

MACHINE ACTIONABLE INTEGRATION OF DATACITE AND DDI METADATA

Wolfgang Zenk-Möltgen, GESIS – Leibniz Institute for the Social Sciences

Presentation at EDDI14 – 6th Annual European DDI User Conference

at IOE - Institute of Education, University of London

December 2-3, 2014

OUTLINE

- Documentation of research data at the GESIS Data Archive
 - Metadata in DDI-C and DataCite format
 - DOI persistent identifiers for research data
- Usage of DOIs for data citations
- Machine actionable processing of metadata
 - IsMetadataFor and HasMetadata properties in the DataCite metadata schema
 - Example implementation
- Lessons learned and other possible solutions

DOCUMENTATION OF RESEARCH DATA AT THE GESIS DATA ARCHIVE

- Documentation for the research data from the social sciences at the GESIS Data Archive is being done on several levels:
 - The Data Catalogue DBK has the main level of information for finding appropriate studies
 - Further documentation includes methodological reports, questionnaire and dataset documentation, English and German translations
 - The DBK metadata is compatible with DDI-C and DDI-L and can be exported
 - The DBK metadata is (mostly) compatible with DataCite and is transferred to da|ra when DOIs are registered
 - Further documentation is only available for parts of the collection which are documented in depth (with a Codebook). This documentation is produced internally in DDI-C compatible form, but distributed to end-users only in PDF format

USAGE OF DOIS FOR DATA CITATIONS

- To enable users of datasets from the archive to cite the data:
 - Every dataset for a study gets a DOI
 - Each new version of the dataset gets a new DOI
- Metadata for the study is transferred to da|ra, and from there to DataCite
- Users can find a pre-formatted citation for the dataset at each of the systems
- They can copy&paste the citation,
- But they can also read and change it according to their citation style

MACHINE ACTIONABLE PROCESSING OF METADATA

- To enable computers to process the study level metadata
 - DBK provides a webservice to access the study level metadata in DDI format
 - DataCite provides content negotiation for the metadata of the DOI record
- To access further metadata, so-called „rich metadata“, that may be kept in other files, a link to it is provided by the DataCite metadata schema:
 - relatedIdentifier of relationType „HasMetadata“
 - relatedIdentifier of relationType „IsMetadataFor“ (to link back)
 - Additional information for the related identifier:
 - relatedIdentifierType
 - relatedMetadataScheme, schemeURI, schemeType

EXAMPLE IMPLEMENTATION

- European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)
- DBK metadata

ZA4804: European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)

Bibliographic Citation	Content	Methodology	Data & Documents	Errata & Versions
Further Remarks	Publications	Groups		

Citation ⓘ

Study No.

Title

Current Version

Date of Collection

Principal Investigator/
Authoring Entity,
Institution

EVS (2011): European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008). GESIS Data Archive, Cologne. ZA4804 Data file Version 2.0.0, doi:10.4232/1.11005

ZA4804

European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)

2.0.0, 2011-12-30, doi:10.4232/1.11005

01.03.1981 - 15.03.2010

1

2

3

4

5

6

7

8

9

10

>>

▾ Gedeshi, Ilir - Center for Economic and Social Studies, Albania

▾ Zulehner, Paul M. - University of Vienna, Austria

▾ Faradov, Tair - International Center for Social Research (ICSR), Azerbaijan

▾ Rotman, David - Belarus State University, Belarus

▾ Titarenko, Larissa - Belarus State University, Belarus

▾ Billiet, Jaak - Catholic University of Leuven, Belgium

▾ Dobbelaere, Karel - Catholic University of Leuven, Belgium

▾ Kerkhofs, Jan - Catholic University of Leuven, Belgium

▾ Swyngedouw, Marc - Catholic University of Leuven, Belgium

▾ Voyé, Liliane - Catholic University of Leuven, Belgium

Categories ⓘ

▾ Society, Culture

Religion and Weltanschauung

EXAMPLE IMPLEMENTATION

- European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)
- da|ra metadata

dataset	European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)		gesis
DOI	10.4232/1.11005		
Version	2.0.0		
Resource Type	Dataset		
Creator	• EVS		
Publication Date	2011-12-30		
Contributor	• Gedeshi, Ilir (Center for Economic and Social Studies, Albania) (Researcher) • Zulehner, Paul M. (University of Vienna, Austria) (Researcher) ... see more		
Language	English		
Classification	• ZA: ◦ Society, Culture ◦ Religion and "Weltanschauung"		
Description	• Abstract Moral, religious, societal, political, work, and family values of Europeans. Compilation of the data sets from 1981, 1990, 1999, ... see more		
Temporal Coverage	• 1981-03-01 - 1981-05-31 Belgium (1st wave) • 1982 Canada (1st wave) • 1981-03-01 - 1981-05-31 Denmark (1st wave) • 1981-03-01 - 1981-05-31 F ... see more		
Geographic Coverage	• Albania (AL) Albania (AL)		

EXAMPLE IMPLEMENTATION

- European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)
- DataCite metadata



doi:10.4232/1.11005

This page represents DataCite's metadata for *doi:10.4232/1.11005*.

