

Creating the structured semantic wiki

Yaron Koren
Wikimania 2008
July 19

About me

Freelance developer, living in Brooklyn, New York.

Doing Semantic MediaWiki-related development since June 2006.

First experience with SMW was creating the semantic wiki Discourse DB, at discoursedb.org.

Author of all the extensions discussed here, except where otherwise noted.

Two basic ways to organize and store digital information:

Traditional content-management systems (CMS) and wikis

CMS

Strengths: Good handling of structured data

Weaknesses: Difficult to set up, editing of data limited to a small number of users

Wiki

Strengths: Flexible, easy to set up, allow mass collaboration

Weaknesses: Poor handling of structured data

The goal

Can we create a system that combines the structure of a CMS with the flexibility and mass-collaboration ability of a wiki?

Our “platform”: Semantic MediaWiki

MediaWiki extension created in 2005 by Markus Krötzsch, Denny Vrandečić et al.

Enables adding queryable data to wiki pages.

Basic syntax:

[[Property name::value]]

Structured approach to semantics

All semantic properties and category declarations should be contained in templates.

- Prevents regular users from having to think about or understand properties and categories
- Standardizes data structure for each page type
- Standardizes visual display

Functionality provided by a CMS

- User management
- Viewing data
- Adding/editing data
- Aggregating data
 - Lists, tables, calendars, maps...
- Search
- Drilldown (faceted browsing)
- Workflow
- Importing/exporting data

Let's go through the set of
functionality of a CMS...

User management

Already handled by MediaWiki

Viewing data

Semantic data for a page is shown through the combination of SMW properties and templates.

Many MediaWiki extensions can be used to display videos, maps, RSS feeds, etc.

Adding/editing data

Semantic Forms extension

Forms definitions created in a page in the “Form:” namespace, using a custom scripting language.

Form definition, templates, properties and existing page are all parsed on the fly to create a form.

Forms associated with a category or namespace, to allow automatic “edit with form” tab.

Aggregating data

Semantic MediaWiki handles much aggregation, through **#ask** parser function: lists, tables, etc.

Semantic Calendar extension: displays data in a monthly calendar

Semantic Google Maps extension (Robert Buzink & Yaron Koren): displays data in a Google Map

Search

Since all data is stored in text form, regular MediaWiki search can search on all fields simultaneously

- A nice benefit of using semantic wikis

Drilldown / faceted browser

Semantic Drilldown extension

Displays all top-level categories, their subcategories and pages.

Filters can be defined for each category, using pages in the "Filter:" namespace

Workflow

Semantic Tasks extension (Stereon Giannini and others at Creative Commons)

Lets users be notified when certain task dates are reached.

- More work can be done here

Importing/exporting data

SMW allows for export of semantic data in RDF format.

Data Transfer extension: export data in simple XML form.

- No good solution yet for importing data

Usage in the real world

Currently, around 50 public wikis use Semantic Forms (the most popular of the “structured” extensions)

Around the same number of private wikis

Public wikis used for many purposes; most are for listings of people, businesses, places, etc.

Private wikis seem to be mostly used by corporations for employee and project management

Benefits of a semantic wiki

- Allow for global editing of data
- Can be set up without any programming
- Easy to set up and maintain
- Flexible data structures

Revolutionary implications

Worth noting

This is not the only approach to using Semantic MediaWiki – some wikis use semantic properties in free-form text.

Halo extension (created by Ontoprise GmbH) offers extensive support for this approach.

Links

semantic-mediawiki.org – main site for Semantic MediaWiki, with information on the other extensions

mediawiki.org/wiki/Extension:Semantic_Forms – extension page for Semantic Forms

referata.com – a hosting site that allows free creation of wikis that use all these extensions