomplete

- Data are not labeled properly or are not saved in correct format (additional problems with older data)

- Data are not anonymised / interviews might be in digital (audio-visual data) format – not transcribed

- Open ended questions not coded

- Which is the last version of my data?

- Consent form did not include necessary wording
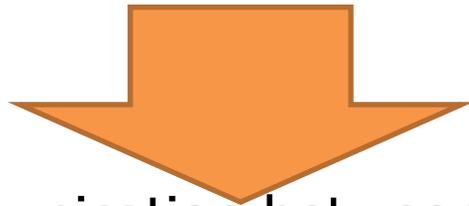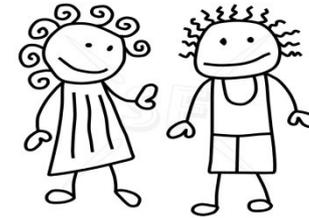
A lot of communication between DA and researchers to prepare complete documentation and clean data. It usually does not happen after survey itself. Researchers forget detailed information.

## DATA MANAGEMENT PLAN

A **data management plan** is a formal document that outlines how you will **handle your data** both **during** your research, and **after** the project is completed.

The goal of a data management plan is to consider the many aspects of **data management**, **metadata generation**, **data preservation**, and **analysis** before the project begins.

This ensures that <u>data are well-managed in the present, and prepared for preservation in the future</u>.

<p align="right">WIKI</p>

Responsible researcher should have a DMP.

# SHARING RESEARCH DATA +

Research data are valuable resources, usually requiring much **time** and **money** to be produced. Many data have a significant value **beyond usage for the original research**.

Sharing research data:

- encourages scientific enquiry and debate,

- promotes innovation and potential new data uses,

- leads to new collaborations between data users and data creators,

## SHARING RESEARCH DATA +

- maximises transparency and accountability,
- 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,
- provides important resources for education and training.

(van den Eynden, 2011)

Open access principles.
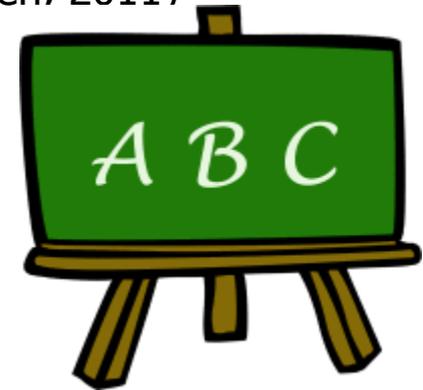
## SHARING  RESEARCH  DATA -

In most fields, the reward comes from **publication**, not from data management.

SI: study equivalent to publication

The categorization of scientific publications using Slovenian Research Agency:

Inclusion of study in **bibliography of researcher** as **Complete Scientific Data base** or Corpus

----> gaining scientific points



People are hired and promoted based on their publication record rather than on the quality of their metadata.
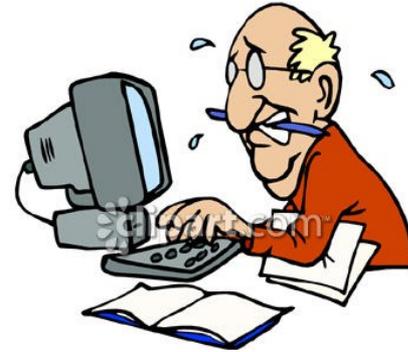
There's got to be a better way to get an employer's attention. . .

Prime Personnel Consulting

Documenting data might be
a labour-intensive process.

Documenting methods,
instrumentation, procedures, and producing metadata at a level
that the data are interpretable by others, can require much more
labour than documentation for use by oneself or one's team.

(Borgman, 2010)

Researchers are concerned about establishing the priority of
their claims on research findings in the face of competitions.

Embargo periods

Public founders of research increasingly follow

[OECD Principles and Guideliness for Access to Research Data from Public Funding](#)

Which say that publicly funded research data
should be as much as possible openly available
to the scientific community.

In the UK, funding bodies such as the Economic and Social Research Council (ESRC), the Natural Environment Research Council (NERC) and the British Academy **mandate researchers to offer all research data generated during research grants** to designated data centres – the UK Data Service and NERC data centres.

(van den Eynden, 2011)

Some others funders in UK have similar data policies (BBRSC, MRC and Wellcome Trust).

Proposals submitted to above mentioned funders should include **Data management plan**.

Same goes for **USA** (NSF, NIH), having in mind that all researchers are expected to be able to explain and defend their results.