For a landing page of this dataset please follow <http://dx.doi.org/10.4232/1.11005>

Citation

EVS; (2011): European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008); GESIS Data Archive. <http://dx.doi.org/10.4232/1.11005>  

Descriptions

Abstract

Moralische, religiöse, gesellschaftliche, politische, ökonomische und soziale Wertvorstellungen der Europäer. Der Online Variable Overview ermöglicht den Vergleich von Trend Variablen über die vier EVS Wellen 1981, 1990, 1999 und 2008 sowie einen Vergleich von originalsprachlichen Fragen aus den Wellen 1999 und 2008. Online Variable Overview Themen: Einige Fragen wurden nicht in allen Ländern und zu allen Befragungszeiträumen gestellt. Freizeit: Wichtigkeit der Lebensbereiche Familie, Freunde, Freizeit, Politik, Arbeit, Religion (nur in Schweden: Einsatz für andere); Glücksgefühl; subjektiver Gesundheitszustand; Gefühle von: Aufregung oder Interesse, Rastlosigkeit, Stolz aufgrund von Komplimenten, Einsamkeit, Freude über Vollendung einer Sache, Langeweile, Wohlfühlen, depressiv oder unglücklich, alles gelingt mir, Traurigkeit aufgrund von Kritik; Gefühlslage des Befragten zu Hause: Entspannung, Angst, Zufriedenheit, Aggression oder Sicherheit. Respekt und Liebe für die Eltern;

EXAMPLE IMPLEMENTATION

- European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)
- DataCite content negotiation: different representations of the metadata

Other formats

[text/html](#)

[application/x-datacite+xml](#)

[application/vnd.datacite.datacite+xml](#)

[application/x-datacite+text](#)

[application/vnd.datacite.datacite+text](#)

[application/rdf+xml](#)

[text/turtle](#)

[application/x-bibtex](#)

[application/x-research-info](#)

[application/citeproc+json](#)

[application/vnd.citationstyle](#)

[text/x-bibliography](#)

```
- <rdf:RDF>
- <rdf:Description rdf:about="http://dx.doi.org/10.4232/1.11005">
  <j.0:creator>EVS</j.0:creator>
  <j.0:publisher>GESIS Data Archive</j.0:publisher>
- <j.0:title>
  European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)
  </j.0:title>
  <j.0:date>2011</j.0:date>
  <owl:sameAs>doi:10.4232/1.11005</owl:sameAs>
  <owl:sameAs>info:doi/10.4232/1.11005</owl:sameAs>
  <j.0:identifier>10.4232/1.11005</j.0:identifier>
</rdf:Description>
</rdf:RDF>
```

EXAMPLE IMPLEMENTATION

- Rich metadata: DDI-C Codebook for “European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)”
- DBK

ZA4804: European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)

Bibliographic Citation	Content	Methodology	Data & Documents	Errata & Versions
Further Remarks	Publications	Groups		

Dataset

Number of Units: 166.502

Number of Variables: 865

Analysis System(s): SPSS

[Availability](#) ⓘ

A - Data and documents are released for academic research and teaching.

Download of [Data and Documents](#) ⓘ

All downloads from this catalogue are free of charge. Data-sets available under access categories B and C must be ordered via the shopping cart. [Charges apply!](#)

↓

Datasets	Codebooks	Other Documents	DDI Documents
----------	-----------	-----------------	---------------

⇓

- [ZA4804_EVS-WVS_ParticipatingCountries.pdf](#) List of participating countries (EVS and WVS) (Table) 102 KBytes
- [ZA4804_EVS-WVS_ParticipatingCountries.xls](#) List of participating countries (EVS and WVS) (Table) 221 KBytes
- [ZA4804_EVS_1981-2008_Overview.xlsx](#) Method Overview - National Surveys (Table) 137 KBytes
- [ZA4804_EVS_ParticipatingCountries.pdf](#) List of participating countries (EVS only) (Table) 45 KBytes
- [ZA4804_EVS_VariableCorrespondence.pdf](#) variable correspondence list (EVS only) (Table) 429 KBytes
- [ZA4804_EVS_VariableCorrespondence.xlsx](#) variable correspondence list (EVS only) (Table) 92 KBytes
- [ZA4804_v2-0-0_ddi20_Codebook_20140901.xml](#) (DDI-C Codebook) 11 MBytes
- [ZA4804_ValuesSurveys_DataFile_Information.pdf](#) Information on the merged

