MetaDEx – from Data Dictionary to Complete Blaise Code Generation

Farit Vakhetov
Population Health Research, Cancer Care, Alberta Health Services

The 13th International Blaise User Conference
Baltimore, USA
October 2010
About PHR department

- Focused on population-based epidemiological research in the areas of cancer incidence, causes, survival, and prevention.
- Made up of PhD-level epidemiologists with their research teams, Statistical Analysis group and Data Support Unit (DSU).
- DSU designs all in-house surveys, other data collection and transformation tools.
- Two main data collection tools:
  - Cardiff TeleForm® - paper-based questionnaires.
  - Blaise® - CAPI and CATI instruments.
- Blaise has been used since early 2000s
Outline

• Background
• MetaDEx: Design
  – Data Dictionary object model
  – Metadata storage format
• MetaDEx: Implementation & User Interface
  – Data Dictionary specification
  – Blaise specific parameters & code generation
  – Extensions and reports
• Conclusion
Background

• Growing volume and diversity of data
• Need for data documentation
• Data Standardization Initiative (DSI) was proposed in 2007
• Data Documentation Tool becomes a vital part of DSI
• MetaDEx was initially designed as Data Dictionary tool
• Later, it was extended to generating complete Blaise instrument code
Main Features

• Developed in C#.NET with Visual Studio 2008
• Data stored in XML format
• Tool independent Data Dictionary specification
• Blaise parameters specification and code generator
• Data Dictionary part provides sufficient definition for unified access to the data and for data documentation purposes
• May be used in two-step Blaise instrument design process (Subject matter specialist -> IT specialist)
• Because no manual modifications in code needed, always keeps Data Dictionary, Blaise data model and actual data in sync
Object model – Data Dictionary

**) only main properties shown, methods not shown**
Object model – Blaise implementation

DDInstrumentImpl

Can be attached to any DDEntity object (DDField or DDBlock)

DDBlaiseField
- Tag
- RangeMax
- RangeMin
- AllowDontKnow
- AllowRefusal
- AllowEmpty
- ArrayElements

DDBlaiseBlock
- Tag
- TypeName
- BlockType
- IsEmbedded
- BlockSettings
- BlockParameters
- BlockFunctions
- BlockTypes
- BlockLocals
- BlockFields
- BlockAuxfields
- BlockRules

DDBlaiseBlockTypes
- Block
- Table
- Datamodel

*) only main properties shown, methods not shown
XML storage (blocks and fields)

<DOEntity xsi:type="D0Field" Name="B4_Breastfed" LogicalType="SingleChoice" Code=""F|YesNo" FldLength="1" Precision="0">
  <Description>84. Did you breastfeed any of these babies?</Description>
  <ProgCode>B4. Did you breastfeed any of these babies?</ProgCode>
</DOEntity>

<DOEntity xsi:type="D0Block" Name="BS_BreastfedBlock" NumElements="1" TypeName="BS_BreastfedBlock">
  <Description>85. Breastfeeding Table</Description>
  <ProgCode>B85. Breastfeeding Table</ProgCode>
</DOEntity>

<XMLBlock name="BS_BreastfedBlock">
  <BlockParameters>XMLBlock</BlockParameters>
  <BlockLocals>
    <BlockLocals>
      <string>FOR I = 1 TO 20 DO</string>
      <string>IF I citr= NOT THEN</string>
      <string>BreastfedTime[I]</string>
      <string>ENDIF</string>
      <string>ENDDO</string>
    </BlockLocals>
  </BlockLocals>
</XMLBlock>

<EndXMLBlock>XMLBlock</EndXMLBlock>

<Element xsi:type="D0Block" Name="BreastfedTime" NumElements="10" TypeName="BS_BreastfedTime">
  <Description>For each pregnancy, please indicate how long you breastfed the baby/babies</Description>
  <ProgCode>For each pregnancy, please indicate how long you breastfed the baby/babies</ProgCode>
</DOEntity>

<XMLBlock name="BreastfedTime">
  <BlockParameters>XMLBlock</BlockParameters>
  <BlockLocals>
    <BlockLocals>
      <string>BTime</string>
      <string>INPUT</string>
      <string>IF BTime AUX = Yes THEN</string>
      <string>BTime = "Please check time and units. Probably, breastfeeding time can not be more than 5 years"</string>
      <string>ENDIF</string>
    </BlockLocals>
  </BlockLocals>
</XMLBlock>

<EndXMLBlock>XMLBlock</EndXMLBlock>

<XMLBlock name="BreastfedTime">
  <BlockParameters>XMLBlock</BlockParameters>
  <BlockLocals>
    <BlockLocals>
      <string>BTime</string>
      <string>INPUT</string>
      <string>IF BTime AUX = Yes THEN</string>
      <string>BTime = "Please check time and units. Probably, breastfeeding time can not be more than 5 years"</string>
      <string>ENDIF</string>
    </BlockLocals>
  </BlockLocals>
</XMLBlock>

<EndXMLBlock>XMLBlock</EndXMLBlock>
XML storage (User Defined Types)

<DCodingType Name="THighestGrade">
  <Description>Highest grade completed</Description>
  <Codes>
    <DCodePair Code="1" Value="grade5" Descriptor="Fifth grade or less" />
    <DCodePair Code="2" Value="grade6_8" Descriptor="Sixth to Eighth grade" />
    <DCodePair Code="3" Value="grade9_11" Descriptor="Ninth to Eleventh grade" />
    <DCodePair Code="4" Value="HighSchool" Descriptor="High School Graduace or GED" />
    <DCodePair Code="5" Value="PostHSC" Descriptor="Post High School training other than College (vocational, technical, etc.)" />
    <DCodePair Code="6" Value="SomeCol_Ebn" Descriptor="Some College/University" />
    <DCodePair Code="7" Value="GradCol_Ebn" Descriptor="Graduate from College/University" />
    <DCodePair Code="8" Value="PostGrad" Descriptor="Post Graduate" />
    <DCodePair Code="96" Value="DontKnow" Descriptor="Don't know" />
    <DCodePair Code="99" Value="Refused" Descriptor="Refused" />
  </Codes>
</DCodingType>

