Collaborative Financial Analysis using Semantic MediaWiki

Benedikt Kämpgen, Tobias Zundel

SMWCon Fall 2012, Cologne, Germany, October 25, 2012

Work supported by BMBF under grant 02PJ1002 (SyncTech)
Motivation

- Good companies use for analysing financial data [1]
  - Data Warehouse + ETL pipeline
  - Excel spreadsheets

- Collaborative financial analysis difficult
  - Discussion of financial reports done detached per PowerPoint, mail
  - Changes to planning done without annotating the rationale

- Collaboration crucial
  - Lessons learned for later situations
  - Rationale are retained for later reproduction by others
Problems

- Can Semantic MediaWiki replace common Data Warehousing and Excel spreadsheet analysis?

- Can Semantic MediaWiki allow collaborative decision making on financial data?
Problems

- Can Semantic MediaWiki replace common Data Warehousing and Excel spreadsheet analysis?

- Can Semantic MediaWiki allow collaborative decision making on financial data?
Case Study: Semantic MediaWiki for ls3finance

- **ls3finance**
  - **3 people** responsible for management of third-party funding at research group of around **30 people**

- **Data sources**
  - Metadata about personnel, projects etc. (Excel, SVN)
  - Planning of expenses and budget (Excel, SVN)
  - Accounting transactions (University Data Warehouse)

- **Tasks**
  - Daily creation and editing of expense and budget plans
  - Monthly import and check of accounting transactions
  - Collaboration crucial
**Is3finance: Requirements**

- Spreadsheet analyses
  - Overviews of people, projects etc.
  - Pivot analyses – Comparison of budget planning, expense planning and accounting transactions
  - Personell table – Overview of project allocation per project/person

- Collaboration
  - Flexible annotation and commenting
  - Synchronous editing, reproducible editing
  - Security
Is3finance: Requirements

- Spreadsheet analyses
  - Overviews of people, projects etc.
  - Pivot analyses – Comparison of budget planning, expense planning and accounting transactions
  - Personell table – Overview of project allocation per project/person

- Collaboration
  - Flexible annotation and commenting
  - Synchronous editing, reproducible editing
  - Security
Architecture

Input:
- Semantic Forms
- Import script

Output:
- Inline Queries
- Result Formats for visualisations
- Spark [2] for SPARQL queries and visualisations
Expense and Budget Plans

- Person
- Project
- Personell plan
- Travel plan

...
Expense and Budget Plans – Linking entities
Expense and Budget Plans – Linking entities

[Diagram showing relationships between entities such as Person, Kategorie, Projekt, Sachkonto, Kontonummer, Projektkonto, Projektzeitraum, Datum, Betrag, Kommentar, Nummer, and Startdatum, Enddatum with arrows indicating relationships and cardinalities.]
Expense and Budget Plans – Modelling and Querying Statistics

- Data Cube with
  - Account transactions (planned)
- Dimensions
  - Project
  - Person
  - Date
  - Type (expense, budget,...)
  - Finance category
- Measures
  - Sum of amount

Import and reuse of RDF Data Cube Vocabulary
http://www.w3.org/TR/vocab-data-cube/
Monthly Import and Check of Accounting Transactions

B. Kämpgen – Collaborative Financial Analysis using Semantic MediaWiki
Implicit link between plans and accounting transactions

Explicit links through helper entities and properties (e.g., project through project account)
Monthly Import and Check of Accounting Transactions (3)

- Import script reusing RDF Data Cube Vocabulary
- Adding accounting transactions to Data Cube
- Preprocessing using construct queries
- Simultaneous querying of transactions from SMW and from Data Warehouse
Monthly Import and Check of Accounting Transactions (3)

- Import script reusing RDF Data Cube Vocabulary
- Adding accounting transactions to Data Cube
- Preprocessing using construct queries
- Simultaneous querying of transactions from SMW and from Data Warehouse
Pivot analyses

