Semantic OLAP: Using OLAP Functionality to Analyse Semantic MediaWiki Data

SMWCon Fall 2016

Tobias Weller, Maria Maleshkova, Martin Wagner, Lena-Marie Ternes and Hannes Kenngott
Motivation

Martin often calls me to get information from our wiki about the patients.

Don’t worry Martin! I’ll have a look!

Hi Tobi!
Can you please look up how many patients in our system has asthma and what their age is?
Motivation

- Users want to analyse data
  - Have ad-hoc requests
  - Are not always familiar with wiki language

- Analyse data in Semantic MediaWiki
  - Integrated System

- Search functionality does not always answer questions
  - Overview about data is preferable
  - Aggregation functionality is preferable

- Export functionality to reuse data in other tools
Related Extensions

- Google Analytics Integration, Open Web Analytics, Piwik Integration
  - Statistics about wiki pages (e.g. Access)

- EventLogging
  - Captures and analysis reads and edits of wiki pages

- SemanticWebBrowser
  - Faceted Browsing
  - Does only allow to browse the data and not really analyse
Contribution

- Extension for SMW to analyse the data in Semantic MediaWiki by using OLAP functionality

- OLAP functionality
  - Pivoting
  - Roll Up
  - Drilldown
  - Slice

- Export data in Excel
Sneak Preview

Semantic OLAP: Using OLAP Functionality to Analyse Semantic MediaWiki Data
Implementation

- First Step is to Query the Data

- Users may be not familiar with ASK Queries
  - Help with a form

SemanticOLAP

Select at least one Category

Select at least one Property

Submit
Implementation

- Form provides Autocomplete

- At least one category and property must be selected
  - Otherwise ASK Query does not work
  - Multiple Categories and Properties can be selected
Implementation

- Errors help users to comprehend false input
  - Missing input

- No results gained

![Error in Query](image)

Semantic OLAP: Using OLAP Functionality to Analyse Semantic MediaWiki Data
Implementation

- Results from ask Queries are processed and presented

Semantic OLAP: Using OLAP Functionality to Analyze Semantic MediaWiki Data
Implementation

- For properties that are used multiple times on wiki pages, the amount of usage is given additionally
  - Marked with Hashtag (#)

Has Anforderungszeit 1 January 2015 10:18:25 +
Has Asthma false + and true +
Has Ausschleusen 1 January 2015 19:55:00 +

Semantic OLAP

Select Data

- Rows: 
- Columns: 
- Export: XLS

Fields
- Has Asthma
- Has Asthma
- Has BMI
- # WikiPage

Data
- Has Alter (sum)

Wiki Page
Implementation

- Units of Properties of type **Quantity** are considered
  - List of units are not yet completed
  - Unknown units are not displayed
Implementation

- Drag & Drop fields to columns or rows
- OLAP functionality: Pivoting / Split

SemanticOLAP

Select Data

Rows: Σ Columns: Export: XLS

Fields

Has Asthma Has BMI # WikiPage

Data

Has Asthma

RektumPatient 1 66
RektumPatient 10 62
RektumPatient 100 15

Wiki Page

RektumPatient 1 66
RektumPatient 10 62
RektumPatient 100 15
RektumPatient 1000 63
RektumPatient 1001 70
Implementation

- Drag&Drop fields to columns or rows
- OLAP functionality: Pivoting / Split
- Row/Columns Sum can be toggled
Implementation

- Filter for specific values
  - OLAP functionality: Slice/Dice
  - Filter with matches
- Sort values
Implementation

- Export Pivot table into xls
Conclusion

- SemanticOLAP allows to analyse Semantic MediaWiki Data
  - Providing OLAP functionality
  - Not restricted to specific use-case

- Usable by non wiki-experts
  - More trust in the analysis and system
  - Users are not dependent of wiki experts

- We use SemanticOLAP to
  - Allow Physicians to query data
  - Analyse data
    - Select patients for studies