Documenting, Maintaining, and Sharing Standard Variables with DDI Version 3.0: the ISCO example

Joachim Wackerow
GESIS-ZUMA
German Social Science Infrastructure Services / Centre for Survey Research and Methodology

IASSIST 2007, May 16
Outline

- ISCO
  - Description
  - Revisions / Versions
- ISCO in DDI 3
  - Variable Description
  - Comparison of Variables
  - Standalone Resource Package
- DDI 3 as Metadata Repository
Intention of presentation

- ISCO as an example to show reuse of metadata by the means of DDI 3.0
- Not a presentation on the content of the ISCO classification
- Focus on documenting standard variables in standalone resources
- Report on first experiences with DDI 3.0
ISCO Description

- International Standard Classification of Occupations
- Maintained by the International Labour Organization (ILO)
- ISCO is a tool for organizing jobs into a clearly defined set of groups according to the tasks and duties undertaken in the job.
- ISCO organizes occupations in a hierarchical framework. At the lowest level is the unit of classification - a job - which is defined as a set of tasks or duties designed to be executed by one person.
- Jobs are grouped into occupations according to the degree of similarity in their constituent tasks and duties.
- ISCO is defined in four levels of aggregation, for example ISCO-88(COM) consists of:
  - 10 major groups
  - 28 sub-major groups (subdivisions of major groups)
  - 116 minor groups (subdivisions of sub-major groups)
  - 390 unit groups (subdivisions of minor groups)
ISCO Example: Major Group 6
Skilled agricultural and fishery workers

61 Skilled agricultural and fishery workers
   611 Market gardeners and crop growers
      6111 Field crop and vegetable growers
      6112 Gardeners, horticultural and nursery growers
   612 Animal producers and related workers
      6121 Dairy and livestock producers
      6122 Poultry producers
      6129 Animal producers and related workers not elsewhere classified
   613 Crop and animal producers
      6130 Crop and animal producers
Revisions / Versions of ISCO

- 1958, 1968, 1988 (current revision), 2008 (draft)
- Versions with minor differences exist, for example:
  - ISCO-88(COM) for the European Union (Eurostat)
  - ISCO-88(CIS) for the Commonwealth of Independent States (CIS Statistical Committee)
  - ISCO-88(OCWM) Occupational Classification of Workers in Migration under ISCO-88 of ILO/UNDP Asian Regional Programme on International Labour Migration (ILO, Bangkok, 1992)
- Country-specific variants
- Country-specific classifications which can be converted into ISCO
Examples for ISCO usage in major studies

- Direct usage
  - ESS - European Social Survey
  - ISSP - International Social Survey Programme
  - The Eurobarometer Survey Series
  - SOEP – German Socio-Economic Panel Study

- Conversion from other occupation classification
  - GSS (US General Social Survey) - ISCO-68, converted from 1970 U.S. Census Classification of Occupations and Industries
  - German Microcensus – ISCO-88(COM), converted from classification KldB92
Standard Variables

- ISCO as a standard variable facilitates the usage of an occupation variable in multiple studies.
- Comparison of studies is one of the major goals of standard variables.
- By using standard variables within an organization or across a group of organizations central documentation and maintenance of these variables is improved.
- Standard variables can be maintained in a variable repository.
- The variable repository should be processable by applications for better exploitation.
- DDI 3 provides the metadata structure for a machine-actionable variable repository.
DDI: Metadata Structure for Variable Repositories

- Different parts of a variable are documented in different schemes. This enables reuse of these parts.
- Central description of a variable in a variable scheme
  - Variable scheme
    ISCO-88(COM)
    … other variables
- Category codes are listed in a code scheme which can be referenced by a one or more variables
  - Example: Occupation of respondent, occupation of spouse of respondent
DDI: ISCO Hierarchy of Codes

Example: Major Group 6

61 Skilled agricultural and fishery workers
   611 Market gardeners and crop growers
      6111 Field crop and vegetable growers
      6112 Gardeners, horticultural and nursery growers
   612 Animal producers and related workers
      6121 Dairy and livestock producers
      6122 Poultry producers
      6129 Animal producers and related workers not elsewhere classified
   613 Crop and animal producers
      6130 Crop and animal producers
DDI: ISCO Hierarchy of Codes

Example: Major Group 6

61

611

6111
6112

612

6121
6122
6129

613

6130
DDI: Metadata Structure for Variable Repositories

- Different parts of a variable are documented in different schemes. This enables reuse of these parts.
- Central description of variable in a variable scheme
- Codes for categories are listed in a code scheme which can be referenced by a one or more variables
- Description of categories in a category scheme which can be referenced by one or more code schemes
Example: Major Group 6

61 Skilled agricultural and fishery workers
   611 Market gardeners and crop growers
      6111 Field crop and vegetable growers
      6112 Gardeners, horticultural and nursery growers
   612 Animal producers and related workers
      6121 Dairy and livestock producers
      6122 Poultry producers
      6129 Animal producers and related workers not elsewhere classified
   613 Crop and animal producers
      6130 Crop and animal producers
DDI: ISCO Hierarchy of Codes

