

# DDI3 and Re-use of Metadata in Archival Processes

Uwe Jensen, GESIS

1st. EDDI 4 Dec. 2009 - IZA – Bonn, Germany

## **Background** - Work on DDI3 is one topic of WP8 - CESSDA PPP

“Preparatory Phase Project for a Major Upgrade of the Council of European Social Science Data Archives (CESSDA) Research Infrastructure”

## **Relevance**

- DDI 3 provides new features to support consequent re-use of metadata along the study & data life-cycle.
- Great interest for data archives to apply these potentials in future workflows for data documentation & managing collections of complex survey data.

## **Major focus of the presentation: Grouping & Resource Package**

- Two test cases to explore & understand principles in handling large complex studies (top down <> bottom up country perspective by FSD)
- It is work in progress in cooperation with DDI experts

# 1. Basic options to re-use of metadata in DDI 3

## Group > inheritance mechanism

- ❖ Structure two or more studies > series (by design) or “ad-hoc” ....
- Commonalities > upper level (Group > Subgr. > Study Unit) inherited down the hierarchy as long as ...
- Differences > ADD or DELETE > lower level

## Resource Package > reference mechanism

- ❖ schemes that may be used outside of a Group structure by single Studies or grouped studies
- Rational: they are independent of a study (or related groups)

## 2. Which modules to use for what purpose?

- Re-use of metadata requires
  - overview on the scope of studies in mind
  - & decisions on how to handle or organise respective metadata
- Several solutions / paths possible in DDI 3 to structure metadata
- What life-cycle event and or process is considered
  - Data processing ..., metadata documentation ..., exchange ..., preservation ...
  - publish > Browse or Search by year, topic, languages, country ...
  - tools design ...
- What are the Group / Resource Package / + comparison able to do?
- How to? in applying Group / Resource Package / + comparison

## TEST CASE 1: Group 10 topical modules ISSP

- 23 temporal instances (1985-2007), 522 country samples, 44 different countries,

### Subgroup - Environment 2000 / 1993

- What is **module dependent** > Concepts
- What is **module instance independent** > ...
- What is **study independent** > Question scheme

### What is **instance dependent & study independent**

> Question scheme (Q1; Q16 not asked)

### What is **study dependent** (e.g. deviations)

Country 1

- Sample design / Country variables
- Country questionnaire / language

### What is **instance dependent & study independent**

> Add > geographical coverage  
> ....

### What is **study dependent** (e.g. deviations)

Country 1

- Sample design / Country variables
- Country questionnaire / language

### Subgroup Module Environment

Conceptual Component .....  
Data Collection > Q.1 to Q.20  
Logical Product > V.1 to V.40

### Subgroup Environment 1993

Data Collection ....  
Logical Product ....

Study Unit Country 1  
...

Study Unit Country 6

### Subgroup Environment 2000

Data Collection ...  
Logical Product ...

Study Unit Country 1  
....

Study Unit Country 15

### Sg M2

...

Subgroup  
Year ...

Subgroup  
Year ...

Subgroup  
Year ...

### Sg M3

...

Subgroup  
Year ...

### Sg M4

...

Subgroup  
Year ...

Subgroup  
Year ...

### Sg M...

...

Subgroup  
Year ...

