Semantic MediaWiki and the "Semantic Bundle"

Yaron Koren
Ontolog
December 19, 2012

About me

- MediaWiki developer and consultant, based in New York City.

- Founder and head of the MediaWiki consulting company WikiWorks (**wikiworks.com**)

- Run the MediaWiki-based wiki farm Referata (**referata.com**)

- Wrote the recent book *Working with MediaWiki* (**workingwithmediawiki.com**)

MediaWiki

An open-source wiki application, best known for running Wikipedia, but also a very popular engine for regular wikis.

Homepage: **mediawiki.org**

Semantic MediaWiki

An extension to MediaWiki that allows for storing and querying text in the wiki as data, turning the wiki into a "collaborative database".

Homepage: **semantic-mediawiki.org**

Abbreviated as "SMW".

Please don't call it "Semantic" or "Semantic Wiki" - those both mean something else!

Storing data in SMW

On a page called "Susie's Shoes":

```
The phone number of this store is
[[Has phone number::123-4567]].
```

This will store the data as a semantic triple, of the form:

Susie's Shoes Has phone number 123-4567

In SMW terms, "Has phone number" is the *property*.

Querying data in SMW

The #ask and #show "parser functions" are used to query the data, after it is stored.

To display a table of all stores, and their phone numbers, call the following:

```
{{#ask:[[Category:Stores]]|?Has phone number}}
```

This query looks for all pages tagged as being in the category "Stores", and then displays a single semantic property for each.

Querying a specific value using #show

#show is used to get just one value. To get just the phone number for Susie's Shoes:

```
{{#show:Susie's Shoes|?Has phone number}}
```

Semantic MediaWiki and templates

Data is only rarely stored in free text, like in the previous example. Instead it is usually stored via "templates", which are like macros in MediaWiki.

Sample template call:

```
{{template name
|param1=value1
|param2=value2
}}
```

Why use templates for SMW?

Templates containing SMW properties do three things at the same time:
• handle all the storage of data, behind the scenes
• handle the display of data (usually in the style of Wikipedia's right-hand-side "infoboxes")
• standardize the set of data for each page type

SMW and templates: an example

Most likely, the page "Susie's Shoes" would contain something like:

```
{{Store
|Store type=Shoes
|Address = 1000 Shoe Ln.
|Phone number=123-4567
...
}}
Susie's Shoes is a conveniently-located, ...
```

N-ary/compound data in SMW

Not all data can be stored using regular SMW properties.

Example: "Susie's shoes is open on Monday from 9 AM to 5 PM."

This is the kind of information that is put into a table – i.e. n-ary/compound data.

It is accomplished in SMW using *subobjects*.

#subobject calls

```
{{#subobject:|Store=Susie's Shoes|Day of
week=Monday|Start time=9 AM|End time=5
PM}}
```

This would most likely also go into a template, so the page might look like:

```
{{Store

...

|Hours={{Hours for day|Monday|9 AM|5 PM}}
{{Hours for day|Tuesday|9 AM|5 PM}}...

...}}
```

Semantic MediaWiki and RDF

SMW can be configured to store its data in an RDF triplestore (Virtuoso, 4store, etc.)

Its data can then be queried using SPARQL.

Semantic MediaWiki and Wikidata

Wikidata is an amazing project that will provide a giant data store for Wikipedia.

Wikidata and SMW have some shared code (and developers), but they are two different software projects!

Wikidata is intended for massive, multilingual wikis; SMW is intended for regular-sized, single-language wikis.

Semantic MediaWiki and Wikidata

Wikidata is an amazing project that will provide a giant data store for Wikipedia.

Wikidata and SMW have some shared code (and developers), but they are two different software projects!

Wikidata is intended for massive, multilingual wikis; SMW is intended for regular-sized, single-language wikis.

"Semantic Bundle"

- A collection of Semantic MediaWiki-based extensions that can be downloaded together – I help to maintain it.