Basic data management is required by the *Australian Code for the Responsible Conduct of Research* and ARC and NHMRC require compliance with it.

# JOURNALS

Journals increasingly require data that form the basis for publications to be **shared or deposited within an accessible database or repository.**

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. Therefore, a condition of publication in a **Nature** journal is that **authors are required to make materials, data and associated protocols promptly available to readers without undue qualifications**.

Availability of data is nowadays required by most of most cited journals. And it is shown that higher the journal's impact factor more likely it is that author / researchers will make data available without any questions (Piwowar and Chapman, 2009).

There are known surveys that show that publicly available data are significantly associated with an increase in citations, independently of journal impact factor, date of publication, and author, country of origin (Piwowar, Day, Fridsma, 2007).

On the other hand some journals require that **data are made available upon request**.

In empirical study that Savage and Vickers ([2009](#)) have made they received only one of ten raw data sets requested. This suggests that journal policies requiring data sharing do not lead to authors making their data sets available to independent investigators.

Data archives are introducing **persistent identifiers** for referencing and citing data. **DOI** and **URN** are most common.

# THE DATA CITATION INDEX – WEB of SCIENCE
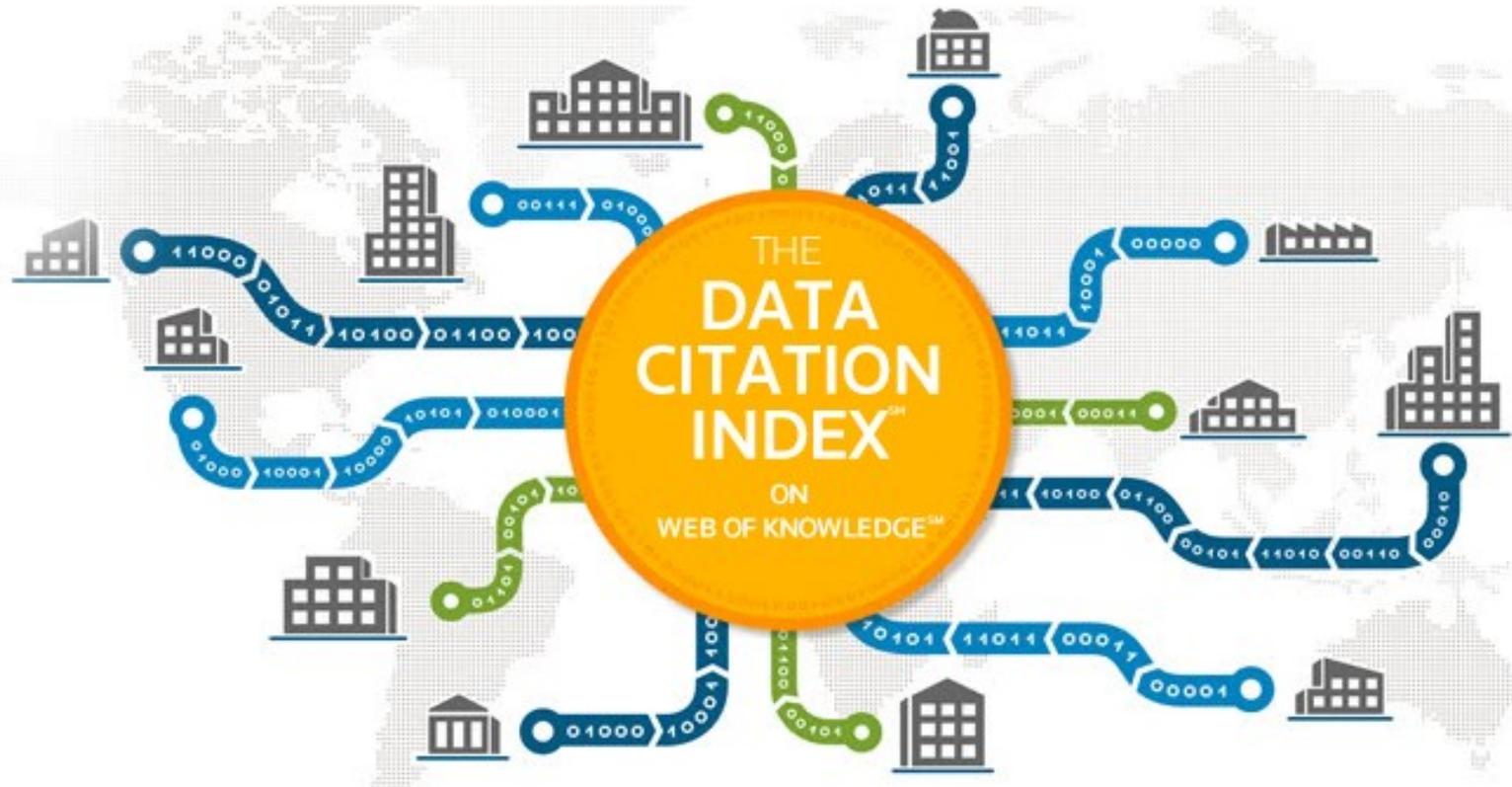


The Data Citation Index on the Web of Knowledge platform provides a single point of access to quality research data from repositories across disciplines and around the world.

# ESRC GUIDELINES TO PREPARE DATA MANAGEMENT PLAN

Make sure you have: The ability and motivation to manage your time effectively.

ESRC guidelines to prepare Data management plan. It should include the following topics:

1. ASSESSMENT OF EXISTING DATA
2. INFORMATION ON NEW DATA
3. QUALITY ASSURANCE OF DATA
4. BACK-UP AND SECURITY OF DATA
5. EXPECTED DIFFICULTIES IN DATA SHARING
6. COPYRIGHT (INTELLECTUAL PROPERTY RIGHT)
7. RESPOSIBILITIES
8. PREPARATION OF DATA FOR SHARING AND ARCHIVING

**Should the budget and its justification specifically address the costs of implementing the Data Management Plan?**

Yes, include costs for data preparation in proposal of your project. One can cost the additional expenses – above standard research procedures and practices – that are needed to make research data sharable beyond the primary research team.

Some DA will help you with that - UK Data Service have [Data management costing tool and checklist](#) prepared.

There are several examples of data management plans you can find on-line from [ICSPR](#), [DCC](#), [UK DA](#), [MIT Libraries](#).

