The role of librarians in the research data lifecycle

Sonja Bezjak, ADP

Academic libraries' support for open science in the European Research Area, Staff training days at the University of Ljubljana, 15-18 May 2017
Content

1) Who we are and what we do
2) Challenges in research environment
3) Librarians as stakeholders in RDLC
4) Social science data repositories

5) Feedback
Social Science Data Archives, University of Ljubljana

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)
- Member of CESSDA
- International projects (DwB, Foster, SEEDS, SaW, SERISS ...)
ADP’s mission

Social Science Data Archives = in Slovenian →
Arhiv družboslovnih podatkov (ADP)

- **Social Science Data Archives** offer access to data that are interesting for **social science analysis**, with emphasis on problems related to Slovenian society.

- **Priority** is given to **theoretically significant** and **methodologically well designed** studies, especially data gathered over a period of time and international comparative data that include Slovenia.

The ADP’s data catalogue: [http://www.adp.fdv.uni-lj.si/opisi/serije/]
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
- ...

ADP Selection
Added value
Curation
Access
...
Consortium of European Social Science Data Archives

- 15 CESSDA members
- Austria - AuSSDA
- 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’s objectives

• To support national and international research and cooperation in areas expected to be of great importance in the future

• To facilitate access to social science (and related areas) data resources for researchers regardless of the location of researcher or data

• To provide large scale, integrated and sustainable data services to the social sciences and facilitate and support research, teaching and learning
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
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
CESSDA’s European projects

- Strengthening and widening the European infrastructure for social science data archives
- A cluster project aiming at tackling important socio-economic challenges for Europe
- Dealing with new technologies in ICT infrastructure building for big data management and data manipulation for the societal challenges that Europe is facing
- European Research Infrastructures in the International Landscape

RISCAPE
... in social sciences exist all over Europe but researchers are still not motivated to publish and share data.
Open Data project in Slovenia (2010-2013)

1) 22 semi-structured interviews:
- researchers, librarians, heads of research institutes
- 10 research institutes, 6 faculties
- 17 research disciplines: physics, biology, medicine, civil engineering, archeology, social work, economy, musicology, anthropology, languages...

2) 3 workshops:
- Workshop 1: Problems and Solutions in the Field of Data Services in Slovenia
- Workshop 2: Policy of Research Data Management in Slovenia
- Workshop 3: Advanced Technology for Establishing Data Infrastructure in Slovenia

3) Individual working visits/consultations:
- Information Commissioner
- Intellectual Property Institute
- Research Centre of the Slovenian Academy of Sciences and Arts
- DARIAH, CLARIN representatives in Slovenia...

• 'OPEN DATA' Action plan for establishing a system of open access to research data funded by public sources (2013)
• National strategy of open access to scientific publications and research data in Slovenia 2015-2020
Regarding documentation, retention and access to research:

- Researchers have **different habits and views**
- Research institutes follow **unwritten rules and practices**

But mainly they face identical problems:

lack of knowledge, time and finances for dealing with research data, 
extremely competitive scientific environment

*OPEN DATA* Action plan for establishing a system of open access to research data funded by public sources (2013)
What we found out? Open data in Slovenia (2010-2013)

Major barriers in achieving open data:

1) big differences in the development of data infrastructure and services
2) dilemmas and fears emerging from questions related to law (intellectual property, personal data protection)
3) absence of culture of data sharing
4) absence of framework policy, which would help in research data management

Researchers mainly understand that open access is a must, they are aware of disorder.

And express the need of developing a framework policy and rules by which the system would be justly organized and would provide service to researchers (data creators and data users).

OPEN DATA Action plan for establishing a system of open access to research data funded by public sources (2013)
Reality ...

Scene from the past?

Data sharing

Research data alliance meeting 2014

Research data alliance meeting 2014
“...constitute primary sources that underpin scientific research and enable derivation of theoretical or applied findings.”

All you need is in the article...
There is a need to...

- Raise awareness
- Share the knowledge
- Change culture of sharing
- Support researchers

DATA SCIENCE ROLES

Collaboration of various stakeholders is needed:

- Political and funding level
- Institutional
- Research Offices
- Libraries
- Repositories...