- These also represent, in my opinion, the ideal set of extensions to be used jointly with SMW (although there are many other helpful MediaWiki extensions).

Some of the Semantic Bundle extensions

- Semantic MediaWiki
- Semantic Result Formats
- Semantic Maps
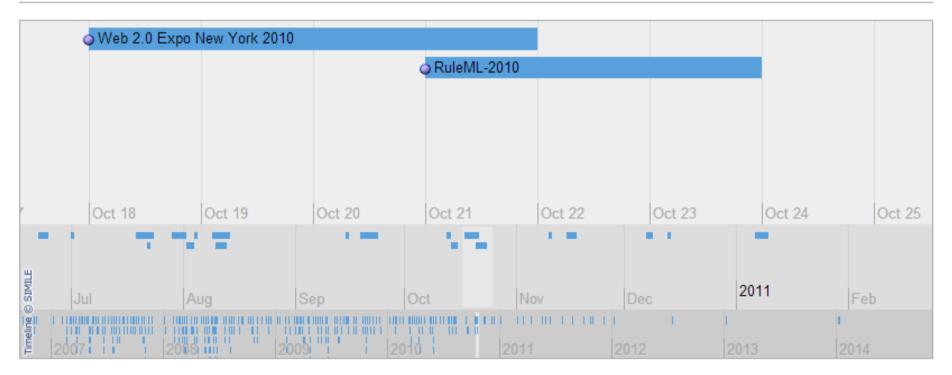- Semantic Compound Queries
- Semantic Forms
- Semantic Forms Inputs
- Semantic Drilldown

- External Data
- Data Transfer
- Approved Revs
- ParserFunctions
- Widgets
- Header Tabs
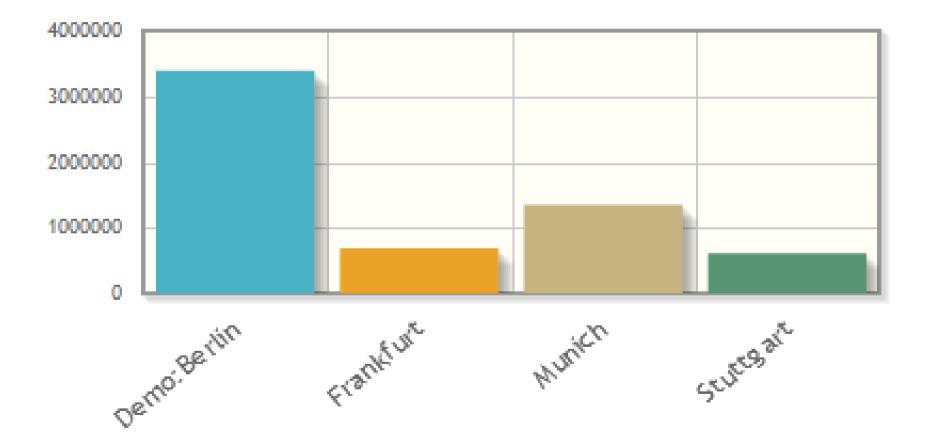- Replace Text
- Admin Links

Semantic Result Formats

Additional display formats for Semantic MediaWiki: calendars, charts, graphs, tag cloud, gallery, etc.

Semantic Result Formats: timeline format

Semantic Result Formats: jqplotbar format

Semantic Result Formats – other notable formats

- **calendar** – displays a monthly calendar
- **eventcalendar** – similar, but uses Javascript instead of HTML
- **filtered** – similar to Semantic Drilldown (covered later), but uses Javascript instead of HTML
- Many other charting libraries, besides jqPlot (D3, Dygraphs, Google Charts, etc.)
- Math operations (**sum**, **average**, etc.)

Semantic Maps

Map-based display and input of coordinate data.

Semantic Maps: Google Maps

Semantic Compound Queries

Can display multiple Semantic MediaWiki queries at the same time.

