

# The Next Generation of the Microdata Information System MISSY: An Integrated Solution for the Documentation of European Microdata

European DDI User Conference, Paris 2013

Jeanette Bohr, Alexander Mack, Thomas Bosch & Florian Thirolf  
GESIS – Leibniz Institute for the Social Sciences, Mannheim

# Overview

- MISSY - Metadata for scientific research
- Development as a DDI use case
- MISSY III Project: Documentation of EU Data
- MISSY & Data without Boundaries
- The new MISSY-Editor

# Scientific service for official microdata

- Increasing accessibility of individual data from official statistics for scientific research
  - Scientific support for the use of official microdata is required
- ➡ Microdata Information System MISSY:  
provides structured metadata for official statistics

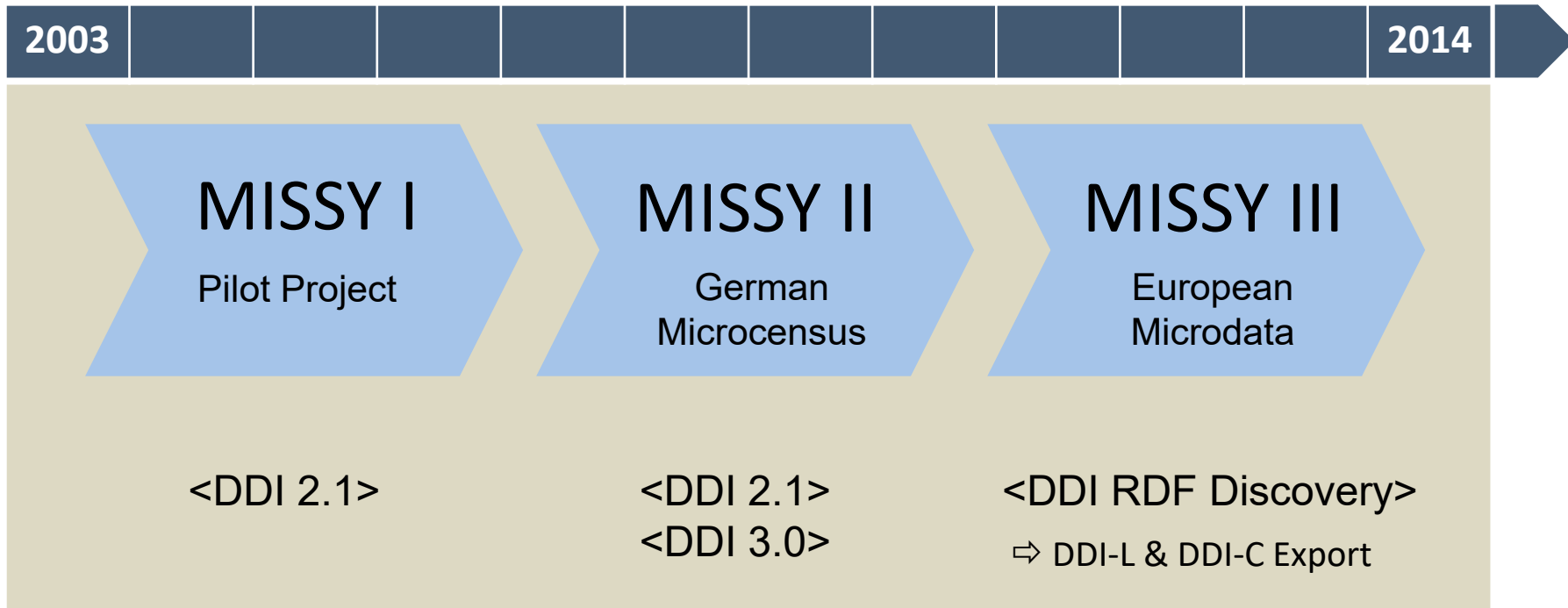
# MISSY - Metadata for scientific research

- Facilitates professional and efficient use of official microdata for the social sciences
- Contains detailed metadata for all available German Microcensus Scientific Use Files (since 1973)
- Documentation for integrated European microdata from official statistics is underway
- ➔ in cooperation with Data without Boundaries Project (WP5, Task 3)

# MISSY – Scope of metadata

- Detailed information on variable level
- Comparability of variables over time
- Study description and data collection materials
- Concepts, classifications and definitions
- Setups and program routines ('microdata tools')
- Literature database

