Metadata Processing

Zuzana Nevěřilová

<xpopelk@fi.muni.cz>

28th April 2010
## Contents

1. DML-CZ Metadata
2. Visualization
3. Server Side
4. Client Side
5. Conclusion
DML-CZ Metadata

- article: authors, title, abstract, language …
- issue: name, date, volume number
- serial: name, periodicity, languages used, description …
- similar articles: number between 0 and 1 that represents similarity
- references
DML-CZ Metadata

• article: authors, title, abstract, language …

• issue: name, date, volume number

• serial: name, periodicity, languages used, description …

• similar articles: number between 0 and 1 that represents similarity

• references
DML-CZ Metadata

- article: authors, title, abstract, language …
- issue: name, date, volume number
- serial: name, periodicity, languages used, description …
  - similar articles: number between 0 and 1 that represents similarity
  - references
DML-CZ Metadata

- article: authors, title, abstract, language ...
- issue: name, date, volume number
- serial: name, periodicity, languages used, description ...
- similar articles: number between 0 and 1 that represents similarity
- references
DML-CZ Metadata

• article: authors, title, abstract, language …
• issue: name, date, volume number
• serial: name, periodicity, languages used, description …
• similar articles: number between 0 and 1 that represents similarity
• references
Visualization

“the visual representation of large-scale collections of non-numerical information”

so far Mathematica Bohemica is visualized

data conversion has preceded to the visualization
“the visual representation of large-scale collections of non-numerical information”

so far *Mathematica Bohemica* is visualized

data conversion has preceded to the visualization
Visualization

“the visual representation of large-scale collections of non-numerical information”

so far Mathematica Bohemica is visualized

data conversion has preceded to the visualization
Data Conversion to RDF

- selecting only the appropriate data for visualization
- assigning IDs for articles, issues, journals and authors
- adding short titles for the visualization
- conversion of the `lang` attribute according to RFC 3066 section 2.3
Data Conversion to RDF

- selecting only the appropriate data for visualization
- assigning IDs for articles, issues, journals and authors
- adding short titles for the visualization
- conversion of the `lang` attribute according to RFC 3066 section 2.3
Data Conversion to RDF

- selecting only the appropriate data for visualization
- assigning IDs for articles, issues, journals and authors
- adding short titles for the visualization
- conversion of the lang attribute according to RFC 3066 section 2.3
Data Conversion to RDF

- selecting only the appropriate data for visualization
- assigning IDs for articles, issues, journals and authors
- adding short titles for the visualization
- conversion of the `lang` attribute according to RFC 3066 section 2.3
RDF Server

• Joseki\(^1\) (just a few lines in config and ./bin/rdfserver)
  
  • so far running on neptun04
  
  • so far data are stored in a file
  
  • provides SPARQL for queries

\(^1\) <http://joseki.sourceforge.net>
RDF Server

- Joseki\(^1\) (just a few lines in config and ./bin/rdfservlet)
- so far running on neptun04
- so far data are stored in a file
- provides SPARQL for queries

\(^1\)\texttt{<http://joseki.sourceforge.net>}
RDF Server

- Joseki\(^1\) (just a few lines in config and ./bin/rdfserver)
- so far running on neptun04
- so far data are stored in a file
- provides SPARQL for queries

\(^1\)<http://joseki.sourceforge.net>
RDF Server

- Joseki\(^1\) (just a few lines in config and ./bin/rdfserver)
- so far running on neptun04
- so far data are stored in a file
- provides SPARQL for queries

\(^1\)<http://joseki.sourceforge.net>
result of query is either

- a projection (using SELECT, similar to SQL SELECT)
- a graph (using CONSTRUCT)

Example:

```
CONSTRUCT { ?x ?y ?z }
WHERE { ?x ?y ?z.
  FILTER( ?x = dmlcz:vector )
}
```
result of query is either

- a projection (using SELECT, similar to SQL SELECT)
- a graph (using CONSTRUCT)

Example:

CONSTRUCT { ?x ?y ?z }
WHERE { ?x ?y ?z.
  FILTER( ?x = dmlcz:vector )
}
result of query is either

- a projection (using SELECT, similar to SQL SELECT)
- a graph (using CONSTRUCT)

Example:
CONSTRUCT { ?x ?y ?z }
WHERE { ?x ?y ?z.
    FILTER( ?x = dmlcz:vector )}
}
Interaction between Visual Browser and HTML page

• Java $\rightarrow$ HTML (via Java Applet methods)

• better comfort: AJAX$^2$

• processing results of SPARQL queries
  result is a graph, conversion to HTML using XSLT $\leftarrow$ not an easy job

$^2$ Asynchronous Javascript And XML
Interaction between Visual Browser and HTML page

- Java $\leftrightarrow$ HTML (via Java Applet methods)
- better comfort: AJAX$^2$
- processing results of SPARQL queries
  result is a graph, conversion to HTML using XSLT $\leftrightarrow$ not an easy job

$^2$Asynchronous Javascript And XML
Interaction between Visual Browser and HTML page

- Java $\leftrightarrow$ HTML (via Java Applet methods)
- better comfort: AJAX$^2$
- processing results of SPARQL queries
  result is a graph, conversion to HTML using XSLT $\Leftarrow$ not an easy job

$^2$ Asynchronous Javascript And XML
Conclusion and Future Work

Visual Browser provides an alternative interface to the library data.

Is it useful? Will it be attractive for users?

Is this alternative interface inspirative for NLP?

Future work comprises processing all the DML-CZ data.