Also common tools, standards, services, infrastructure...
Open data is data that is free to access, reuse, repurpose, and redistribute.

https://www.openaire.eu/opendatapilot
Change in EU: Open access & Open data

1. Background – Extension of the Open Research Data Pilot in Horizon 2020

⚠️ Please note the distinction between open access to scientific peer-reviewed publications and open access to research data:

- **publications** – open access is an **obligation** in Horizon 2020.
- **data** – the Commission is running a flexible pilot which has been extended and is described below.

See also the guideline: [Open access to publications and research data in Horizon 2020](https://example.com).

How ADP as service provider supports RDM

- Workshops
- Webinars
- Brochures
- Discussions
- Consultations...


Also for librarians!
ADP’s RDM Activities for librarians

• 2 one day workshops

→ 18. junij 2014, Ljubljana: Role of librarians in opening up research data and managing bibliography of researchers (also in EN)
→ 26th October 2016, Ljubljana: Knjižničarji kot podpora pri načrtovanju ravnanja z raziskovalnimi podatki (only in SI)

• Lectures at several different occasions

Topics shared with librarians:

1. Benefits of publishing and sharing
2. What is research data
3. Types of data
4. Research data life-cycle
5. What is DMP
6. ...

ADP's RDM Activities for librarians
How librarians see their role in RDM

Q4 How do you imagine you could help researchers in research data management?

In general:
• advise, give information, suggesting literature, references... (3)

Discovering data sources:
• Help when searching for content, select of data sources, help in data citing (3)

Long-term curation and access:
• To support preparation of research data for publishing, I advice researcher to get in contact with ADP, to help digitalization of data ... (5)

Results of the survey carried out before ADP’s workshop for data librarians, 26.10.2016, N= 12)
Obstacles:

• They feel that they lack the knowledge and confidence to talk to researchers on this topic
• They are overloaded with their own work
Indicate how strongly you agree/disagree with the statements

1. I understand the basic stages of the data lifecycle.
   - Strongly Disagree
   - Disagree
   - Neither Agree nor Disagree
   - Agree
   - Strongly Agree

2. I feel confident in my ability to assist researchers with their data.
   - Strongly Disagree
   - Disagree
   - Neither Agree nor Disagree
   - Agree
   - Strongly Agree

Benefits of sharing data

For the public:

• Research findings based on existing data on certain topics can improve quality of life; in Social Sciences gaining a better understanding of the community is important, including building collective identity and eliminating prejudice, discrimination, inequality etc.

• Open access to data also introduces savings relating to public spending on research and education.

Benefits of sharing data

For the scientific community:

• **New research findings** which can only be achieved by using and combining available data from multiple sources, e.g. studies of trends or international comparative research, studies of rare or small populations, meta-analysis.

• More robust research findings can be derived as a result of validating and through **further analysis based on data related to publication**.

• Increasing the number of findings by enhancing the number of users of the same data, with efficiencies gained through **removal of duplication of effort** and/or the need to repeat raw data collection due to a shortage of funds and time.

• Training of future scientists, **using research data for teaching purposes**.

Benefits of sharing data

For the data creators:

• Future reuse of data by data creators themselves is made easier if data are properly documented and preserved.

• Indirect benefits for data creator’s academic reputation and career development include:
  • increased citation counts and wider dissemination of their research findings;
  • data may be formally assessed for scientific excellence and form part of researcher’s bibliography thus adding points to the score of scientific excellence.

Data, data...

NO USE, CAN NOT DRINK IT

DATA OCEAN

DATA HANDLING

IS THAT SECURE?
... „is defined as recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings; although the majority of such data is created in digital format, all research data is included irrespective of the format in which it is created.“

https://www.epsrc.ac.uk/about/standards/researchdata/scope/
Data varies by how it’s...

... *produced:*
- observation,
- experimentation,
- simulation,
- derivation,
- compilation.

... *represented:*
- text,
- numerical,
- multimedia,
- model,
- software,
- discipline-specific,
- instrument-specific.

... *stored:*
- ASCII,
- PDF,
- SPSS,
- Excel,
- TIFF,
- Java
- ...

... *conceptualized:*
- life-sciences,
- physical sciences,
- social sciences,
- humanities.

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

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

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

CESSDA SaW Webinar available on – Research Data Management Community Training – 23-06-2016
Data sharing – various possibilities

Different channels
- web page (own, institutional)
- data journal
- repository

Different approaches of data publishing:
- self-archiving
- peer-review

Repositories with different scope:
- institutional
- general
- disciplinary

Different types of repositories:
- certified / licensed or non
- open access
- ...