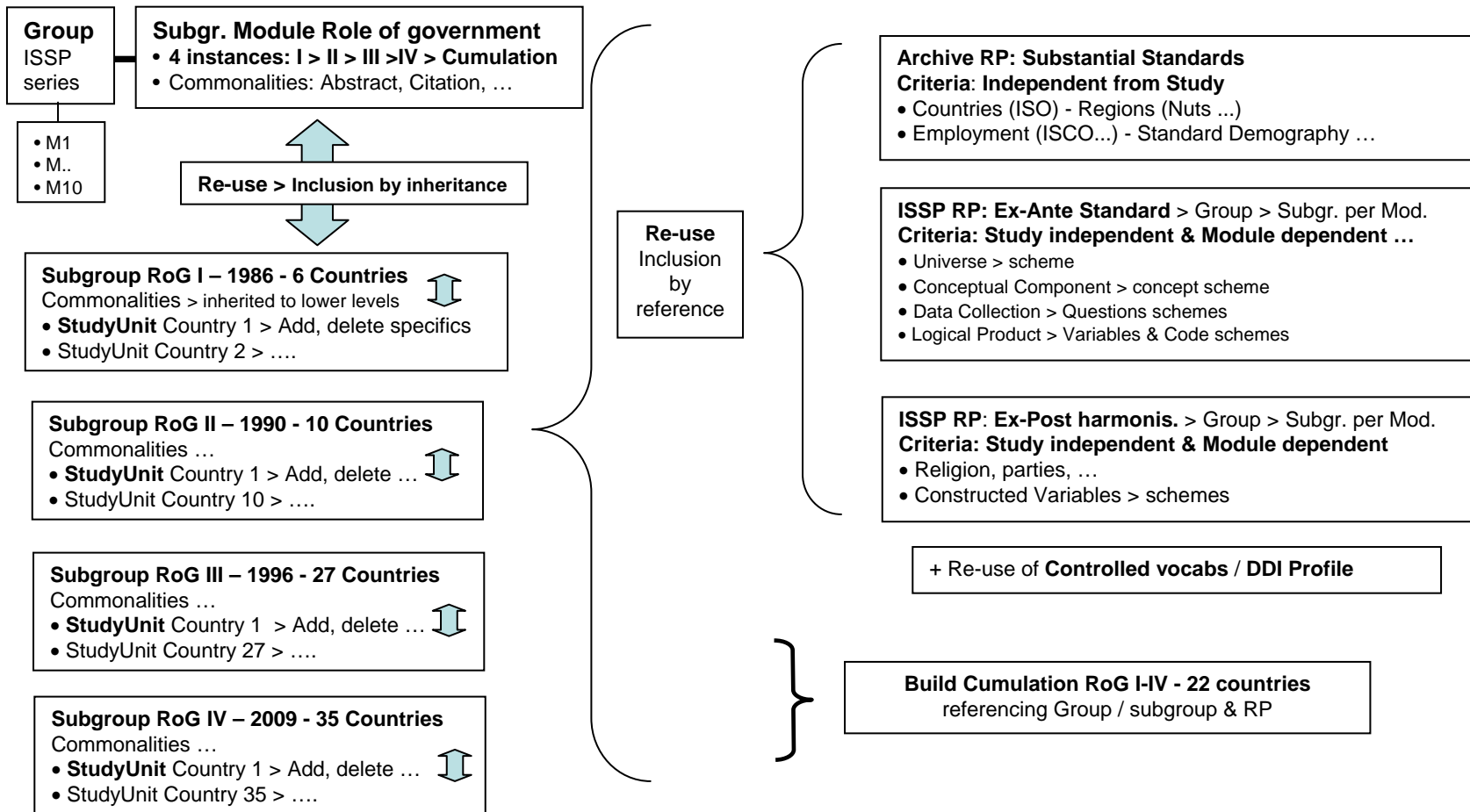
### Commonalities for ISSP as a whole

#### What is **independent from modules + studies**?

- Series Name
- Series description
- Concept (> Conceptual Component)
- Data Collection > .....
- Logical Product > .....
- Archive > integrating instance / country instance



## Test Case 2: Advanced re-use combining **Group & Resource Package (RP)** - Comparison (not yet considered) – work in progress



# Summary I: Metadata re-use with DDI 3 modules

## 1. Group allows

- to organise dependent studies (comparative by design; ad-hoc by topics, ....)  
to capture metadata one time (group > subgroup > study unit)

## 2. Resource Package advance

- Re-use Study independent metadata (agency internal re-use) or
- To publish metadata for public re-use (e.g. Demography / > 3CDB / QDB)

## 3. Group & RP & Comparison combination foster

- economic metadata handling & provision of advanced content
  - Reuse of schemes for Questions, Variables, Codes etc. for comparative studies ...
  - Standardised documentation of changes over time ...
  - Extended indication of comparable metadata ...
- **Example: Future of DDI at GESIS & ICPSR - IASSIST 2009 [presentation](#)**

## Summary II: Metadata Re-Use & Lifecycle

Applying the Group along the lifecycle would foster increased re-use of metadata for archives and beyond:

1. **Provision of ex-ante** defined **standard** (questions / variables) to prepare the fieldwork (Research team > fieldwork agency)
2. **Transfer** when the fieldwork / documentation is finished at one place (> from data producer to data provider)
3. **Publication & presentation** of a completed documentation with a fix (dedicated) metadata structure
  - Publish at a public portal > for Browsing in grouped structure
  - Search for comparable items based on metadata from CM



Thank  
you very much  
for your attention

Final paper will be published from the DDI Alliance

Access CESSDA-PPP information  
> the final slide

## Information on CESSDA – PPP

Preparatory Phase Project for a Major Upgrade of the Council of European Social Science Data Archives (CESSDA) Research Infrastructure

Funded by EU 7th. Framework programme

Duration 1.2008 - 12.2009 (most WPs)

<http://www.cessda.org/project/>

**Tasks & Work package**

**Outputs & Publications**

Completed reports, recommendations,  
presentations and other public documents

like tendered Report:

[Technical Specifications for a European Question Data Bank](#)

and some more ....