

# Core Information in DDI 3.0

Exporting SAS Datasets to DDI 3 XML files - Joachim Wackerow / Larry Hoyle

## Study Unit

- Identification
  - Coverage
    - Topical
    - Temporal
    - Spatial
  - Conceptual Components
    - Universe
    - **Concept (stub)**
    - Representation (optional replication)
  - Purpose, Abstract, Proposal, Funding
- Identification is mapped to Dublin Core and basic Dublin Core is included as an option
  - Geographic coverage mapped to FGDC
    - bounding box
    - spatial object
    - polygon description of levels and identifiers
  - Universe Scheme, Concept Scheme
    - link of concept, universe, representation through Variable
    - also allows storage as a ISO/IEC 11179 compliant registry

## Data Collection

- Methodology
- Question Scheme
  - Question
  - Response domain
- Instrument
- Coding Instructions
  - question to raw data
  - raw data to public file
- Question and Response Domain designed to support Question Banks
  - Question Scheme is a maintainable object
- Organization and flow of questions into Instrument
  - Used to drive CAI systems
- Coding Instructions
  - Reuse by Questions, Variables, and comparison

## Physical Storage

- Physical Data Structure
  - Links to Data Relationships
  - Links to Variable or N-Cube Coordinate
  - Description of physical storage structure
    - fixed, delimited or proprietary
- Physical Instance
  - One-to-one relationship with a data file
  - Coverage constraints
  - **Variable and category statistics**
- Data set
  - **In-line data items**

## Logical Product

- **Category Schemes**
- **Coding Schemes**
- **Variables**
- N-Cubes
- Variable and N-Cube Groups
- Data Relationships
- Used as both question response domains and variable representations
- Used as both question response domains and variable representations
- Link representations to concepts and universes through references
- Built from variables (dimensions and attributes)
  - Map directly to SDMX structures
  - More generalized to accommodate legacy data

**Blue – the generated metadata**

## DDI XML Samples

```
<pi:VariableStatistics>
  <pi:VariableReference>
    <r:ID>ISCO88COM-Variable</r:ID>
  </pi:VariableReference>
  <pi:SummaryStatistic>
    <pi:SummaryStatisticTypeCoded codeListID="Summary
Statistic Type"
codeListAgency="DDI">ValidCases</pi:SummaryStatisticT
ypeCoded>
    <pi:Weighted>>false</pi:Weighted>
    <pi:Value>13732</pi:Value>
  </pi:SummaryStatistic>
  <pi:CategoryStatistics>
    <pi:CategoryValue>6130</pi:CategoryValue>
    <pi:CategoryStatistic>
      <pi:CategoryStatisticTypeCoded
codeListID="Category Statistic Type"
codeListAgency="DDI">Frequency</pi:CategoryStat
isticTypeCoded>
      <pi:Weighted>>false</pi:Weighted>
      <pi:Value>57</pi:Value>
    </pi:CategoryStatistic>
  </pi:CategoryStatistics>
</pi:VariableStatistics>
```

```
<l:CategoryScheme id="ISCO88COM-CategoryScheme">
  ...
  <l:Category id="Level4Category6130">
    <r:Label>Crop and animal producers</r:Label>
  </l:Category>
  ...
  <l:CodeScheme id="ISCO88COM-CodeScheme">
    ...
    <l:Code>
      <l:CategoryReference>
        <r:Scheme>
          <r:ID>ISCO88COM-CategoryScheme</r:ID>
        </r:Scheme>
        <r:ID>Level4Category6130</r:ID>
      </l:CategoryReference>
      <l:Value>6130</l:Value>
    </l:Code>
    ...
  <l:Variable id="ISCO88COM-Variable">
    <r:Label>ISCO-88 (COM), International Standard
Classification of Occupations (for European Union
purposes), 1988 version</r:Label>
```