1. Use an external **data archive** or **repository already established for your research domain** to preserve the data according to recognised standards in your discipline.

2. If available, use an institutional research data repository, or your research group’s established data management facilities.

3. Use a cost-free data repository such as Zenodo.

4. Search for other data repositories here: [re3data.org](http://re3data.org). On top of specific research disciplines you can filter on access categories, data usage licenses, trustworthy data repositories (with a certificate or explicitly adhering to archival standards) and whether a repository gives the data a persistent identifier.

[How to select repository](https://www.openaire.eu/opendatapilot-repository)
In re3data you can search repositories by

- Subjects
- Content Types
- Countries
- AID systems
- API
- Certificates
- Data access
- Data access restrictions
- Database access
- Database access restrictions
- Database licenses
- Data licenses
- Data upload
- Data upload restrictions
- Enhanced publication
- Institution responsibility type
- Institution type
- Keywords
- Metadata standards
- PID systems
- Provider types
- Quality management
- Repository languages
- Software
- Syndications
- Repository types
- Versioning
How do you imagine role of a librarian in a research data life cycle?
• 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.

• Provides information about accessioning datasets into a bibliographic system for the purpose of scientific evaluation and resource discovery.

• Provides data citation guidance.

• Provides information on datasets, modes of access and possible use at information literacy training.

• Provides information about linking data with publications and research projects.

URN:NBN:SI:DOC-G0DPXMZ1 from https://www.dlib.si, p. 15
About the Czech Social Science Data Archive

Mission

The Czech Social Science Data Archive (ČSDA) at the Institute of Sociology of the Czech Academy of Sciences accesses, processes, documents and stores data files from social science research projects and promotes their dissemination to make them widely available for secondary use in academic research and for educational purposes.

ČSDA promotes the objectives of the OECD Principles and Guidelines for Access to Research Data from Public Funding (OECD 2007) in the area of social science research in the Czech Republic.

ČSDA participates in international networks of data organisations and it is the Service Provider for the Consortium of European Social Science Data Archives (CESSDA). The major objective for CESSDA is to provide seamless access to data across repositories, nations, languages and research purposes.
About LiDA

LiDA provides virtual digital infrastructure where researchers can deposit, search, browse, make online analyses and download almost 300 quantitative and qualitative social sciences and humanities data sets.

Training

LiDA serves as a tool to increase methodological competence of researchers by providing methodological assistance and training by means of distance learning, organization of data confrontation and methodological seminars.

Data

As the National Member Institution at the Inter-university Consortium for Political and Social Research (ICPSR) LiDA also provides access to international empirical data resources for Lithuanian researchers at major science and research institutions.

Research

LiDA coordinates international data collection initiatives in Lithuania, e.g. European Social Survey (ESS), European Election Study (EES) and International Social Survey Programme (ISSP).
Denmark - Rigsarkivet
Online Study Catalogue - ZACAT

ZACAT is a social science data portal allowing you to search for, browse, analyse and download social science survey data. The studies in this catalogue present a small selection of the studies available at the GESIS Data Archive for the Social Sciences - however the ZACAT holdings are extended on a regular basis. Beside documentation on study level (e.g. abstract, related publications), all studies come along with documentation of full question and answer texts on variable level. Moreover, for some of the collections documentation on variable level is offered in two or more languages.

The available study collections at ZACAT

- International Social Survey Programme (ISSP) - web pages
- Comparative Study of Electoral Systems (CSES) - web pages
- Eurobarometer:
  - Standard and Special Topic Eurobarometer - web pages
  - Candidate Countries Eurobarometer - web pages
  - Central and Eastern Eurobarometer - web pages
- European Values Study:
  - Longitudinal Data File 1981-2008 - web pages
  - 1981 1st wave - web pages
  - 1990 2nd wave - web pages
  - 1999/2000 3rd wave - web pages
  - 2008 4th wave - web pages
- ALLBUS (German General Social Survey) - web pages
- Studies from Eastern Europe:
  - Comparative Studies - web pages
  - Election Studies - web pages
France - Progedo

Développer la culture des données

The PROGEDO large infrastructure ensures the implementation of a public policy for social sciences and humanities.
Social sciences and humanities data archive

Researcher: Johan Surkyn – Promotor: Patrick Deboosere
Partners: Université Catholique de Louvain (UCL), National Archive
2014-2016 | federal funding

Abstract:

