Ensuring well timed publishing of data in case of public emergency

Session: Best practices in research data lifecycle during pandemic times 1

Dr. Sonja Bezjak, Slovenian Social Science Data Archives (presenting author)
Dr. Janez Štebe, Slovenian Social Science Data Archives

ESRA 2023 Conference, Milano, Italy
"Survey research in times of crisis: Challenges, opportunities, and new directions".
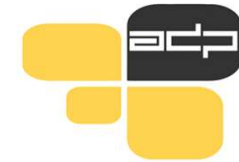
19th July 2023

Content

1. Slovenian Social Science Data Archives

2. CESSDA: Data services and support

3. Covid-19 data collections

4. Challenges in publishing research data

5. Good practice in publishing research data

Slovenian Social Science Data Archives

- **National research data centre for social sciences**

- Founded in 1997
- Located at the Faculty of social sciences, University of Ljubljana, Slovenia
- Data depositors from public and private universities, research centers
- Involved in EU and national projects
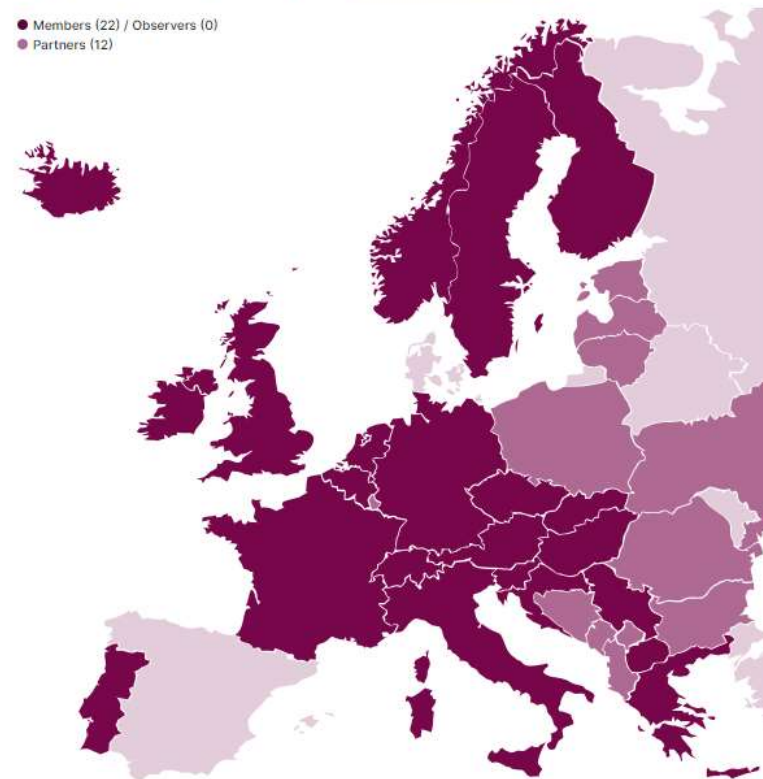- Service provider at **CESSDA ERIC**

https://www.adp.fdv.uni-lj.si/

3

Consortium of European Social Science Data Archives

"**CESSDA** provides large-scale, integrated and sustainable data services to the social sciences. **It brings together social science data archives across Europe**, with the **aim of promoting the results of social science research and supporting national and international research and cooperation**."
(https://www.cessda.eu/About)

https://www.cessda.eu/About/Consortium

To find out more, use the map below or see the list of members and partners.

- Members (22) / Observers (0)
- Partners (12)

4

Social science data to be discovered in data catalogues

CESSDA Data Catalogue

- Contains **metadata of more than 40.000 data collections** held by SPs
- Datasets from more than 20 countries
- Descriptions in English and/or local languages
- Quanti and Quali-data, Mixed-modes data
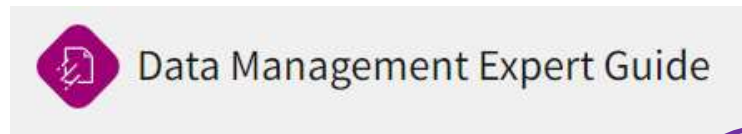- Crossectional, Longitudinal
- New, recent & historical data

cessda
DC Data Catalogue

ADP Catalogue

- Cca. **800 data sets and descriptions**
- Accompanying materials (questionnaires, reports, articles…)
- Metadata in Slovenian and English
- Different types of data
- Oldest from 1960s

Trainings for data discovery and data sharing

Training

Learn how to discover, use, manage and preserve research data

Data Management Expert Guide

https://dmeg.cessda.eu/

„The research data lifecycle is a model that illustrates the stages of data management and describes how data flow through a research project from start to finish."

