MediaWiki & Semantic MediaWiki tutorial

Yaron Koren
Spring 2011 SMWCon
Arlington, VA

April 28, 2011
Yaron Koren

- New York City-based software developer
- Wikipedia enthusiast since 2005
- Semantic MediaWiki enthusiast since 2006
- MediaWiki developer and consultant since 2007
- My extensions: Semantic Forms, Semantic Drilldown + 9 more
My MediaWiki consulting company:

WikiWorks
MediaWiki consulting

(wikiworks.com)

Via WikiWorks, also run the Semantic MediaWiki-based wiki farm Referata, at referata.com.
MediaWiki

- Open-source, PHP-based wiki application
- Created in 2003 for use on Wikipedia
- Rewritten twice since then, most recently by Brion Vibber
- Now maintained by hundreds of developers, some paid by the Wikimedia Foundation, others volunteers
MediaWiki is popular

• Powers most of the major public wikis
• Popular within corporations and organizations as well (is it the most popular? hard to tell)
• Used by Wikia (wikia.com), the most popular wiki farm
MediaWiki wiki-text
(http://smwcondc.referata.com)

==Header 1==
===Header 2===

''italics''
'''bold'''
''''bold and italic''''
[[Link to wiki-page]]
[[Some page|link with custom text]]
[http://some.url Link to an outside URL]
* bulleted list
# numbered list
: indentation
Namespaces

- none (Main) – regular pages
- "User:" - user pages
- "Category:" - category pages
- "File:" - images and other uploaded files
- "Template:" - template pages
- "MediaWiki:" - system messages
- "Special:" - provide functionality; not editable

- "Talk", "User talk", "Category talk" etc. - discussion of contents on corresponding page
Categories

Add a page to a category by putting this anywhere in the page (though usually at the bottom):

[[Category:Category name]]
Templates
Pages meant to be transcluded within other pages, and usually called with one or more parameters.

Calling a template:

```
{{Template name|Param1=Value1|Param2=Value2|...}}

{{Template name|Value 1|Value 2}}

In template: {{{{Param1|}}}}
```
Redirects

#REDIRECT [[Actual page]]

Useful for synonyms, misspellings, content subsets (like a book redirecting to its author)
Images

To embed an uploaded image in a wiki:

[[File:image name
|border/frame/frameless/thumb
|widthpx
|xheightpx
|left/right-center/none
|link=
|alt=]]
Dynamic elements

Three kinds:
- Parser functions: `{{#function:...}}`
- Tags: `<tag>...</tag>`
- Magic words: `{{MAGICWORD}}, __MAGICWORD__`
Parser functions

{{#function-name: param1 | param2 | ...}}

Important functions (all defined in ParserFunctions extension):

#if, #ifeq, #switch, #sub, ...

{{#if: {{{param1|}}} | | ... | {{{!}}} | ...}}

These are especially important in templates.
MediaWiki tags

Look like HTML tags, but they're different!

Important tags:
<nowiki>, <noinclude>, <includeonly>, <ref>
(used for citations)

The way to do "'italics'" is
<nowiki>''like this!''</nowiki>
Useful magic words

```
{{PAGENAME}} – displays name of current page

__NOTOC__ - hides table of contents
```
Skins

Main MediaWiki skin is Vector, developed last year by the Wikimedia Foundation – before that, MonoBook was the main skin.

Site admins can set another skin as the default, and create their own.

Users can also choose their skin, in Special:UserPreferences.
MediaWiki is highly customizable

LocalSettings.php is MediaWiki's initialization file – many features can be modified via lines added to that file.

All system messages can be modified – each one has a page in the MediaWiki: namespace.

The sidebar, for instance, is defined by the page MediaWiki:Sidebar.
MediaWiki extensions

MediaWiki has hundreds of extensions, covering many different areas of functionality.

Most extensions have a page on mediawiki.org.
Some crucial extensions

- ParserFunctions - #if, etc., plus string operators
- Cite – references
- ConfirmEdit – anti-spam CAPTCHAs
- Widgets – secure embedding of HTML and Javascript
- Replace Text – global search-and-replace
- Semantic MediaWiki!
The Semantic Web

The Semantic Web = lots of triples!

Triple:

Subject  Predicate  Object

William Shakespeare  Has occupation  Playwright
RDF and OWL

RDF = Resource Description Framework
- a framework for storing semantic triples

RDF/XML – a file format for storing RDF data

OWL = Web Ontology Language
- a superset of RDF – used to store both data and information about data structure
SPARQL

SPARQL = SPARQL (formerly Simple) Protocol and RDF Query Language

- standard language for querying and modifying RDF data
Semantic MediaWiki

A MediaWiki extension, originally developed by Markus Krotzsch, Denny Vrandecic and several others in 2005.

Current main developers: Markus Krotzsch and Jeroen De Dauw.

In use on hundreds of sites, has > 30 spinoff extensions.
Semantic MediaWiki syntax

In any regular page:

`[[Property name::Value]]`

This defines the semantic triple:

```
Page name  Property name  Value
```
SMW special properties

Pre-defined properties in SMW – these include:

Modification date, Has type, Allows value
SMW property types

Each property has its own page, in the “Property:” namespace. There the type and allowed values are defined.

Standard SMW data types:
String, Page, Number, Date, Boolean, Text, URL, Email, ...

Allowed values defined using “Allows value”.
Viewing a page's semantic data

“Browse properties” link in sidebar – takes you to:

Special:Browse/Page_name
Querying SMW data

Two parser functions: `#ask` and `#show` (`#ask` is the main function).

```
{{#ask: [[Category:Countries]] [Has continent::Europe]
  | format=table
  | mainlabel=Name
  | ? Has population}}
```

The Special:Ask page is the easiest way to generate a query.
SMW query formats

list, ol, ul, table, template, count,
rss, csv, json, ...

(Many more defined in other SMW-based extensions, like Semantic Result Formats and Semantic Maps)