Semantic Compound Queries: map

Airports

The following map shows the Köppen climate classification for aerodromes across North America which are listed on SKYbrary:

Semantic Compound Queries: calendar

October 2012

◀ Today ▶

October ▾ 2012

Go to month

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
30	**1**	**2**	**3** Re-Elect the President	**4**	**5**	**6**
7 For president: Mitt Romney / Obama for president: A second term for a serious man	**8**	**9**	**10**	**11** A Possibly Fatal Mistake	**12**	**13**
14 Barack Obama for president (Sacramento Bee)	**15** Obama will do a better job	**16**	**17** Time for another change	**18** Ken Burns: Why I am voting for Barack Obama	**19** Barack Obama for president (Denver Post) / Our pick for president: Romney / Too Many Mitts: Obama has earned another term	**20** Mitt Romney: new leadership for more economic growth / Mitt Romney for president / Nevada needs a change now; elect Mitt Romney president / Obama for president

Semantic Forms

- The most popular SMW-based extension.

- Allows for making forms for creating and editing template-based wiki pages.

- Forms can also be used to run search queries.

Semantic Forms

Features include:

- many input types (text, textarea, dropdown, radio button, checkboxes, listbox, combo box, date field)

- autocompletion (single and multiple values)

- other extensions can define additional input types (Semantic Maps defines a map-based input)

- field validation (mandatory fields, etc.)

Semantic Drilldown

Provides a drill-down interface for SMW data.

You can filter on the values of each property, and number and date values can be aggregated into buckets.

At every state, the set of pages matching the filters is shown below.

Semantic Drilldown – example (from broadcastforschools.co.uk)

Semantic Drilldown – example (after clicking 1980s)

Choose a category:

| Images (2470) | Programmes (2896) | Schedules (34) | Subjects (13) |

Programmes > Decade: 1980s ⊗

Click on one or more items below to narrow your results.

▶ **DECADE:** (click arrow to add another value)

▼ **CURRICULUM AREA:**

Art (9) · Citizenship (272) · English (133) · Geography (69) · History (66) · Languages (79) · Mathematics (32) · Miscellany (79) · Movement (17) · Music (35) · Religion (24) · Science (83) · Teachers (5)

▼ **AGE RANGE:**

Age 11-13 (197) · Age 13-16 (441) · Age 16-19 (122) · Age 3-5 (15) · Age 5-7 (53) · Age 7-9 (92) · Age 9-11 (152)

▼ **BROADCASTER:**

BBC Schools Radio (422) · BBC Schools TV (330) · Channel 4 Schools (11) · ITV Schools (153) · S4C Ysgolion (1)

▼ **COMPANY:**

ATV (13) · Anglia (2) · Ash Films (1) · Associated-Rediffusion (2) · BBC (742) · BBC Wales (1) · Central (14) · Childsplay (1) · Ffilmiaur Nant (1) · Grampian (10) · Granada (26) · HTV (19) · KCET (1) · Lodestar (1) · Opus 30 (1) · SFTV (3) · Scottish (8) · TSW (1) · Television Techniques (1) · Thames (32) · Ulster (7) · Yorkshire (27) · independent companies (11)

▼ **PRESENTED IN LANGUAGE:**

In Bengali (1) · In British Sign Language (1) · In Cantonese (1) · In Cornish (1) · In English (793) · In French (34) · In German (16) · In Greek (1) · In Gujarati (1) · In Irish Gaelic (2) · In Italian (6) · In Klaptonian (1) · In Punjabi (1) · In Russian (3) · In Scottish Gaelic (9) · In Spanish (9) · In Turkish (1) · In Urdu (1) · In Welsh (43)

That's it!

You can find links to these extensions, and others, at:
https://www.mediawiki.org/wiki/Semantic_Bundle

Shameless plug: again, you can find my book at:
http://workingwithmediawiki.com/