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:

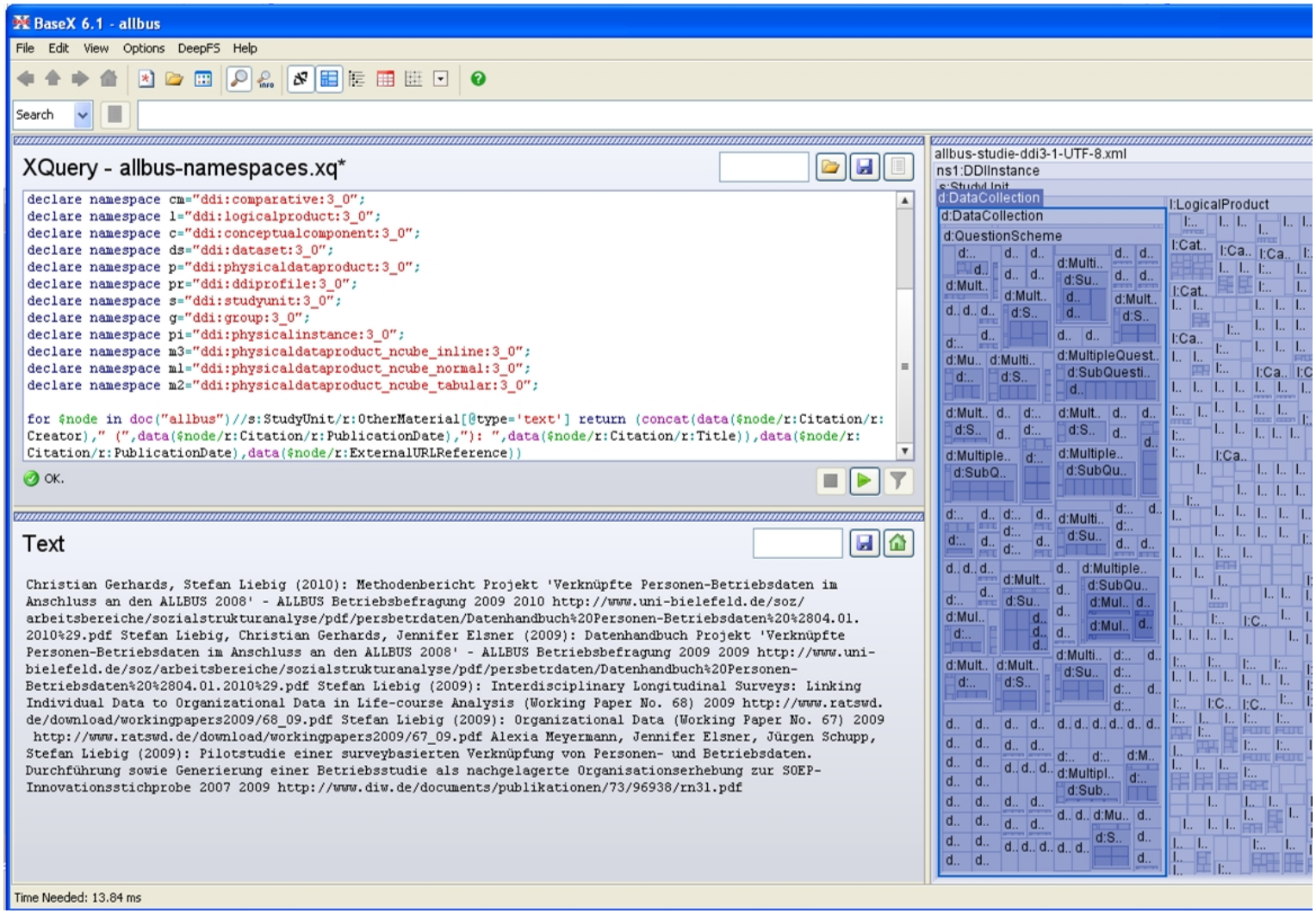
  JavaScript Framework for Rich Internet Applications (EXT-JS)
  Website: `http://www.sencha.com/products/js/`

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



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
    'Verknüfte 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*

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