Example: Major Group 6

Skilled agricultural and fishery workers
Market gardeners and crop growers
Field crop and vegetable growers
Gardeners, horticultural and nursery growers
Animal producers and related workers
Dairy and livestock producers
Poultry producers
Animal producers and related workers not elsewhere classified
Crop and animal producers
Crop and animal producers
DDI: Metadata Structure for Variable Repositories

- Different parts of a variable are documented in different schemes. This enables reuse of these parts.
- Central description of variable in a variable scheme
- Codes for categories are listed in a code scheme which can be referenced by a one or more variables
- Description of categories in a category scheme which can be referenced by one or more code schemes
- Multiple languages can be used in the documentation, important for labels
DDI: Category Example
Multiple Languages

<l:Category>
  <r:Identification>
    <r:ID>Level2Category61</r:ID>
  </r:Identification>
  <r:Label xml:lang="en">Skilled agricultural and fishery workers</r:Label>
  <r:Label xml:lang="de">Fachkräfte in der Landwirtschaft und Fischerei</r:Label>
  <r:Label xml:lang="fr">Agriculteurs et ouvriers qualifiés de l'agriculture et de la pêche</r:Label>
</l:Category>
DDI: Concepts

- A variable represents a theoretical concept about the reality, for ISCO the concept is “Work”.
- Concepts are described in DDI 3 in a concept scheme. Similar variables can point to the same concept.
- Additionally the concepts can be organized in concept groups. The concept for ISCO is more precisely “Paid work”. “Work” is the more general concept which can comprehend other types of work.
DDI: Minor changes of variables
New revisions of variables

- Variables, code and category schemes undergo changes, i.e. error correction. This change can be documented in a new version of the same DDI scheme.
  - The responsible agency can be documented. Both version and maintaining agency are part of the global unique identifier of an item
    urn:ddi:3_0:CodeScheme:gesis.org:lp:2_0.ISCO88:1_0
  - Additional documentation: version date, version responsibility, version rationale
DDI: Minor changes of variables
New revisions of variables

- Variables, code and category schemes undergo changes, i.e. error correction. This change can be documented in a new version of the same DDI scheme.
- For ISCO a new revision is published every 10 years. Different versions exist for example for ISCO-88. Most of these versions are still used.
- So they can be documented in DDI as separate variables, not as separate versions.
DDI: Relationship between Similar Variables

- The relationship between these similar variables can be documented in comparison tables:

<table>
<thead>
<tr>
<th>ISCO-88</th>
<th>ISCO-08</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>6112</td>
<td>6113</td>
<td>Tree and shrub crop growers</td>
</tr>
<tr>
<td>6114</td>
<td>6115</td>
<td>Mixed crop growers</td>
</tr>
</tbody>
</table>

- Derived categories can be described in the category scheme:
  - 6123 and 6124 (ISCO-88) to 6124 (ISCO-08) Apiarists and sericulturists
  - Human-readable description
  - Derivation command of the used program / package
## DDI: Resource Package

### Resource Package

**Concept Group „Work“**
- Concept „Paid Work“
  - ...

### Variable Scheme

**Variable „ISCO-68“**
- Variable „ISCO-88“

### Code Scheme „ISCO-88“
- Code 6114
  - ...

### Category Scheme „ISCO-88“
- Category „Mixed crop grwr“
  - ...

### Comparison Table Codes

<table>
<thead>
<tr>
<th>Ref. to ISCO-88</th>
<th>Ref. to ISCO-08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. to 6114</td>
<td>Ref. to 6115</td>
</tr>
</tbody>
</table>

---

[Comparison Table Codes](#)
DDI: Resource Package as Reusable Set of Metadata

- Resource packages can be used by multiple studies documented in DDI.

- The same metadata is reused and this implies potential comparability.
DDI: Resource Package as Reusable Set of Metadata

- Schemes in resource packages can be used in multiple ways.

**Study**

- Variable „Occupation of Spouse“
- Variable „Occupation of Father“

**Resource Package**

- Concept „Work“
- Variable „ISCO-88“
- Code Scheme „ISCO-88“
- Category Scheme „ISCO-88“
DDI: Resource Package
Metadata Repository

- Concepts, variables, code schemes, and category schemes as a whole create a metadata registry
- DDI 3 borrowed ideas from the standard ISO/IEC 11179 (description of the structure for metadata registries on an abstract level)
- A metadata registry (for offline or online usage) requires an application in addition to the metadata structure
- Full advantage of a metadata registry by a search front end and by web services for programs
Conclusion

Advantages in usage of DDI metadata repositories
- Easy exchange of the documentation of standard variables
- Enabling metadata mining for finding potential comparable studies
- Support for survey designers, when questionnaires are included
- Can support the process of establishing a standard variable
- Basis of metadata registry application

Requirements / Limitations
- Requires often a well established standard variable like ISCO
- Acting in a controlled environment is required, quality assurance
- Trusted relationship to maintaining organization is necessary, especially when using automated access like a web service
- Large variety of “standard variables” for one concept can cause problems

Using DDI 3 resource packages is one step into the direction of “Building global knowledge communities with open metadata”
Thank you for your attention

Basic DDI example on ISCO-88(COM) available at DDI 3.0 page: http://www.ddialliance.org/ddi3/

joachim.wackerow@gesis.org

African Scops Owl