<DCodingType Name="TMarriedStatus">
  <Description>Married (Relationship) Status</Description>
  <Codes>
    <DCodePair Code="1" Value="Married" Descriptor="Married" />
    <DCodePair Code="2" Value="LvlMarried" Descriptor="Living as married" />
    <DCodePair Code="3" Value="Divorced" Descriptor="Divorced" />
    <DCodePair Code="4" Value="Separated" Descriptor="Separated" />
    <DCodePair Code="5" Value="Other" Descriptor="Other" />
  </Codes>
</DCodingType>

<DCodingType Name="TTimesLongUnits">
  <Description>Amount of time units</Description>
  <Codes>
    <DCodePair Code="1" Value="Weeks" Descriptor="Weeks" />
    <DCodePair Code="2" Value="Months" Descriptor="Months" />
    <DCodePair Code="3" Value="Years" Descriptor="Years" />
    <DCodePair Code="96" Value="DontKnow" Descriptor="Don't know" />
    <DCodePair Code="99" Value="Refused" Descriptor="Refused" />
  </Codes>
</DCodingType>
MetaDEx – Main Screen
Data Structure representation
MetaDEx – User Defined Types
MetaDEx – Optional Properties

**Question:**
- Can be specified as RTF
- Simple RTF editor with Copy/Paste and formatting functions
MetaDEx – Optional Properties

Skip patterns:
- Name (optional)
- Condition
- Description (optional)
- GoToElement

Restrictions:
- Name (optional)
- Condition
- Description (optional)
- CanBeSuppressed
MetaDEx – Blaise Block Properties
MetaDEx – Blaise Field Properties
MetaDEx – Code Generation Form

```c
{***** BLOCK B2B_BaselineMain for B2B_Baseline *****}
(* First sections of B2B Baseline Questionnaire *)
BLOCK B2B_BaselineMain

TYPE
{***Simple type for Yes/No responses with Don't Know and Refusal***}
TYesNo = {
    Yes (1) "Yes",
    No (2) "No",
    DKnow (98) "Don't know",
    Refused (99) "Refused"
}

{***Months with DK and RF***}
TMonthDKRF = {
    Jan (1) "January",
    Feb (2) "February",
    Mar (3) "March",
    Apr (4) "April",
    May (5) "May",
    Jun (6) "June",
```
Entity Copy/Paste operations

• Facilitates reusing questions or sections
• Implements Copy/Paste operations for any selected field or block between different surveys:
  – Copies entire definition
  – Automated check for name uniqueness
  – Special handling for User Defined Types
• With Cut/Paste operations implements move question/section within one survey
MetaDEx - Options

New fields generation template:
- Name prefix
- Auto incremented index
- Type
- Type details

Version information:
- Version
- User Name
- Date
- Auto update
Reporting – Settings

Options dialog box showing checkboxes for various properties to be included in the report.
Reporting – Simple Plain Report

### HTML Report Browser

<table>
<thead>
<tr>
<th>5.4. A4_EverMarried</th>
<th>A4. Before (mainRefDateStr), have you ever been married or lived as married? SingleChoice Yes:.................[1] No:.................[2] Divorced:...........[3] Refused:...........[99]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6. A5a_OtherRelStatus</td>
<td>A5a. Other relationship status 6. B_PregnancyHist</td>
</tr>
<tr>
<td>6.1. B1_WasPregnant</td>
<td>In this section we will be asking you questions concerning your pregnancy history. This includes all live births, stillbirths, miscarriages, abortions, and tubal, molar and other ectopic pregnancies. SingleChoice Yes:.................[1] No:.................[2] Divorced:...........[3] Refused:...........[99]</td>
</tr>
<tr>
<td>6.2. B2_NumPreg</td>
<td>B2. Before (mainRefDateStr), how many times have you been pregnant (include pregnancy on Reference Date)? Number(2.0)</td>
</tr>
<tr>
<td>6.3. B3_PregMore6m</td>
<td>B3. Of all your pregnancies on or before (mainRefDateStr), how many lasted six months or longer? Number(2.0)</td>
</tr>
<tr>
<td>6.5. B5_Breastfed</td>
<td>6.5.1. BreastfeedTime: ARRAY[1..20]</td>
</tr>
<tr>
<td>6.5.1.1. BFTime</td>
<td>Pregnancy #: 1 Breasfteeding Time Number(4.1)</td>
</tr>
</tbody>
</table>
Extensions

• Load Blaise metadata utility:
  – Built on Blaise API
  – Create metadata definition in MetaDEx format
• Server data access tool:
  – Uses Data Dictionary in MetaDEx format
  – Provide unified access to the data on SQL Server
  – Field information and selection
  – Assign/Load field aliases
  – Create SQL Server view from user field selection
Conclusion

• MetaDEx provides:
  – interface for comprehensive survey/data specification
  – complete Blaise instrument code generation
  – configurable Data Dictionary reporting
• MetaDEx was successfully used for design of more than 10 production Blaise instruments for different studies
• Several historical questionnaires was documented with MetaDEx and unified access to the data was provided
• We are still working on further improvements
Contact Information

Farit Vakhetov
Population Health Research, Cancer Care, Alberta Health Services, Alberta, Canada
E-mail: Farit.Vakhetov@albertahealthservices.ca
Thank you