Research Data MANTRA is a course designed for PhD students and others who are planning a research project using digital data.

A course *Data Intelligence 4 Librarians* designed by 3TU.Datacentrum and DANS.

## Tools

DMPTool: Web-based tool to build and edit a customized plan according to select funder requirements.

DMP Online: Web-based tool to build and edit a customized plan according to select funder requirements. Also available is a Template for a Data Management Plan (Word), which draws together details researchers often are asked to provide in data management plans. (Digital Curation Centre)

## ADDITIONALLY TO THE DATA

It may be important to manage the following **research records** both during and beyond the life of a project:

- Correspondence (electronic mail and paper-based correspondence)
- Project files
- Grant applications
- Ethics applications
- Technical reports
- Technical Appendix
- Research reports
- Research publications
- Master lists
- Signed consent forms
- Internal social media communications such as blogs, wikis etc.

**Survey among secondary school pupils**

In proposal of the project we can read:

… One of the WP 7 products will be a publicly accessible international database with the corresponding documentation (questionnaires, technical information, sampling information) based on which anyone interested will be able to perform secondary analyses. The international database will be published on the web page and publicly after the publication of the final report.

Prior agreement between researchers (**7 partners = countries**) about making data available and when and which DA will host international database.

DDI standard used for metadata documentation – easy transfer / translation to other national catalogues.

# REFERENCES AND FUTURE READING

- van den Eynden, V., Corti, L., Woollard, M., Bishop, L. and Horton, L. (2011). Managing Research Data:. Best Practice for Researchers. Colchester: UK Data Archive, University of Essex. [www.data-archive.ac.uk/media/2894/managingsharing.pdf]

- Borgman, Christine L. (2010). Research Data: Who will share what, with whome, when, and why? China-North Americal Library Conference. http://works.bepress.com/borgman/238/]

- National Science Foundation (2010). Data Management & Sharing FAQ. NSF. [http://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp]

- Australian Government, National Health and Medical Research Council, Australian Reseach Council (2007): Australian Code for the Responsible Conduct of Research [http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/r39.pdf]

- Digital Curation Centre – DCC(2010). Data management plans. [http://www.dcc.ac.uk/resources/data-management-plans]

- Jones, Sarah (2011). Develop a Data Management and Sharing Plan. Digital Curation Centre. [http://www.dcc.ac.uk/resources/how-guides/develop-data-plan]

- MIT Libraries (2010). Data management and publishing: organizing your files. [http://libraries.mit.edu/guides/subjects/data-management/organizing.html]