(Princeton Research Data Service, https://researchdata.princeton.edu/research-lifecycle-guide/research-lifecycle-guide)

Trainings for researchers, PhD students, mentors, librarians…

PLAN
DISCOVER
ORGANISE & DOCUMENT
PUBLISH
PROCESS
PROTECT
STORE

Shifts toward open science and open data

MY DATA

MY TAX MONEY

SOCIETY

MONEY FOR RESEARCH

DATA SHARING

- EU Horizon 2020 pilot for DMP
- EU Horizon Europe mandate for DMP

Slovenia
- DMP mandatory at the Faculty of Social Sciences, University of Ljubljana
- DMP mandatory for PhD students at the University of Ljubljana

- Slovenia's new Scientific Research and Innovation Activities Act demands FAIR data

7

Experience from a real life → more training needed

Common & Challenging situations in ADP

1) I FORGOT TO ASK THE RESEARCH PARTICIPANTS FOR THEIR CONSENT TO SHARE DATA

2) I PROMISED THE PARTICIPANTS THAT I WOULD USE THE DATA EXCLUSIVELY FOR THIS PROJECT.

3) I NEED A DOI „ASAP", BUT I DON'T HAVE TIME TO TRANSCRIBE ALL 50 INTERVIEWS AND HAND THEM OVER TO THE ARCHIVES.

Activities during coronavirus pandemic

How CESSDA is supporting researchers during COVID-19
Fri 5 Jun 2020, **https://www.cessda.eu/News/CESSDA-Newsitem-nid118**

Despite the coronavirus pandemic, CESSDA and its national Service Providers strive to serve the research community and online services continue to function as usual.
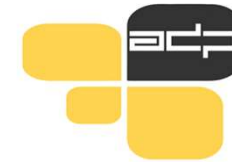
CESSDA's official **COVID-19 ambassador** has been appointed

Service providers:
- **established Covid 19 data collections**
- promoted reuse of data

https://www.cessda.eu/Covid-19

Data collections on Covid 19 at CESSDA SPs

COVID-19 activities from CESSDA Service Providers

COVID-19 resources by country are listed below.

Austria

Belgium

Czech Republic

Denmark

Finland

France

Germany

Netherlands

Portugal

Slovenia

Sweden

Switzerland

United Kingdom

Slovenia

- A webpage on data resources relevant to the pandemic has been made available on the website of ADP: Collection of COVID-19 data (in Slovenian and English).
- ADP offers researchers the opportunity to deposit their COVID-19 data via a fast-track procedure: read more on how to deposit COVID-19 data (in Slovenian and English).
- ADP created a list of data resources by topic where researchers can find variables related to pandemics: read more on the themed studies in the ADP data catalogue (in Slovenian and English).

https://www.cessda.eu/Covid-19

10

Studies on Covid 19 in Slovenia

- **Step 1: Mapping of projects and consultation with data creators**
  - From September to November 2020, we identified around 40 studies on the social aspects of Covid-19 through desktop research in the fields of sociology, communication studies, tourism, psychology, mental health, health, economics, medicine, ethnology, food studies, criminology and more.
- **Step 2: Call for contributions from lead authors to the thematic collection**
  - Around 20 researchers were interviewed individually in the form of an online meeting, during which the researcher or group was briefed on the mission of the ADP, the importance of the collection, how to collaborate, and the benefits of archiving and publishing data for researchers, the research community, and society.
- **Findings:**
  - **higher motivation to collaborate & raised awareness of the importance of sharing data**

Studies on Covid 19 in Slovenia

- 20 studies on Covid-19 published in the ADP catalogue
  - https://www.adp.fdv.uni-lj.si/eng/uporabi/covid-19/zbirka/
- Broad thematic coverage
  - the impact of covid-19 on cultural and economic life, changes in specific sectors (education, social services), diet, health habits etc.
- Different types of data
  - surveys , interview transcripts, personal diaries
- Different origin
  - Academic researchers, private research agencies and a NGO
- Sensitive data about people
  - Protection needed: anonymisation, access control and regulation

Challenges on quality

There were limitations of data collection during an epidemic

- Ad hoc projects with a small number of variables,
- Addressing narrower aspects of the topic,
- Methodological limitations (such as, non-probability samples)
- Lower levels of participation.

- No extra funding provided during the pandemic.

What we were searching for

- In the evaluation of the materials, we gave greater weight to the **relevance of the topic**,
- **completness of the documentation & related materials**,
- ensuring **secondary use as soon as possible**.

Some old challenges

- Incomplete consent forms for archiving and reusing data (GDPR)
- Lacking time and effort needed to clean data and prepare the last version.

CESSDA Training Team (2017 - 2022). *CESSDA Data Management Expert Guide.*
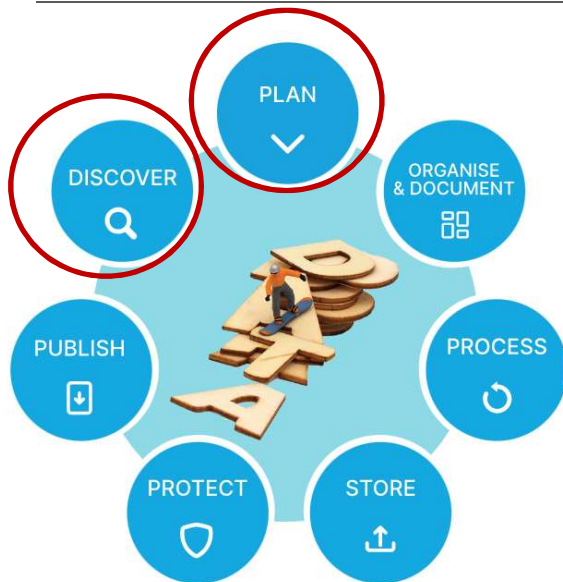Bergen, Norway: CESSDA ERIC. Retrieved from https://dmeg.cessda.eu/

Good practice

Data Publication should 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.

