

Web-based Information System for Business and Organizational Data (DSZBO)

Johanna Vompras

Bielefeld University Library

`johanna.vompras@uni-bielefeld.de`

Utrecht, December 2010

Specific Aspects of our Work

- Web-based Information System for sociological Studies coded in DDI3.x.
- Easy to use Graphical Interface
- General approach for mapping an 'infinite' set of DDI elements into a visualization (e.g. nested questions)
- Data Encapsulation for a simplified modification of queries, adaptation to future DDI release and modification of graphical interface
- In future: Secondary Analysis, Comparison of Longitudinal Data from our databases

Data Layer

- Import of DDI3.x files
- Export of DDI3.x files, PDF, STATA, SPSS (SUF, PUF)
- Focus on Quantitative Data
- Support of Qualitative Data perspectively
- Linkage of Questions/Concepts to Thesauri
- Using the Comparative Module for longitudinal organizational data (e.g. BEATA)
- Persistent Identifiers of StudyUnits/Modules (DOIs, URNs)

Presentation Layer

- User and Role Management
- Search, Browse, Visualization
 - Selection of studies, display of general study information
 - Listings of data collections, questions, concepts, etc.
 - Linking of data with questionnaires and publications, and other materials
 - Search by keywords, tags, concepts
 - Filtering (e.g. number of conducted interviews)
 - Matrix of panel variables included in a time wave

Technical Realization

- Data storage:
 - DDI File: XML Database (BaseX)
 - Other Metadata: Relational Database
- Data Queries/Modification:
 - XQUERY/MySQL
- Data Visualization:
 - JavaScript Framework for Rich Internet Applications (EXT-JS)
 - Website: <http://www.sencha.com/products/js/>

PHP Script with XQUERY ⇒ Results ⇒ JSON ⇒ Input for Components (e.g. GridPanel or DataView)

BaseX XML Database

- Developed by DBIS Group at the University of Konstanz (Open Source)
- BaseX is a powerful and very fast XML database system and XPath/XQuery processor
- Highly conformant support for the latest W3C Updates
- Text, attribute, full-text and path indexes
- Client/Server architecture, supporting ACID safe transactions, user management, logging.
- Interactive visualizations, support. very large XML documents
- Debugging and testing: XQuery editor with syntax highlighting and error feedback

XQuery Editor and Visualization

The screenshot displays the BaseX 6.1 XQuery Editor. The main window is titled "XQuery - allbus-namespaces.xq*" and contains the following XQuery code:

```

declare namespace cm="ddi:comparative:3_0";
declare namespace l="ddi:logicalproduct:3_0";
declare namespace c="ddi:conceptualcomponent:3_0";
declare namespace ds="ddi:dataset:3_0";
declare namespace p="ddi:physicaldataproduct:3_0";
declare namespace pr="ddi:ddiprofile:3_0";
declare namespace s="ddi:studyunit:3_0";
declare namespace g="ddi:group:3_0";
declare namespace pi="ddi:physicalinstance:3_0";
declare namespace m3="ddi:physicaldataproduct_ncube_inline:3_0";
declare namespace m1="ddi:physicaldataproduct_ncube_normal:3_0";
declare namespace m2="ddi:physicaldataproduct_ncube_tabular:3_0";

for $node in doc("allbus")//s:StudyUnit/c:OtherMaterials[(@type='text')] return (concat(data($node/x:Citation/z:Creator)," ",data($node/x:Citation/z:PublicationDate),":",data($node/x:Citation/z:Title),data($node/x:Citation/z:PublicationDate),data($node/x:ExternalURLReference))

```

Below the code, the "Text" pane displays the results of the query, which are a list of bibliographic references:

Christian Gerhards, Stefan Liebzig (2010): Methodenbericht Projekt 'Verknüpfte Personen-Betriebsdaten im Anschluss an den ALLBUS 2008' - ALLBUS Betriebsbefragung 2009 2010 <http://www.uni-bielefeld.de/soz/arbeitbereiche/sozialstrukturanalyse/pdf/persbetdaten/Datenhandbuch%20Personen-Betriebsdaten%20%2804.01.2010%29.pdf> Stefan Liebzig, Christian Gerhards, Jennifer Elmer (2009): Datenhandbuch Projekt 'Verknüpfte Personen-Betriebsdaten im Anschluss an den ALLBUS 2008' - ALLBUS Betriebsbefragung 2009 2009 <http://www.uni-bielefeld.de/soz/arbeitbereiche/sozialstrukturanalyse/pdf/persbetdaten/Datenhandbuch%20Personen-Betriebsdaten%20%2804.01.2010%29.pdf> Stefan Liebzig (2009): Interdisciplinary Longitudinal Surveys: Linking Individual Data to Organizational Data in Life-course Analysis (Working Paper No. 68) 2009 http://www.ratswd.de/download/workingpapers2009/68_09.pdf Stefan Liebzig (2009): Organizational Data (Working Paper No. 67) 2009 http://www.ratswd.de/download/workingpapers2009/67_09.pdf Alexia Meyerarona, Jennifer Elmer, Jürgen Schupp, Stefan Liebzig (2009): Pilotstudie einer surveybasierten Verknüpfung von Personen- und Betriebsdaten. Durchführung sowie Generierung einer Betriebsstudie als nachgelagerte Organisationserhebung zur SOEP-Innovationsstichprobe 2007 2009 <http://www.diw.de/documents/publikationen/73/96938/rn31.pdf>