Pilot year for the Belgian data archive for the social sciences and humanities under the European CESSDA framework – Consortium of European Social Science Data Archives. A Data Archive is a virtual space where survey or register data from the Social Sciences and Humanities can be stored, documented and made available for further research. ID will work in collaboration with BELSPO, the UCL and the National Archive to explore the alternative paths towards such a Data Archive in Belgium, in analogy with data archives in many other European countries. Furthermore, a new operational website will be dedicated to archived data originating from and beyond past studies carried out at both research centers.
Latest Data:

**SN173**
**ISTAT**

The Italian Labour Force Survey is the main source of statistical information on the Italian labor market. The information gathered from the population constitutes the basis on which official estimations [...] 

**SN151**
Household Budget Survey (2000)

Latest News:

**30/03/2017**
Flash Eurobarometer study released by the GESIS archive

The GESIS Data Archive has made available new Flash Eurobarometer study, listed below. The data are freely downloadable from the GESIS website, following the DOI address listed alongside the individual [...] 

**14/02/2017**
ICPSR Summer Program 2017 in Quantitative Methods of Social Research
RODA is the national Romanian institution specialised in archiving electronic data collections obtained by social research.

The archive contains data collections accessible for the academic community and the interested public, for secondary and comparative analysis, under certain access conditions ranging from free access to some level of restriction imposed by owners.

The main goal of the data archive is to preserve these data indefinitely, and to serve as an intermediary between the data owners and data users.

RODA is integrated into a global archive network, being affiliated since February 2002 to CESSDA (Council of European Social Science Data Archives) and since December 2002 to IFDO (International Federation of Data Organisations). By these links, international data exchanges are enabled; through RODA, users have access to national and international datasets from other European data archives.

In December 2013, CESSDA has become an ERIC - European Research Infrastructure Consortium, a research entity with a pan-European priority, located in Bergen, Norway. Since our IT system has been recently upgraded, RODA is prepared to interface the new institution.

RODA is now organised an institutional consortium, governed by the University of Bucharest, the National Institute of Statistics and the Centre for Quality of Life Research of the Romanian Academy of Sciences.
Estonia - Eesti

Eesti Sotsiaalteaduslik Andmearhiiv
ESTA

<table>
<thead>
<tr>
<th>Eesti Sotsiaaluruingute Andmebaas</th>
<th>:: NESSTAR @ ESTA ::</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andmearhiivi tutvustus</td>
<td></td>
</tr>
<tr>
<td>Andmearhiivi statuut</td>
<td></td>
</tr>
<tr>
<td>Kontaktinfo</td>
<td></td>
</tr>
<tr>
<td>CESSDA organisatsioonid</td>
<td></td>
</tr>
<tr>
<td>ESSO - online ajakiri</td>
<td></td>
</tr>
</tbody>
</table>

TEMAATILISED ANDMEBAASID
A. Kõrgharidusuuringute andmebaas 2000-2009
   B. Eesti vanemad sotsiaaluruingud

ESTA LEPINGUD
A. Kasutamisleping
   B. Talletamisleping
   C. Taotluse vorm

MITTEELEKTROONISED MATERJALID
A. Eesti Raadio arhiivimaterjalid
   B. Tartu Ülikooli Sotsioloogiaakonna arhiivimaterjalid

Tartu Ülikoolile avatud e-ajakirjad
   ANNA INFOT TEHTUD UURINGUST
   ANNA INFOT PUBLIKATSIOONIST/ANALÜÜSIST

Estonian Social Science Data Archive webpage in English will be completed soon. Old webpage is here. Vana kodukeelu jõuga tutvumiseks vasta siia.
Croatia - CRODAS

DATA ARCHIVE FOR THE SOCIAL SCIENCES

GET DATA

USE DATA

MANAGE DATA

DEPOSIT DATA

EXPLORE OUR DATA COLLECTION

Data catalog
МАКЕДОНСКИ АРХИВ НА ИСТРАЖУВАЧКИ ПОДАТОЦИ ОД ОПШТЕСТВЕНИНАУКИ

- никое безбедно ги куваме ваши податоци -
Spain - CIS

The CIS
Surveys
Barometer
Publications
Training and grants
Actividades
Library
Press centre

INFO CIS
12.05.2017
• 09.05.2017 Advance on study results 2173
Baròmitz de abril de 2017. [view +]

Baròmetric Indicators (April 2017): Indicators of political and economic trust and confidence in the Government/opposition systems; ideological self-positioning indicator and voting intention and estimation series.

