

ADP - SOCIAL SCIENCE DATA ARCHIVES

Analyze data! Deposit study! Promote science!

FAIR social science data enhancement

Janez Štebe, UL/ADP

11TH INTERNATIONAL CONFERENCE ON SOCIAL SCIENCE
METHODOLOGY, **Naples**, Thursday, September 25th,

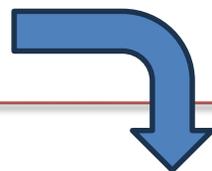
11:00-13:00 Session 12C: 8-Embracing Open Science: Methodological and
Epistemological Challenges in Research Infrastructures for the Social Sciences II



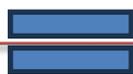
CESSDA ERIC



Increasing volume of data accessible



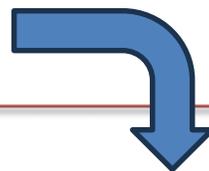
Intent: Facilitate data Re-use via secondary analysis



FAIR principles is all about increasing data reusability



One of the mission of data repositories: to support data producers in FAIR data enhancement



By providing resources and guidance in planning and processing data for increasing the reuse potential

There is room for improvement in making *data specifically* more FAIR

Examples of resources and their utilization for data FAIRification

FAIR principles references for Data (without ,meta'-)

FAIR	ID	Indicator
F1	RDA-F1-01D	Data is identified by a persistent identifier
F1	RDA-F1-02D	Data is identified by a globally unique identifier
A1	RDA-A1-02D	Data can be accessed manually (i.e. with human intervention)
A1	RDA-A1-03D	Data identifier resolves to a digital object
A1	RDA-A1-04D	Data is accessible through standardised protocol
A1	RDA-A1-05D	Data can be accessed automatically (i.e. by a computer program)
A1.1	RDA-A1.1-01D	Data is accessible through a free access protocol
A1.2	RDA-A1.2-01D	Data is accessible through an access protocol that supports authentication and authorisation
I1	RDA-I1-01D	Data uses knowledge representation expressed in standardised format
I1	RDA-I1-02D	Data uses machine-understandable knowledge representation
I2	RDA-I2-01D	Data uses FAIR-compliant vocabularies
I3	RDA-I3-01D	Data includes references to other data
I3	RDA-I3-02D	Data includes qualified references to other data
R1.3	RDA-R1.3-01D	Data complies with a community standard
R1.3	RDA-R1.3-02D	Data is expressed in compliance with a machine-understandable community standard

Research Data Alliance FAIR Data Maturity Model Working Group. 'FAIR Data Maturity Model: Specification and Guidelines', 2020. <https://doi.org/10.15497/RDA00050>.



Utilisation of standard demographics variables in new data production planing and management

Example:

- ◇ Standard occupation coding (ISCO by ILO)
- ◇ Based on it, the international ESeC and Occupational prestige score can be derived
- ◇ Shared social science concepts and operationalisations
- ◇ Classifications and historical harmonisation tables, computer code syntax, reference data and metadata sets etc.

Some resources:

- Harry Ganzeboom's Tools for deriving occupational status measures from ISCO-08 with interpretative notes to ISCO-08 <http://www.harryganzeboom.nl/isco08/index.htm>
- Hughes BT, Srivastava S, Leszko M, Condon DM (2024): Occupational Prestige: The Status Component of Socioeconomic Status. *Collabra: Psychology* 16 January 2024; 10 (1): 92882. <https://doi.org/10.1525/collabra.92882>
- GML, German Microdata Lab (2025): [MISSY - Metadata for Official Statistics](https://www.gesis.org/en/missy/materials/EU-SILC/tools/datahandling): Data Handling & Analysis: EU-SILC. GESIS. Accessed on May 2025.
- BIBB-FDZ (2025): Metadata Catalogue. The Federal Institute for Vocational Education and Training (BIBB). Accessed May 2025. <https://metadaten.bibb.de/en/>
- Van Leeuwen, MHD, 2016, "Files from HISCO database", <https://hdl.handle.net/10622/JA9B80>, IISH Data Collection, V1.

Consistent use of standard identifiers for named entities (locations, political parties, peoples...)

Examples:

- ◊ GeoNames identifiers resolving the language and political changes over time (e.g. Mezzoli, 2022)
- ◊ Measured characteristics of the named entities from parallel sources as contextual data
 - ◊ ‚Macro‘ data kept as a resource ready for integration with ‚micro‘ data (e.g. ESS 2019 and CSES 2023)
 - ◊ Cross disciplinary data integration challenging (e.g. Harrison, 2024; Gregory et. al. 2023)

References:

- Mezzoli, E (2022): Historical data of the labour force, 1945-1947: Zone A-AMG-Venezia Giulia (Allied Military Government) [Data file]. Ljubljana: UL/Slovenian Social Science Data Archives. https://doi.org/10.17898/ADP_HDS47_V1.
- CSES Comparative Study of Electoral Systems (2023): Data Bridging. Accessed on May 2025. <https://cses.org/data-download/data-bridging/>
- ESS European Social Survey (2019): Contextual information for Round 8 now available. Accessed on May 2025. <https://www.europeansocialsurvey.org/news/article/contextual-information-round-8-now-available>
- Harrison, E (2024). Climate and Air Quality Data in Attitudinal Research: A Feasibility Use-Case. Zenodo. <https://doi.org/10.5281/zenodo.10581886>
- Gregory A, Wackerow J, Orten H (2023) Reuse and Reproducibility: Describing Cross-Domain Research Data in the Science Project *Climate Neutral and Smart Cities*. ARPHA Preprints. <https://doi.org/10.3897/arphapreprints.e115047>

Existing concepts and operationalisations of key social science variables for Findability and Reuse

- ◊ Concerns about the proliferation of constructs and measures ("jingle/jungle fallacies") (NASEM, 2022).
- ◊ Difficulties in variables standardisation and data harmonisation from multiple single studies
- ◊ Challenges in reusing existing (psychological) measures in different research design settings (Ryser, 2023).
- ◊ Ontology resources (e.g. ELSST Thesaurus)
- ◊ Repositories of existing measures like European Question Bank (EQB) and REPOPSI

References and resources:

- NASEM, National Academies of Sciences, Engineering, and Medicine (2022): *Ontologies in the Behavioral Sciences: Accelerating Research and the Spread of Knowledge*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/26464> .
- Ryser, V.-A. (2023). Measuring Psychological constructs. FORS Guide No. 22, Version 1.0. Lausanne: Swiss Centre of Expertise in the Social Sciences FORS. <https://doi.org/10.24449/FG-2023-00022>
- EQB and ELSST on CESSDA Digital Tools (2025). <https://www.cessda.eu/Tools>
- FAIRsharing.org (2024b): REPOPSI; Repository of Psychological Instruments in Serbian, <https://doi.org/10.25504/FAIRsharing.d8fec9>

Conclusions

- FAIRification improves data quality and utility
- Data transparency fosters cumulative knowledge production, and other
- Motivating and training data producers for engaging in deeper annotation and use of resources
- Enabling research community and infrastructure providers efforts for sustaining such resources



Thank you!

University of Ljubljana
Faculty of Social Sciences
Social Science Data Archive
Kardeljeva ploščad 5
1000 Ljubljana
Slovenia



www.adp.fdv.uni-lj.si



arhiv.podatkov@fdv.uni-lj.si



[Arhiv.Druzboslovnih.Podatkov](https://www.facebook.com/Arhiv.Druzboslovnih.Podatkov)



[@ArhivPodatkov](https://twitter.com/ArhivPodatkov)