# Development as a DDI use case



## Project MISSY III

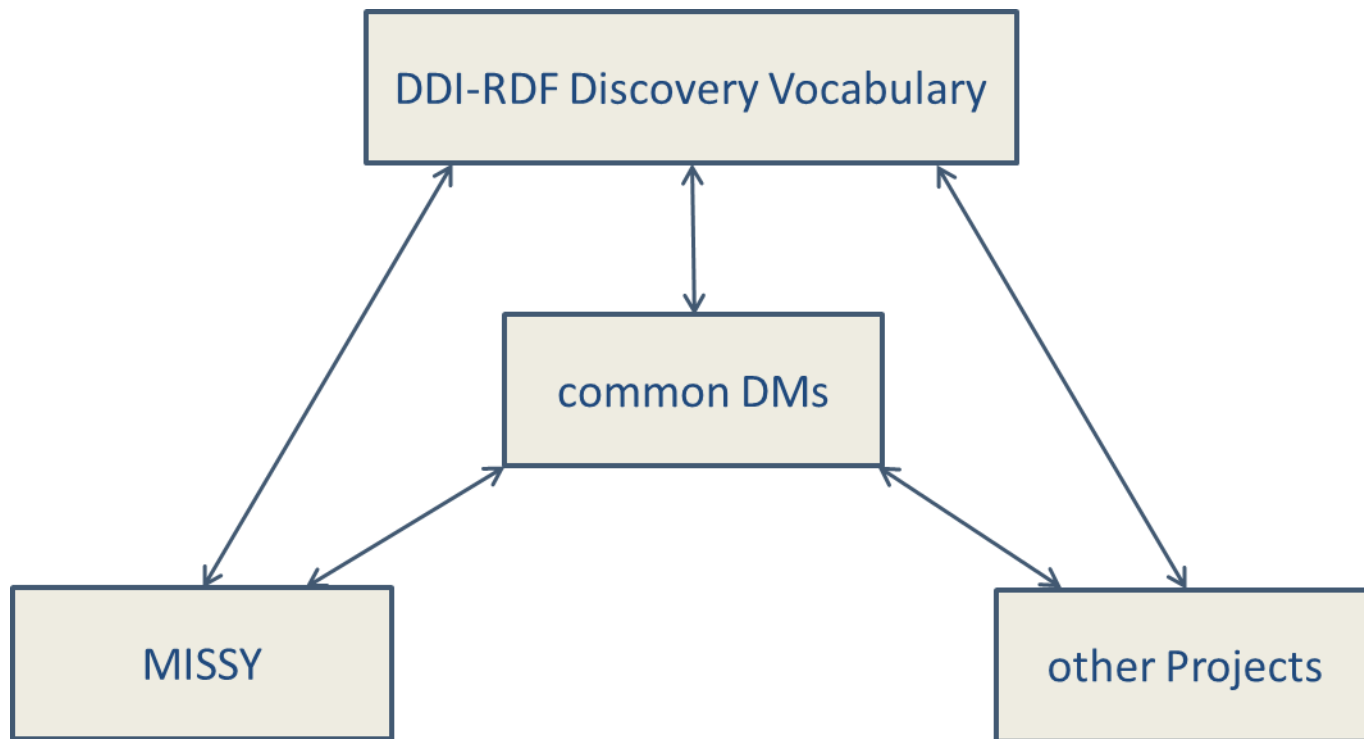
- Central, systematic and user-friendly access to information about official EU microdata via MISSY
- Potential to conduct thematic searches over multiple studies
- Support for cross-cultural and inter-temporal analysis
- Re-use of metadata in other systems (export to multiple formats such as DDI-L, DDI-C, DDI-RDF)

## MISSY III – Back-End

- Implementation of the new MISSY-Editor (Web-Application)
- Metadata import (from SPSS, XML, .csv)
- Documentation of EU microdata via MISSY-Editor
- Open source software  
<https://github.com/missy-project/>



