Volume 52, Number 2, 2011 ACTA TECHNICA NAPOCENSIS
Electronics and Telecommunications

SHARING THE KNOWLEDGE: SEMANTIC MEDIAWIKI

Cornelia Flavia M. VEJA[1], Mircea-Florin VAIDA[1], Gregor HAGEDORN[2]
[1]*Technical University of Cluj-Napoca, Str. Memorandumului nr. 28,400114, Cluj-Napoca, Romania*
Tel: +40 264 592 055
{cornelia.veja, mircea.vaida}@com.utcluj.ro
[2]*Julius Kühn Institute - Institute for Epidemiology and Pathogen Diagnostics, Königin-Luise- Straße 19*
D-14195 Berlin, Germany
Tel: +49 308 3042220
gregor.hagedorn@jki.bund.de

Abstract: Broadly discussed in recent scientific papers, Mass Collaboration Systems (MC), enlist users who explicitly or implicitly collaborate to build a long-lasting artifact that is beneficial to the community. Wikis are systems that let users build artifacts and require users to edit and merge one another's inputs. Wiki's is a semi-closed world and the use and re-use of wiki's information by external non-humans agents are cumbersome tasks. The step from unstructured data but strongly connected information to structured data and annotated information is brought by adopting Semantic technologies into wikis. This paper presents a survey and comparative study of Semantic Wikis, with the emphasis on Semantic Media Wiki model and mechanisms and gives insight about the beneficial to use semantic wikis onto different scientific projects to share knowledge.

Keywords: *Mass Collaboration Systems, MediaWiki, Semantic MediaWiki, Semantic Web, Semantic Wiki, Social Software, Semantic Social Software, Wiki System.*

I. INTRODUCTION

Mass Collaboration systems (MC) enlist a multitude of users to explicitly and/or implicitly collaborate to build a long-lasting artifact that is beneficial to the whole community. The Web can help recruit a large number of users, enable a high degree of automation, and provide a large set of *social software* (e.g., email, wiki, discussion group, blogging, and tagging) that MC systems can use to manage their users [1].

Wikis are systems that let users build artifacts such as Wikipedia and require users to edit and merge one another's inputs.

Many scientific projects employ MediaWiki[1] as a collaborative platform to explicitly exchange knowledge. The usage of wikis as infrastructure is targeting different purposes: documentation and collaborative authoring, media sharing, learning, and entertainment.

The goal of this article is to introduce the motivation for a novel layered approach accomplished within the scope of two separate scientific projects, aiming at achieving seamless interoperability and collaboration by introducing Mass Collaboration systems, aka wikis, into design.

This article has three main contributions. Firstly, it provides an up-to-date insight into semantic wikis initiatives related to social software impact. Secondly, it presents a layered novel interoperable infrastructure focusing on the social software platforms, which allows a smooth collaboration and integration among its actors and resources. Thirdly, by comparing different initiatives in this domain, it provides recommendations towards adopting MediaWiki (MW) or Semantic MediaWiki[2] (SMW) as underlying technology to address interoperability and exchange knowledge in an evolutionary environment.

The rest of the paper is organized as follows: section II discusses about the recent term of Social Network and not only, section III points out the wiki's place from Social Software perspective and as collaborative tool. In section IV we start to present comparative Semantic wiki's features and the next section (V) emphases Semantic MediaWiki technology advantage. Afterwards, we present discussion and use-cases of MediaWiki and Semantic MediaWiki model (section VI). At the end of the paper we present conclusions and a summary of use cases as application profiles for the presented models.

II. SOCIAL SOFTWARE IMPACT

The recent use of the term 'social network' is introduced by Albert Laszlo-Barabasi and Mark Buchanan in the popular science world, and Clay Shirky and others in the social software world [3]. These authors build on the definition of the social network as 'a map of the relationships between individuals.' Other definition, following the sociologist Karin Knorr Cetina and having Jyri Engeström's3 affiliation is the term of 'object centered sociality' [2]. Approaching sociality

[1] www.mediawiki.org

[2] http://semantic-mediawiki.org
[3] http://www.zengestrom.com/blog/2005/04/why-some-social-network-services-work-and-others-dont-or-the-case-for-object-centered-sociality.html

Manuscript received February 20, 2011; revised June 15, 2011

as object-centered is to suggest that when it becomes easy to create digital instances of the object, the online services for networking on, though, and around that object will emerge too.

For many social websites, success has come from enabling communities formed around common interests, where the users are active participants who as well as consuming information also provide content and metadata. They will continue to move closer to towards simulating their real-life social interaction.

The main drawbacks of this object-centered social networks architecture are:

- the difficulty to reuse and identify common data across these sites;
- it is not possible to identify a user's contribution across all different types of social software sites.

Adding semantics will enable richer application to be built [3]. Making use of semantics is meant to connect this user's contribution across different types of social software sites.

Contrary of other forms of social software, aka blogs, which reflect the opinion of a known set of writers, wikis use an open approach whereby anyone can contribute to add value to the community [4].