EXAMPLE IMPLEMENTATION

- Rich metadata: DDI-C Codebook for “European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)”
- DBKEdit: Create DOI for DDI-C document

DBKEdit 1.8 stable Menu

Edit DOI for document

Dokument mit ID 54500 zu Studie 4804

Primärmaterial-Datei

[ZA4804_v2-0-0_ddi20_Codebook_20140901.xml](#) DOI erzeugen: 10.4232/2.4804.54500
<http://dbk.gesis.org/dbksearch/download.asp?id=54500>

Version: ☒ registriert

Versionsdatum: (TT.MM.JJJJ)

[zur Studienbeschreibung](#)

DBKEdit 1.8 stable © GESIS Datenservice DAS 28.11.2014

Angemeldet als Moeltgen. Session alive

EXAMPLE IMPLEMENTATION

- Rich metadata: DDI-C Codebook for “European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)”
- da|ra

Availability	 Download A - Data and documents are released for academic research and teaching.
Rights	All metadata from GESIS DBK are available free of restriction under the Creative Commons CC0 1.0 Universal Public Domain Dedication. However ... see more
Alternative Identifiers	• ZA4804 (Type: ZA-No.)
Related Identifiers	• Has metadata DOI: 10.4232/2.4804.54500 • Is new version of DOI: 10.4232/1.4804
Publications	• The European Values Study: A Third Wave. Source book of the 1999/2000 European Values Study Surveys. Halman, Loek, Tilburg: EVS, WORC, Til ... see more
Update Metadata: 2014-09-09 Issue Number: 171 Registration Date: 2012-03-01	

text

DDI-C Codebook for Study ZA4804

gesis

DOI	10.4232/2.4804.54500
Version	1.0.0
Resource Type	Text : DDI-C Codebook
Creator	• GESIS - Data Archive for the Social Sciences
Publication Date	2014-09-01
Language	English
Data and File Information	• File Name: ZA4804_v2-0-0_ddi20_Codebook_20140901.xml
Availability	 Download
Rights	Creative Commons Attribution-ShareAlike 4.0 International
Related Identifiers	• Is metadata for DOI: 10.4232/1.11005

Update Metadata: 2014-09-03 | Issue Number: 1 | Registration Date: 2014-09-03

EXAMPLE IMPLEMENTATION

- Rich metadata: DDI-C Codebook for “European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)”
- DataCite



doi:10.4232/2.4804.54500

This page represents DataCite's metadata for *doi:10.4232/2.4804.54500*.

For a landing page of this dataset please follow <http://dx.doi.org/10.4232/2.4804.54500>

Citation GESIS - Data Archive for the Social Sciences; (2014): DDI-C Codebook for Study ZA4804; GESIS Data Archive.
<http://dx.doi.org/10.4232/2.4804.54500> **RIS** **BIBTeX**

Resource type

Text DDI-C CodebuchDDI-C Codebook

Rights Creative Commons Namensnennung-Weitergabe unter gleichen Bedingungen 4.0 international
Creative Commons Attribution-ShareAlike 4.0 International

Size

Language eng

Version 1.0.0

Related identifiers

IsMetadataFor doi:[10.4232/1.11005](https://doi.org/10.4232/1.11005)

Other formats

[text/html](#)

EXAMPLE IMPLEMENTATION

- European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)
- DataCite metadata

```
<relatedIdentifier
  relatedIdentifierType='DOI'
  relationType='HasMetadata'
  relatedMetadataScheme='DDI-C'
  schemeURI='http://www.ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/codebook.xsd'
  schemeType='XSD'
>10.4232/2.4804.54500</relatedIdentifier>
```

- Rich metadata: DDI-C Codebook for “European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)”
- DataCite metadata

```
<relatedIdentifier
  relatedIdentifierType='DOI'
  relationType='IsMetadataFor'
  relatedMetadataScheme='DDI-C'
  schemeURI='http://www.ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/codebook.xsd'
  schemeType='XSD'
>10.4232/1.11005</relatedIdentifier>
```

EXAMPLE IMPLEMENTATION

- Landing page for study: [10.4232/1.11005](https://doi.org/10.4232/1.11005)

European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008) **gesis** Citation

DOI	10.4232/1.11005
Version	2.0.0
Resource Type	Dataset
Creator	• EVS
Publication Date	2011-12-30
Contributor	• Gedeshi, Ilir • Zulehner, J. ... see more
Language	English
Classification	• ZA: ◦ Soc ◦ Rel
Description	• Abstract Moral, Europe 1999, ...
Temporal Coverage	• 1981-0 Belgium • 1982 Canada • 1981-0 Denmark 1994-0

GESIS - Leibniz Institute for ...

