Rogatus - Questionnaire and Metadata Management System
Rogatus QMMS – a planned Open Source solution for the whole lifecycle

Agenda

• Introduction
• The current tool situation in DDI-L
• The Generic longitudinal business process model (GLBPM)
• Deriving software tools from the current models
• Introduction to Rogatus QMMS
• Rogatus QMMS toolset
• Example workflow
• Upcoming activities
• Q&A
Introduction

- DIPF is a public service research institute located in Frankfurt am Main and Berlin (Germany)
- In 2007 DIPF started a research cluster called „Technology Based Assessment“ (DIPF/TBA) to accommodate the needs of technical assessment for educational research
- Currently the DIPF/TBA team consists of about 25 IT-professionals and seven psychometricians
- In January 2012 a spin-off company called TBA21 Assessment Systeme GmbH was started to offer the same services for the professional market
- All programming work for TBA21 is currently handled by cooperation partner OPIT Consulting Kft. in Hungary
- Original mission statement of DIPF/TBA: to consult and support German research institutions and universities when they want to implement computer-based testing
Current tool situation for DDI-L

• A lot of metadata editing tools exist:
  • Questasy (CentERdata)
  • Colectica (Algenta = Colectica)
  • DDIEditor (Danish Data Archive)
  • EDO - Easy DDI Organizer (University of Tokyo)
  • NEPS Metadataeditor (DIPF / University of Bamberg)
  • IAB Metadata Management System (TBA21/DIPF/OPIT/Colectica/Amin)
  • Qbee - Questionnaire Builder (DIPF/GESIS/TBA21/OPIT)
  • Metadata Tools by Metadata Technologies USA
  • DesignQuest (GESIS)
  • And a lot more (sorry if I forgot your particular tool)
Current tool situation for DDI-L

• DDI Tools catalog lists additional tools for conversion of different formats to DDI-L (e.g. Blaise) or other small support tools
• No Nesstar-like dissemination tool for analysis
• Currently much more DDI-C users than DDI-L users due to existing tooling
• Question for DDI-L -> which tools does the user really need?

• Answer might be derived from well-known models regarding DDI-L
  • DDI Lifecycle Model itself
  • Generic Statistical Business Process Model (GSBPM)
  • Generic Longitudinal Business Process Model (GLBPM)
• DDI-L might need a complete tool chain for future success
The DDI3 Combined Lifecycle Model
(*DDI* 2009)
The Generic Longitudinal Business Process Model

### Generic Longitudinal Business Process Model: Overview

<table>
<thead>
<tr>
<th>Step</th>
<th>Process</th>
<th>Sub-Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Evaluate / Specify Needs</td>
<td>1.1 Define research questions, universe &amp; high-level concepts, 1.2 Evaluate existing data &amp; publications, 1.3 Establish outputs &amp; needed infrastructure, 1.4 Define specific concepts to be measured, 1.5 Planning, timetable &amp; needed infrastructure, 1.6 Prepare proposal and get funding</td>
</tr>
<tr>
<td>2</td>
<td>Design / Redesign</td>
<td>2.1 Identify sources, 2.2 Design sampling methods, 2.3 Design collection instruments, 2.4 Specify data elements, 2.5 Specify processing / data cleaning methods, 2.6 Specify analysis plan, 2.7 Organize research team, 2.8 Design infrastructure</td>
</tr>
<tr>
<td>3</td>
<td>Build / Rebuild</td>
<td>3.1 Develop data collection instruments, 3.2 Create or enhance infrastructure components, 3.3 Validate instruments, 3.4 Test production systems, 3.5 Finalize production systems</td>
</tr>
<tr>
<td>4</td>
<td>Collect</td>
<td>4.1 Select sample, 4.2 Set up collection, 4.3 Run collection, 4.4 Finalize collection</td>
</tr>
<tr>
<td>5</td>
<td>Process / Analyse</td>
<td>5.1 Integrate data, 5.2 Classify &amp; code, 5.3 Explore, validate &amp; clean data, 5.4 Impute missing data, 5.5 Construct new variables and units, 5.6 Calculate weights, 5.7 Calculate aggregates, 5.8 Anonymize data, 5.9 Finalize data outputs</td>
</tr>
<tr>
<td>6</td>
<td>Archive / Preserve / Curate</td>
<td>6.1 Ingest data &amp; metadata, 6.2 Enhance metadata, 6.3 Preserve data &amp; metadata, 6.4 Ongoing Curation</td>
</tr>
<tr>
<td>7</td>
<td>Data / Dissemination / Discovery</td>
<td>7.1 Deploy release infrastructure, 7.2 Preserving dissemination products, 7.3 Deploy access control system / policies, 7.4 Promote dissemination products, 7.5 Provide data citation support, 7.6 Enhance data discovery, 7.7 Manage user support</td>
</tr>
<tr>
<td>8</td>
<td>Research / Publish</td>
<td>8.1 Discovery of data and relevant research, 8.2 Access data, 8.3 Prepare data, 8.4 Analyze data, 8.5 Prepare research papers, 8.6 Manage disclosure risk, 8.7 Publish research</td>
</tr>
<tr>
<td>9</td>
<td>Retrospective Evaluation</td>
<td>9.1 Establish evaluation criteria, 9.2 Gather evaluation inputs, 9.3 Conduct evaluation, 9.4 Determine future actions</td>
</tr>
</tbody>
</table>

### Project Management / Quality Management

- Metadata Management

### Use of External Standard Metadata (classifications, concepts, questions, variables)
How could a DDI-L tool chain look like?

