DDI Profiles to Support Software Development as well as Data Exchange and Analysis

NADDI 2014
02.04.2014, Vancouver (Canada)

David Schiller, IAB
Ingo Barkow, DIPF
Outline

- When developing software for data documentation one has to think about how to incorporate the DDI standard
- DDI-L 3.2 is a powerful standard that can support sophisticated data models and data production modes
- At the same time software needs to be manageable and adjusted to the given institution
DDI-L - considerations

- DDI Lifecycle 3.2
- XML Schema
  - DDI 3.2 – about 1150 main elements
  - Each element might contain a number of sub-elements
  - Some elements are quite domain specific
- High degree of complexity
- Number of mandatory element quite low due to the demands of different organizations or domains
Challenge when building tools

- Normally there is only a need for 50-60 main elements for a given survey or organisation
- But not always for the same 50-60 main elements

- Huge challenge when building software tools
- Especially when it is about DDI import and export functionalities
- The results are likely not interoperable solutions
Example from the IAB Metadata project

<table>
<thead>
<tr>
<th>DDI Element</th>
<th>IAB Term</th>
<th>Qbee 0.2</th>
<th>Colectica</th>
<th>Questasy</th>
<th>IAB</th>
</tr>
</thead>
<tbody>
<tr>
<td>@time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>@captureInstrument</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>@panel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>@geography</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>@dataProduct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>@languageRelationship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AuthorizationSource</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UniverseReference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SeriesStatement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExPostEvaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FundingInformation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ProjectBudget</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KindOfData</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coverage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coverage TopicalCoverage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coverage TopicalCoverage Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coverage TemporalCoverage ReferenceDate StartDate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coverage TemporalCoverage ReferenceDate EndDate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OtherMaterial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embargo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StudyUnitReference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SubGrouhReference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data unit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding From</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding To</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period Covered from</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period Covered to</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey waves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Solution DDI Profiles by “topic”

- Creating of DDI Profiles is described in an article from 2009 (Ionescu, Sanda)
- The idea is to not have individual profiles but standardized profiles by “topics”
- Examples of topics may include: “survey design”, “data collection” or more specific “administrative data”; what else?
- Institutions can than choose the needed profiles for their institution and be interoperable with others
Standard DDI profiles: Main goals

- Standard DDI profiles should cover the most used DDI elements by topics
- Agencies can choose from those profiles
- This will lead to an harmonized use of DDI
- And to easier exchange between tools and agencies
Use Case I – Data Exchange

- FDZ of BA at IAB has administrative and social insurance data
  - Responsible German law: social code book
- Survey, health, financial and geo data is or will be merged to IAB data
- Normally data stays in the facilities of the BA, due to the availability of high data security standards
- Merging data sources will be easier when using DDI profiles excepted within the community
Use Case II – Data Exchange

- International research becomes more and more important
- Moving confidential data to a central storage is normally no option, due to legal restrictions
- Trusted third party infrastructures could solve that challenge, by providing secure trusted environments
- DDI profiles will ease the job of finding comparable data sources within different countries
Use Case III – Software development

- DIPF/TBA21 provides a DDI-based survey system
- A full DDI-L instance is much too complex to describe a questionnaire for field work
- Needed is only an extract containing e.g. QuestionItems, ResponseDomains and Conditionals (working questionnaires can only be created in DDI-L 3.2)
- Solution to exchange questionnaires between tools is a „simple questionnaire“ DDI profile
Summary

- **Goal is:**
  - to make DDI-L manageable and
  - to ensure exchange possibilities between different software (harmonization between agencies)
  - to prepare shift to DDI-L 4.0
- Useful standard DDI-L profiles need to be discussed and agreed on
Thank you for your attention

david.schiller@iab.de
barkow@dipf.de

http://fdz.iab.de
http://www.dipf.de