Comparison of budget planning, expense planning and accounting transactions

Spark with OAT-Pivot visualisation
Pivot analyses

Comparison of budget planning, expense planning and accounting transactions

<table>
<thead>
<tr>
<th>A</th>
<th>BudgetPlanning</th>
<th>HiwiPlanning</th>
<th>PersonellPlanning</th>
<th>TravelPlanning</th>
<th>expense</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/HiwiReis">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/HiwiReis</a></td>
<td>9600.00</td>
<td>-945.92</td>
<td>0.00</td>
<td>0.00</td>
<td>-972.21</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Reis">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Reis</a></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1364.00</td>
<td>0.00</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Inland_Reis">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Inland_Reis</a></td>
<td>60000.00</td>
<td>0.00</td>
<td>-66000.00</td>
<td>0.00</td>
<td>-55698.39</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Ausland_Reis">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Ausland_Reis</a></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-53.00</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Wiss_Hiwi">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Wiss_Hiwi</a></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-1761.10</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Student_Hiwi">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Student_Hiwi</a></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-2940.39</td>
</tr>
</tbody>
</table>

Spark with OAT-Pivot visualisation
Personell table

### Per project

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>2012-09-01Z</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>2012-10-01Z</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

### Per person

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>2009-09-01Z</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>2009-10-01Z</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>2009-11-01Z</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>2010-01-01Z</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
### Personell table

#### Per project

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Person_a3be75100faeaa71395b5efe84205df">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Person_a3be75100faeaa71395b5efe84205df</a></td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Person_77ede392671a4fe4901dfb10151228bf">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Person_77ede392671a4fe4901dfb10151228bf</a></td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

#### Per person

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_ba29657a6f2b3766d797c89e9422d12">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_ba29657a6f2b3766d797c89e9422d12</a></td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_83a9f6e35d8a79067d2237f75497b5f3">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_83a9f6e35d8a79067d2237f75497b5f3</a></td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_88e8718e55e1feebd318a80b77060f21">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_88e8718e55e1feebd318a80b77060f21</a></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td><a href="http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_0485bca66c47cbf84bd898544a46b7da">http://wiki2.local/F-Wiki/index.php/Special:URIResolver/Project_0485bca66c47cbf84bd898544a46b7da</a></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Lessons Learned

- Maintainability: e.g., Spark + Sesame (CORS, JSON)
- Query capabilities, e.g., flexible SPARQL 1.1 queries
- Performance, e.g., import of 6,011 transactions takes only couple of minutes
- Agility, e.g., adding constant factor to all plans through template
- Robustness, e.g., red links for wrongly spelled projects
Outlook

- Evaluation of applicability at Is3finance
  - Spreadsheet analysis capabilities
    - e.g., who is underestimating their travel costs?

- Collaboration capabilities
  - e.g., discussion through comments and talk pages
  - e.g., recent changes to keep up to date

- Extend current pivot functionality to interactive OLAP interface (e.g., with hierarchies)
Conclusions

- Semantic MediaWiki can be used for spreadsheet analysis
  - e.g., overview tables, pivot tables, diagrams
- Lessons learned not only relevant in financial domain
  - e.g., data integration, performance, agility, robustness
- Collaborative analysis capabilities promising but yet to be demonstrated
  - e.g., discussions, synchronous editing
Thanks for your feedback and questions

<table>
<thead>
<tr>
<th>Property</th>
<th>Anzahl</th>
<th>Anzahl mittels SPARQL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category:Qb:Observation</td>
<td>3442</td>
<td>count 9431</td>
</tr>
<tr>
<td>Property:Qb:dataset</td>
<td>3442</td>
<td>count 9431</td>
</tr>
<tr>
<td>Property:CubeDefinition/Dimension/category</td>
<td>3442</td>
<td>count 8903</td>
</tr>
</tbody>
</table>

![Data diagram](image)

![ER diagram](image)
References