The right pane shows a visualization of the query results as a tree structure, with the root node being "allbus-studie-ddi3-1-UTF-8.xml". The tree is highly nested, showing the hierarchical structure of the data, with many nodes containing "L" or "C" characters, likely representing logical and conceptual components.

Time Needed: 13.84 ms

Data Retrieval and Visualization

- Each 'Study' is stored in a single XML database
- Queried by XQuery language, e.g. *'select all publications related to the study ALLBUS'*
- Results are returned as lists of items and transformed into JSON format

XQuery: find all publications related to ALLBUS study

```
FOR $node IN doc("allbus")//s:StudyUnit/r:OtherMaterial[@type='text']
RETURN
  ( CONCAT(
    data($node/r:Citation/r:Creator), " (", data($node/r:Citation/r:PublicationDate), ")", ", ",
    data($node/r:Citation/r:Title)),
    data($node/r:Citation/r:PublicationDate),
    data($node/r:ExternalURLReference)
  )
```


Data processing and visualization

JSON (Java Script Object Notation)

```
{ "publications":  
  [  
    { "title": "Christian Gerhards, Stefan Liebig (2010): Methodenbericht Projekt  
'Verknüfzte Personen-Betriebsdaten im Anschluss an den ALLBUS'- ALLBUS Betriebsbefragung 2009",  
      "year": " 2010",  
      "ref": "<a href=\" http://www.uni-bielefeld.de/soz/methods.pdf\">File</a>"  
    },  
    { "title": " Stefan Liebig (2009): Interdisciplinary Longitudinal Surveys: Linking Individual  
Data to Organizational Data in Life-course Analysis (Working Paper No. 68)",  
      "year": " 2009",  
      "ref": "<a href=\" http://www.ratswd.de/download/workingpapers2009/68_09.pdf\">File</a>"  
    },  
  ]  
}
```

Visualization of structured JSON data

Components: Windows, (Tree)Panels, Tabs, (Grouping)Grid,
LayoutManagers
Functions GroupingGrid: *Sorting* and *grouping*

The screenshot displays a web application interface for the University of Bielefeld, Faculty of Sociology, specifically the 'Datenservicezentrum Betriebs- und OrganisationsDATEN'. The interface features a navigation menu on the left with 'Studien' and 'Publikationen' tabs. The main content area shows a list of publications under the heading 'Allbus Betriebsbefragung'. The list is grouped by year (2009 and 2010) and includes details such as title, authors, and publication year. A 'Konsolle' window is visible at the bottom left.

Universität Bielefeld Fakultät für Soziologie

INFORMATION.plus!

Datenservicezentrum
Betriebs- und OrganisationsDATEN

Studien
 ▶ Allbus Betriebsbefragung
 ▶ BEATA Befragungen

Überblick Publikationen TEST-GUI

Allbus Betriebsbefragung

Titel

Jahr: 2009

Stefan Liebig, Christian Gerhards, Jennifer Elsner (2009): Datenhandbuch Projekt "Verknüpfte Personen-Betriebsdaten im Anschluss an den ALLBUS 2008" – ALLBUS Betriebsbefragung 2009

Stefan Liebig (2009): Interdisciplinary Longitudinal Surveys: Linking Individual Data to Organizational Data in Life-course Analysis (Working Paper No. 68)

Stefan Liebig (2009): Organizational Data (Working Paper No. 67)

Alexis Meyerermann, Jennifer Elsner, Jürgen Schupp, Stefan Liebig (2009): Pilotstudie einer surveybasierten Verknüpfung von Personen- und Betriebsdaten. Durchführung sowie Generierung einer Betriebsstudie als nachgelagerte Organisationserhebung zur SOEP-Innovationsstichprobe 2007

Jahr: 2010

Christian Gerhards, Stefan Liebig (2010): Methodenbericht Projekt "Verknüpfte Personen-Betriebsdaten im Anschluss an den ALLBUS 2008" – ALLBUS Betriebsbefragung 2009

Konsolle

JavaScript Framework for Data Visualization (EXT-JS)

EXT-JS: Powerful JavaScript (Open Source LGPL 3.0) library that simplifies AJAX development through the use of reusable objects and widgets

Frontend

- Support of Libraries (DOM Selection, Manipulation, Events, AJAX, etc.)
- Layouts and Containers
- Forms and DataStores
- Effects (fade, blur, hide, focus, etc.)

Backend Integration

- XML, JSON, etc.

An Overview of Data Processing