Problem: No tool chain currently complete and working seamlessly
Definition

- **Rogatus (Latin = to be asked) Questionnaire and Metadata Management System (QMMS)**
- **Design ideas:**
  - Framework consisting of several tools for survey design (Qbee, Mbee, Tbee)
  - Support for metadata standards (DDI, SDMX)
  - Support for PAPI, CAPI, CATI and online tests
  - Bases on work on previous projects (NEPS, DaQS, QDDS, Austrian CMS)
  - Compatibility to existing DDI tools (e.g. Colectica, Questasy)
  - Whole toolset is split into two perspectives
    - Researcher creating a study (survey management)
    - Researcher working in data center (data management)
Involved institutions

- DIPF – German Institute for International Educational Research (funding, consultancy)
- GESIS – Leibniz Institute for the social sciences (consultancy)
- OPIT Consulting Kft. (software development)
- TBA21 Assessment Systeme GmbH (project management)
- IAB – Institute for Employment Research (funding)
- Colectica (consultancy)
- Alerk Amin (consultancy)

- Currently negotiations with other partners around the world
History and previous projects

- **NEPS** – Development of a metadata system for survey questionnaires (editor, database, reports, portal) which is not compliant to other studies
  Technology: .NET Framework 4 with SQL Server 2008

- **DaQS** – Development of a metadata system for paper-based cognitive items for re-use (editor, database, portal) which does not adhere to standards like DDI or SDMX
  Technology: LAMP with Zend Framework

- **PIAAC** – Development of a translation system for questionnaires and cognitive items using the XLIFF standard
  Technology: JAVA (picking up abandoned OLT project from SUN)
Some related projects from partners

- Colectica (Algenta Technologies)
- Questasy (CentERdata)
- Case Management System (CMS) and CAPI Infrastructure (Statistics Austria)
- Testing Assisstee par Ordinateur (TAO) (CRP / University of Luxembourg / DIPF)
- ZACAT / da_ra / QDDS (GESIS)
- CBA Itembuilder (DIPF / Softcon)
Rogatus tool overview

Survey Administration Portal
- Questionnaire Builder (Qbee)
- CAPI Server
- Survey / Case Management System
- Delivery Server for Onlinetesting
- External Survey Platforms (Blaise, TAO, Nipo)
- PAPI-based Processes

Data Management Portal
- ETL Processes (Import / Export)
- Internal Dissemination Tool / Portal (?)
- Translation Builder (Tbee)

Tools
- Rogatus Portal
- Reporting
- External Dissemination Tool / Portal
- Variable Shopping Basket

Processes
- Design Survey Build Instruments
- Data Collection
- Process / Analyze
- Archive / Curate
- Data Dissemination
Collaborative approach

- All tools use the same relational database structure for storing metadata
- DDI is used as an import and export format to exchange metadata with other tools, but it can also be used to exchange between different instances within Rogatus
- The toolset also includes several server components to exchange data via web services:
  - Team Server – a component for Qbee for sharing questionnaires between research groups
  - Repository – a read-only version of the Team Server for bigger organizations to share qualified items after quality control procedures
Software platform

• The whole software is developed in C# using .NET Framework 4 using SOAP-based web services
• Local clients like Questionnaire Builder (Qbee) use SQL Server 4.0 Compact Edition for storing local data
• Repository and Team Server are based on SQL Server 2008 R2
• Synchronization and replication e.g. in the Case Management System or Case Builder (Cbee) are basing on Microsoft Message Queuing (MSMQ)
• All base components are or will be released under the Lesser GPL v2 open source license (LGPL v2)
• A very early demo release for Qbee (v 0.1) from August 2012 can be found here: http://www.opit.hu/index.php/en/downloads-en
Example Workflow

- To demonstrate the possibilities of a complete survey design and delivery workflow here some screenshots from Rogatus Qbee plus some of the preceding tools:
  - Creation of DDI-based questionnaire (Qbee – Questionnaire Builder)
  - Translation of Questionnaire (OLT, predecessor of Translation Builder)
  - Delivery of questionnaires/cases to interviewers (PIAAC CAPI System, predecessor of Case Builder Administrator)
  - Overview of paradata (PIAAC CAPI System, predecessor of Case Builder Administrator)

- Similar tools are currently in development also for the data management side (Metadata Builder, Rogatus Portal)
Example Workflow – Creation of DDI-based Questionnaire
Example Workflow – Translation of DDI-based Questionnaire
Example Workflow – Delivery of Questionnaires to Interviewers
Example Workflow – Overview of paradata during survey

Filters

Status information

Disposition code
detailed information

Lawrence, 01.04.2013 | Ingo Barkow and David Schiller | North American DDI User Conference (NADDI 2013)
Future activities

• Based on the existing tools we are re-engineering all our survey and data management software which was just shown to fit the Rogatus model
• Nevertheless this process is very time-consuming and resource-intensive
• Ideas for collaboration, e.g. applying for research funding for components are very welcome
• We also search for testers, experts and other contributors to broaden the spectrum of the software
• We currently are working with GESIS on two research proposals in Germany to get additional funding
  • Qbee / Repository – German Research Foundation (DFG)
  • Cbee / Delivery Server – German Ministry for Economy (ZIM)
Any Questions?

barkow@dipf.de
http://www.dipf.de