Costas, R., Meijer, I., Zahedi, Z. and Wouters, P. (2013). The Value of Research Data - Metrics for datasets from a cultural and technical point of view. A Knowledge Exchange Report, available from www.knowledge-exchange.info/datametrics

Good practice

Data Management Expert Guide

https://dmeg.cessda.eu/

1. Check data catalogues
2. Make a DMP
3. Contact data repository

GOOD AND TIMELY DATA MANAGEMENT PLANNING CAN BE A GUARANTEE OF DATA QUALITY.

Some new challenges for data service providers

Data on Covid-19 in CESSDA Data Catalogue
June 2022: cca. 540 studies → **June 2023: cca. 1290 studies**
(source: https://datacatalogue.cessda.eu/?q=COVID-19)

- Making data understandable across domains

CESSDA Training Team (2017 - 2022). *CESSDA Data Management Expert Guide.*
Bergen, Norway: CESSDA ERIC. Retrieved from https://dmeg.cessda.eu/

BY-COVID will integrate established national and European infrastructures including
- **ELIXIR -** is a European life sciences infrastructure for **life science data**
- **BBMRI -** BBMRI-ERIC is a European research **infrastructure for biobanking**
- **ECRIN -** **European Clinical Research** Infrastructure Network
- **PHIRI -** Research Infrastructure on **Population Health Information**
- **CESSDA**- Consortium of European **Social Science** Data Archives

Questions? Thank you

University of Ljubljana

Faculty of Social Sciences   www.adp.fdv.uni-lj.si

Kardeljeva ploščad 5   arhiv.podatkov@fdv.uni-lj.si

1000 Ljubljana   Arhiv.Druzboslovnih.Podatkov
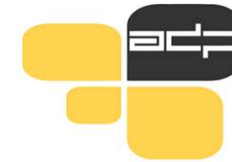
Slovenia   @ArhivPodatkov

Topics

- **<u>Survey data</u>**
    - attitudes towards SARS COV-2 vaccination among health care personnel and in general in Slovenia,
    - on change in travel habits,
    - psychological and behavioural responses to the pandemic in the immediate aftermath of the declaration of the pandemic in Slovenia and Serbia in particular,
    - international survey on well-being and attitudes towards vaccination (Slovenia, Poland, Romania),
    - among personal assistance providers on their work during Covid-19,
    - on students and distance education,
    - life in old people's homes and
    - comparative data on public reaction to the new measures in Slovenia, Croatia, Serbia and BIH.

BY-COVID PROJECT

BY-COVID

The BeYond-COVID project aims to make **COVID-19 data** accessible to everyone, scientists in laboratories, medical staff in hospitals or government officials.

Going beyond SARS-CoV-2 data, the project will provide a framework for making data from other infectious diseases **open and accessible** to everyone, and build a bridge between public health, clinical sciences, social sciences, and molecular sciences.

Subscribe to the newsletter

by-covid.org

The challenge

The world has generated vast amounts of data in response to the COVID-19 pandemic. However, this data comes from many different sources, and identifying, connecting and integrating these for effective analysis remains a challenge.

Public Health Laboratories   Hospital   Medical centres   Consortia   Research Infrastructures   Universities   Research Centres

Data from SARS-CoV-2, other pathogens, their hosts, and the associated diseases come from many different sources.

BY-COVID receives funding from the European Union's Horizon Europe Research and Innovation Programme under grant agreement number 101046203.

What we are doing

1. Mobilising data
2. Connecting data
3. Standardising data
4. Exposing and analysing data

COVID-19 Data Portal

The outcomes

Make SARS-CoV-2 and other infectious disease data easier to access, aggregate and analyse to:

- Enable **scientists in academia and industry** to respond faster to new SARS-CoV-2 strains or other pathogens causing infectious diseases.
- Help **policy-makers** assess the impact of COVID-19 and take appropriate measures to protect people.
- Facilitate **open innovation** in academia and industry across disciplines.

https://by-covid.org/about#structure

- *BeYond-COVID (BY-COVID) is a project that aims to make data on COVID-19 (SARS-CoV-2) and other infectious diseases open and accessible to everyone.*
- *It is important that data, for example genetic sequence and geographical information, is not just available to scientists in laboratories but also to anyone else who can use it such as medical staff in hospitals or government officials.*
- *By making it easy for anyone in the world to find and use the data, it will enable work to prevent and treat infectious diseases to happen more quickly and efficiently.*
- *BY-COVID will integrate established national and European infrastructures including **ELIXIR, BBMRI, ECRIN, PHIRI and CESSDA**.*