• 03.05.2017 Advance on study results 2174
Indicador de Confianza del Consumidor, Mes de abril [view +]


• 11.04.2017 The CIS is carrying out the eighth edition of the European Social Survey. If you have received a letter to your home requesting your participation in this survey, we thank you very much for your participation. You can obtain more information about the survey through the means of contact indicated in the letter and the following link.

• 23.03.2017 His Majesty the King presents the National Sociology and Political Science Prize 2014 to Dr. Emilio Lainez de Espinosa y Michels de Champourcin [view +]

OPEN CALLS
24.04.2017
• 24.04.2017 Notice (BOE 11/13/2017) for the Training scholarships for post-graduate in matters of interest for the organism for the year 2017. [view +]

• 13.03.2017 Notice (BOE 11/13/2017) for the Training scholarships for post-graduate in matters of interest for the organism for the year 2017. [view +]

ACCESES DIRECTOS
• 2016 CIS Publications Catalogue.
• Access to CIS Open Data
• Access to Survey Catalonia bank data.
• (ICC) Consumer Confidence Indicator.
• Accessing the analysis Data of Data Bank platform.

• Latest studies added to the Data Bank.

CONTACT US    LINKS    MAP    NEWSLETTER
A joint project by Institute for Social Studies of the University of Warsaw and Institute of Philosophy and Sociology of the Polish Academy of Sciences
Stawiski 5/7, 00-183 Warsaw, Poland

**Information**

**EAST PaC**

In modern democracy, voters should be able to use elections to control parties and politicians. Citizens lose control when their electoral voice does not compel parties and politicians to act according to the will of the people. Presented free and fair elections are supposed to function as a mechanism of electoral control.

To evaluate electoral control, citizens need real-time data on candidates, parties, and parliamentarians. In this edited book, we present a step towards electoral control by presenting the methodology of the East European Parliamentarian and Candidate dataset (EAST PaC). These data are the universe of candidates and parties who stood for national parliamentary elections in Ukraine, Poland, and Hungary from the 1990s to the 2010s. Candidates are matched over time, rendering a dataset that allows researchers to track the political careers of every candidate, from the thousands who never rose to the few political figures whose parliamentary careers are decades long. With EAST PaC, we can achieve new insights into electoral politics of Central and Eastern Europe.

This book contains everything that scholars need to use EAST PaC. We designed the book for regional specialists and non-specialists: for seasoned scholars of representation, accountability, and political inequality, and for those who are newly interested in these concepts. We hope to intrigue potential users of EAST PaC and to attract new scholars to study electoral politics in Ukraine, Poland, and Hungary.

This book is funded by Poland's National Science Centre (Sonata Bis decision number 2012/05/EAF/S03356)

East European Parliamentarian and Candidate data (EAST PaC) are available only via Polish Social Data Archive (ADS) and can be downloaded here: https://www.ads.org.pl/episzczep.php?id=97

The book Towards Electoral Control in Central and Eastern Europe can be downloaded here.

**New data:**

- 25.05.2016 - Zdrowie, nierówności, dezintegracja społeczna (Mieszkańcy Warszawy 2003), 216 PAN.
- 13.05.2016 - East European Parliamentarian and Candidate data (EAST PaC) 1985 - 2014, 196 PAN.
- 12.03.2014 - Zoom na Uniwersyte Tyczniki Wielkie; Towarzystwo Studentów Stawiski "C".
- 17.01.2014 - Struktura społeczna w Polsce: POLPAN 2008; Zezpold Powierzyston Analiz Nierówności Społecznych, 138 PAN.
- 23.04.2010 - Polskie Generalne Sondage Społeczne 1992-2010 ISS UW.
- 17.01.2014 - Struktura społeczna w Polsce: POLPAN 2008; Zezpold Powierzyston Analiz Nierówności Społecznych, 138 PAN.
There is currently no social sciences data archive (DAS) in Latvia. The strengths and potentials for DAS development in Latvia lie in positive attitudes of researchers towards data sharing; previous experience of a data archive; positive future plans regarding policy setting for RDM.

Finally, there is a variety of initiatives facilitating and/or providing access to social sciences research data for reuse: there are university repositories; OA repository Academia in National Library of Latvia;

(CESSDA SaW T3.2 survey, 2016, not published yet)
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