Integrating Colectica with IBM SPSS Data Collection

Ipsos MORI
Who are Colectica?

- The DDI alliance provides a framework but no tools to work with the files.
- Researchers, 3rd party companies and developers provide tools.
- Some are free, others are commercial.
- Colectica provide arguably the most complete and sophisticated products to work with DDI.
Disconnect between Survey Research & DDI

- No DDI Metadata Data Source Component
- No existing DDI tool to do a direct conversion
- We can use indirect conversion tools from DDI providers or in IBM SPSS Data Collection
- Limitations arise and information is lost when doing conversions, e.g. MDD to SPSS to DDI
- Why the conversion issues?
  - Different Schemas/Unimplemented features
  - Survey Routing/Instructions in different languages
Convert MDD to DDI

- IBM SPSS DC & DDI are two large and diverse schemas
- Due to different features it would be impossible to provide a complete conversion
- We tried a basic conversion in mrScriptBasic:
  - Example MDD
  - Read document object model
  - Write to XML DOM
  - Happy to share this script
- An alternative would be to write your own Data Source Component
  - How would you handle routing?
Or, we could let the experts do it…

- Colectica Designer is the most complete desktop software for working with DDI. You can
- Create surveys, including routing
- Work with a library of questions
- Publish your metadata to the web
- Import from a variety of survey formats
- As of Version 5 Colectica Designer enables importing Dimensions (IBM SPSS DC) MDD files
  - All data types will be supported
  - Routing will be supported
Converting SPSS Dimensions to DDI 3.2

Information in Dimensions MDD XML Files

• Question definitions
  • Question text
  • Data type

```xml
<variable id="e509b6de-477b-4047-a68d-8e93c2c076d5" name="MOYR" type="5">
  <labels context="LABEL">
    <text context="QUESTION" xml:lang="en-GB">
      Our records show that {#cohortMembersFFText} {#WASWERETXT} not living at this address at the last interview on {#FF_Family.FFLSTI00}.
      When did {#cohortMembersFFText} move out of the address {#MOYRTXT} {#WASWERETXT} living at on {#FF_Family.FFLSTI00}?
      ADDRESS AT LAST INTERVIEW WAS:
      {#Sampledata.ALI_AddressLine1}, {#Sampledata.ALI_AddressLine2}, {#Sampledata.ALI_AddressLine3}, {#Sampledata.ALI_Town}, {#Sampledata.ALI_County}, {#Sampledata.ALI_OutwardPostcode}, {#Sampledata.ALI_InwardPostcode}
      INTERVIEWER: IF RESPONDENT HAD TWO ADDRESSES AT PREVIOUS INTERVIEW, AND CURRENT ADDRESS WAS WHERE THEY THEN LIVED SOME OF THE TIME, ASK:
      When did address at last interview stop being {#cohortMembersFFText}'s main address?
    </text>
  </labels>
</variable>
```
Converting SPSS Dimensions to DDI 3.2

Information in Dimensions MDD XML Files

- Source code with logical routing
  - mrScriptBasic language, much like VBScript

```plaintext
if useECS then
    sbCheckOffPath(IMSerial)
    if IMSerial = null then
        IMSerial = fnSetNumeric(IMSerial, addressFromECS)
    end if
    if iom.Info.IsAutoAnswer then
        IMSerial = getRandomAddress(sampleFilenameFull)
        IMSerial.Show()
    end if
else
    IMSERIAL.Ask()
    if iom.Info.IsAutoAnswer then
        IMSerial = getRandomAddress(sampleFilenameFull)
        IMSerial.Show()
    end if
end if
```
Converting SPSS Dimensions to DDI 3.2

Mapping to DDI

1. Extract question information from the XML
2. Create a grammar for mrScriptBasic
3. Parse the mrScriptBasic into an Abstract Syntax Tree
   A. Find patterns that affect control flow
      • Map to DDI ControlConstructs
   B. Find {questionName}.Ask() method calls
      • Map to DDI QuestionConstructs
      • Reference the QuestionItems created in step 1
Typical Workflow

Data Extraction

Dimensions MDD + DDF

Colectica Designer

DDI XML

Export to SPSS

QA
Clean and Edit

Export to SPSS

SPSS
(Edited)