# MISSY III – Data model architecture



## Cooperation DwB - MISSY

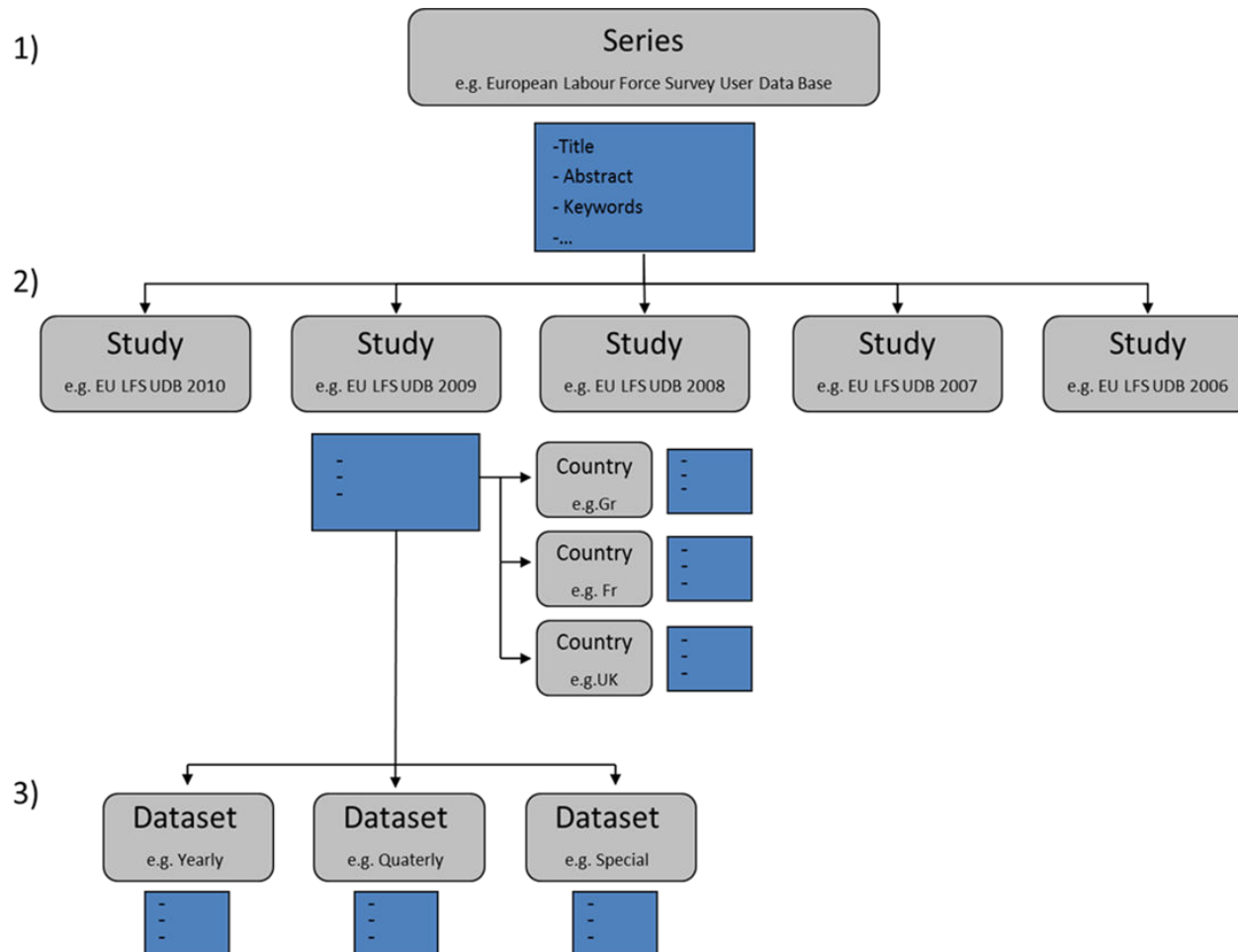


- Metadata schema and controlled vocabularies (WP5)
- Data documentation of SILC and LFS covered by GESIS; AES, SES and CIS by DwB partners (WP5)
- The new web-based MISSY III-Editor was implemented as a web application in order to allow external users
- Data export from MISSY into DwB Portal in DDI 3.2 format (WP8/12)

# Steps in Metadata Preparation

- Read variable information into system from SPSS files
- Customize data input forms and controlled vocabularies for each series
- Structure variables according to thematic classification
- Manually input and/or import metadata at variable and study level

# Metadata Scheme – Study level



Study

Series

MZ

EU-LFS

EU-SILC

Study

2009

Change Study

Series

Study: General Description

Study: Country Specific Information

Dataset

Series: EU-SILC



Data Version

Release Year

?