<https://dbk.gesis.org/dbksearch/sdesc2.asp?no=48048> W - Wikipedia (de)

gesis | Contact | Email to Data Service

GESIS Data Catalogue 2.1

Search Browse Overview News About

Login Shopping Cart

Groups
Products
Studies
Geographic Codes

Member of the **Leibniz** Association

da|ra Registration agency for social and economic data

ZA4804: European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)

Bibliographic Citation	Content	Methodology	Data & Documents	Errata & Versions
Further Remarks	Publications	Groups		

You have requested the DOI for the current version!

Citation

EVS (2011): European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008). GESIS Data Archive, Cologne. ZA4804 Data file Version 2.0.0, doi:10.4232/1.11005

Study No. ZA4804

Title European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)

Current Version 2.0.0, 2011-12-30, doi:10.4232/1.11005

Date of Collection 01.03.1981 - 15.03.2010

Principal Investigator/ Authoring Entity, Institution

1 2 3 4 5 6 7 8 9 10 >>

Gedeshi, Ilir - Center for Economic and Social Studies, Albania

EXAMPLE IMPLEMENTATION

- To programmatically access the rich metadata contained in the DDI-C file, we need direct access to the object
- No landing page, but direct access to rich metadata: [10.4232/2.4804.54500](http://dx.doi.org/10.4232/2.4804.54500)

doi:10.4232/2.4804.54500

This page represents DataCite's metadata for *doi:10.4232/2.4804.54500*.

For a landing page of this dataset please follow <http://dx.doi.org/10.4232/2.4804.54500>

Citation

GESIS - Data Archive for the Social S

<http://dx.doi.org/10.4232/2.4804.54500>

Resource type

Text

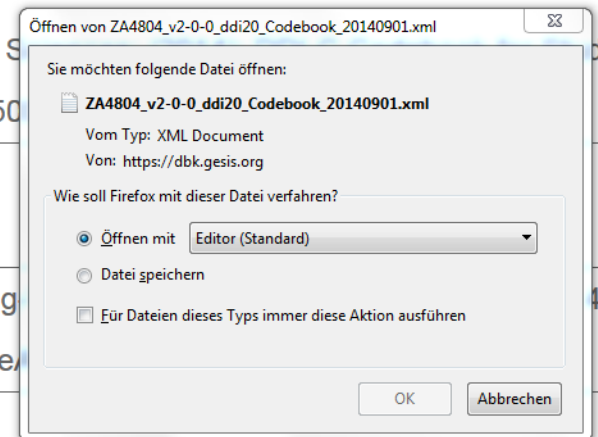
DDI-C CodebuchDDI-C Codebook

Rights

Creative Commons Namensnennung

Creative Commons Attribution-ShareAlike

Size



LESSONS LEARNED

- Current implementation supports
 - Linking from study to rich metadata
 - Machine actionable processing of the metadata and rich metadata
 - Linking back from rich metadata to metadata
- Current implementation lacks
 - Machine actionable processing of the back link from rich metadata to metadata
 - Adherence to policy of using landing pages for DOI resolution

OTHER POSSIBLE SOLUTIONS

- The use of content negotiation
 - implies to have the rich metadata „inside“ the metadata (so it can be sent to the processing machine)
 - Would need explicit support from DataCite in a new metadata schema and for the rich metadata type to be returned
 - Or, a generic content type like „directAccess“ would be supported
- The use of URL instead of DOI for the link between metadata and rich metadata
 - Works only when linking items from the same provider (who has control over changing URLs)

```
<relatedIdentifier
  relatedIdentifierType='URL'
  relationType='HasMetadata'
  relatedMetadataScheme='DDI-C'
  schemeURI='http://www.ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/codebook.xsd'
  schemeType='XSD'
>http://dbk.gesis.org/dbksearch/download.asp?id=54500</relatedIdentifier>

<relatedIdentifier
  relatedIdentifierType='URL'
  relationType='IsMetadataFor'
  relatedMetadataScheme='DDI-C'
  schemeURI='http://www.ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/codebook.xsd'
  schemeType='XSD'
>https://dbk.gesis.org/dbksearch/sdesc2.asp?no=4804&db=e&doi=10.4232/1.11005</relatedIdentifier>
```

THANK YOU!

Questions or comments?

Contact:

wolfgang.zenk-moeltgen@gesis.org

Acknowledgement: Joachim Wackerow: Ways for a Machine-actionable Processing Chain for Identifier, Metadata, and Data. Presentation at the Workshop on Metadata and Persistent Identifiers for Social and Economic Data. 07-08/05/2012, Berlin, http://www.ratswd.de/ver/docs_PID_2012/Wackerow_PID2012.pdf