Deposit Options to Enhance the FAIR Principles

Presenter: **Kelly Ogden-Schuette**
(Authors: **John E Marcotte, Sarah Rush & Kelly Ogden-Schuette**)
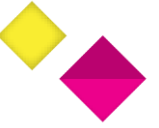*ICPSR, Institute for Social Research*
University of Michigan

ICPSR

60 years

Research Data

- Research data may contain information about individuals and organizations

- Research data are not intended for identifying individuals or organizations

- Information for producing aggregate results such as summary statistics and regression coefficients

- Aggregate results must meet disclosure protection thresholds for cell sizes for tables, sample sizes for regressions, and other specified conditions such as the suppression of certain variables or disallowed sub-samples

Multiple Sources

- Research data are sometimes available from multiple sources

- Ongoing studies with multiple waves often provide data access through the project website

- Creates challenges for projects wanting to make data findable by researchers outside of the project website

Repositories for Research Data

- Most repositories for research data have only one method of deposit: *Data and Documentation*

- These repositories make the research data available for download or through a restricted access mechanism

- Repositories must adhere to the FAIR principles

ICPSR

FAIR Principles

- **F**indable

  Data and metadata should be easy to find for both humans and computers

- **A**ccessible

  Public access through web download or controlled access through an application

- **I**nteroperable

  Data and metadata should interoperate with workflows for analysis, storage, and processing

- **R**eusable

  Data and metadata should be well-described so that they can be replicated and re-analyzed

ICPSR

ICPSR

- One of the largest repositories of research data

- Catalog of research data

- Variables database

- Search for research data by topics or variables

- Online analysis and exploration

- Bibliography

Data Sharing for Demographic Research (DSDR)

- One of the topical archives at ICPSR

- Funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development

- Focus on Maternal and Child Health, Lifecourse, International Comparisons

- DSDR has developed four deposit options

Four Deposit Options

- Standard

- Mirror

- Extensive Metadata

- Study Descriptor

Standard Deposit

- Research Data and Documentation

- Exclusive

- Public and Restricted Access

- Extras:
  - Online Exploration of Data
  - Data Guides
  - Bibliography

- ❖ Should conform to FAIR Principles

Mirror Deposit

- Non-Exclusive

- Different repositories have the same data and documentation

- One repository designated as primary, and other repositories mirror the primary

- Public access only

- ❖ Enhances: Findable, Accessible and Interoperable

ICPSR

Extensive Metadata Deposit

- Data and Documentation available from other sources

- Metadata about study and variables incorporated into multiple repository catalogs

- Ideal for ongoing studies with multiple waves where data are available from a project website

- Researchers can search and compare studies and variables

- ❖ Enhances: Findable and Interoperable

ICPSR

Study Descriptor Deposit

- Data and Documentation available from another source

- Metadata about study only incorporated into other repository catalogs

- Researchers can search and compare studies

- ❖ Enhances: Findable

Metadata Standards

- Metadata standards vary by academic discipline and type of data

- DDI-Codebook or DDI version 2 is the current ICPSR metadata schema
  - DDI: *Data Documentation Initiative*

- Other examples of standards include *CERIF, Dublin Core, DCAT, DataCite, DIF, FITS, PROV*

o Harmonizing metadata elements is an ongoing challenge

Discoverable

- Discoverable is Findable, enhanced
  - A *Digital Object Identifier (DOI)* is necessary but not sufficient

- Search by title, source, author, topic

- Search by variables and question text

- Search by data elements

- Compare search results to ascertain differences and similarities

ICPSR

Research Data Catalogs

- Research Data Catalogs are essential for making data discoverable

- Catalogs are the foundation for searches

- Archives may comprise both data repositories and catalogs or only catalogs

- ❖ Mirror, Extensive Metadata, and Study Descriptor deposits enhance the discoverability of research data by expanding where data are cataloged

Connect with DSDR

📞 (734) 763-8937

🌐 DSDR.ICPSR.UMICH.EDU

✉️ DSDR@icpsr.umich.edu

Subscribe to DSDR News for the latest updates!

ICPSR

60 years