Version Number

?

Title

?

Subtitle

?

Abstract

?

Data Publisher

?

Keywords

Selected Keywords



?

Geograph.Cov.

?

Notes

?

Organization

?

Universe

?

Sampling

?

## Study

### Series

MZ

EU-LFS

EU-SILC

### Study

2009



Change Study

Series

Study: General Description

Study: Country Specific Information

Dataset

## Study: EU-SILC / 2009 / General Description



Title

?

Subtitle

?

Abstract

?

### Keywords

Selected Keywords



?

Geographical Coverage

?

Time Period covered

?

Notes

?

Citation

?



## Study

## Series

MZ

EU-LFS

EU-SILC

## Study

2009



Change Study

## Series

## Study: General Description

## Study: Country Specific Information

## Dataset

## Study: EU-SILC / 2009



AT

BE

BG

CH

CY

CZ

DE

DK

EE

ES

FI

FR

GR

HU

IE

IR

IT

LV

LT

LU

MT

NL

NO

PL

PT

RO

SE

SI

SK

UK

Corresponding national Study

?

Producer

?

Universe

?

## Target Sample Size (cross-sectional) ?

Sampling Units

Households



?

Minimum Effective Sample Size

0

?

Actual Sample Size

0

?

Achieved Sample Size

0

?

Size of Design Effect

0

?

Effective Sample Size

0

?

## Target Sample Size (longitudinal) ?

Sampling Units

Households



?

Minimum Effective Sample Size

0

?

Actual Sample Size

0

?

Achieved Sample Size

0

?

## Study

### Change dataset

Select a series

Microcensus

EU-SILC

EU-LFS

EVS

Select a study



Change Study

Series

Study: General Description

Study: Country Specific Information

Datasets

### Dataset: EU-SILC / 2009 / Panel



Panel

Add Dataset

Title

?



Weighting

?



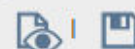
Notes

?



Citation

?





Study

Variables

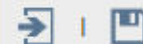
## Change variable

Pos.	Name
0	PB010
1	PB020
2	PB030
3	PB040
5	PB060
7	PB100
9	PB110
11	PB120
13	PB130
15	PB140
17	PB150
19	PB160
21	PB170
23	PB180
25	<b>PB190</b>
27	PB200
29	PB210
31	PB220A
33	PE010

Variable Details

Values / Frequencies

## EU-SILC / 2009 / Cross-sectional / Personal Data / PB190



Variable name PB190 ?

Variable label Marital status ?

Classification ?

Reference Period ?

Description Target Variable ?

Country Specific Comments ?

Other Comments ?

Thematic Classification Marital status ?

Unit of Observation ☐ Household ☒ Person ?

Filter ?

Is ad-hoc module variable ?

Is derived variable type ?

## Question Text

English

Question Wording ?

Comment ?

Study

Variables

## Change variable

Pos.	Name
0	PB010
1	PB020
2	PB030
3	PB040
5	PB060
7	PB100
9	PB110
11	PB120
13	PB130
15	PB140
17	PB150
19	PB160
21	PB170
23	PB180
25	<b>PB190</b>
27	PB200
29	PB210
31	PB220A
33	PE010

Variable Details

Values / Frequencies

## EU-SILC / 2009 / Cross-sectional / Personal Data / PB190



AT	BE	BG	CH	CY	CZ	DE	DK	EE	ES	FI	FR	GR	HU	IE	IR	IT
LV	LT	LU	MT	NL	NO	PL	PT	RO	SE	SI	SK	UK				

ALL

## PB190: Marital status

Current variable Frequencies

## Statistics

Total (Missing) 471937 (3208)

Value	Value Label	Frequency	Total %	Valid %
1	never married	136149	28.7	28.8
2	married	265000	55.8	56.2
3	separated	5157	1.1	1.1
4	widowed	39088	8.2	8.3
5	divorced	26543	5.6	5.6
(M)	Missing	3208	0.7	
	Valid Total	471937	99.3	100
	Total	475145	100	

## PB190\_F: Flag

Current variable Frequencies

for more information...

please visit our

# Poster Presentation

with live demo of the new MISSY Editor

[www.gesis.org/missy](http://www.gesis.org/missy)

[www.dwbproject.org](http://www.dwbproject.org)