III. SHARING THE KNOWLEDGE: WIKI SYSTEMS

Staring from the first definition of a wiki, given by its creator, Ward Cunningham in 1994, "The simplest online database that could possibly work" [13] and until nowadays, the main characteristics of a wiki remain the same: easy to edit and use, collaborative creation of pages, easy interlinking of the pages inside and outside the wiki.

Thereby, wiki system definition could be:

Definition: A wiki system is a form of Social Software web based platform which allows collaboratively digital content creating, maintaining and retrieving.

MediaWiki is the wiki software behind Wikipedia and Wikimedia Commons. It controls user accounts and permissions, converting "wiki markup" to HTML, storing all revisions of all articles, and more. The MediaWiki platform is a strong open content and social networking platform.

Definition: A regular wiki, as Social Software instance, is a collection of collaborative created articles having structured, formatted text (intended for humans to read and understand) and untyped hyperlinks to other related articles within the wiki (again intended for humans to follow) and serving a community.

Following the Object Orientation Programming paradigm, a wiki is an instance (object) of MediaWiki software.

As diverse as wikis can be, they commonly have a straightforward Web-based interface with a simple syntax for editing content and setting hyperlinks to other pages, and loosely restrictive access policies. Majority of wiki systems also provide a rollback mechanism in case of accidental or undesired changes.

Essential features are the support of the requirements of creative commons licenses (perpetuating licenses, tracking contributions and attributing all authors of text and media), a version management and comparison system making changes in a large community transparent to the end user, and a layered development system empowering the community to participate in the functional development of the system.

Wikis often act as informational resources (i. e. reference manual, encyclopedia or handbook). Into wiki, users can add content, edit others content, relying on cooperation, checks and balances of its members, and a belief in the sharing of ideas. This creates a community effort in resource and information management.

A wiki could be used also to set up a second location for files from an external source; known as shared uploads or shared repositories. For example, a company might have several MediaWiki sites and they need a common image repository [5].

Wiki popularity is based on several features:

- *Collaboration aspect:* all information became immediately available for everybody becomes published and visible right away.
- *Simplicity of document creation and information interlinking* - increase the re-usability of information.
- *Openness for reading and editing*: the information is open to readers as well as to editors.
- *Openness for experiments*: this archived by using modification histories for all the information.
- *Fine granularity of information* – the simplicity of document creation and interlinking encourages creation of individual pages for each topic, term, or word.
- *Refactoring*: simplicity in document creation and versioning encourage refactoring of the information; if a document became too big, it can be easily split into many smaller ones.

Wikis such Wikipedia have contained structural metadata in the form of templates and recently, semantic wikis have appeared to address the need for structured data in wikis.

Wiki limitations

Despite of the wiki success as a collaborative tool, the information accumulated in the wiki systems is unstructured, and hence, poses problems for knowledge management and productivity.

Pages cannot be automatically formatted with respect to what kind of information they contain. Sorting entities described in a wiki by arbitrary criteria requires a special extension and special templates usage.

Importing structured data from a database into the wiki and then make queries on the data as if it was still in the database is not possible. The opposite operation, expose or export arbitrary selections of the wiki content to software are tedious tasks in classic wikis. The most basic kind of reasoning cannot be automatically performed inside a traditional wiki, even though the information is often present (but only in a human readable format) [15].

The classic approach for dealing with this issue is for the information to be entered into structured forms or templates thereby becoming much clearer to humans and more accessible to computer aided operations, such as structure-based search, data integration, and data analysis. This approach assumes data allow to be structured. For the unstructured data, semantic systems use the inverse paradigm: not forcing information into given form, as opposed, information into free document format is labeled with semantic tags.

We present in our previous works [4][5] use cases for both of these methods for structuring information into wiki pages. Both methods are in place, the choice is depending on the purpose of wiki usage and the user's level of technical expertise.

IV. SEMANTIC COLLABORATIVE PLATFORMS. SEMANTIC WIKI IMPLEMENTATIONS

Several Semantic wikis are built over time as Semantic Social Software application. In the following section we evaluate only active projects and a selection of the most popular was made.

In order to understand better what can be called a "Semantic wiki" an exploration of available Semantic wiki implementation was conducted. Through exploration, different characteristics, patterns, similarities and differences were identified. A list of available "state of the art" Semantic wikis was retrieved from semanticweb.org5.

Wiki Systems

AceWiki[4] [18] is a slightly unusual wiki which uses the controlled natural language Attempto Controlled English (ACE)[5]. Formal statements are the main content of the wiki itself. In this way, it tries to integrate ontology, rules and query language into one. The editor allows the user to either directly type in the statements or use a guided form of selecting from the existing ontology. AceWiki conceptualizes each wiki page as a concept and produces an OWL output of the underlying ontology.

Kiwi[6], formerly, IkeWiki [10] aims to create instance data based on an existing ontology but also to be a (limited) tool for creating and editing ontologies. Kiwi tries to keep the look and feel of popular wiki systems, such as Mediawiki. It makes use of the recommended Semantic Web technologies RDF and OWL. The presentation of the information makes use of the formal knowledge structure to improve navigation and provide recommendations while editing.

KawaWiki [17] aims to provide a complete formal structure of the data with proper use of RDF and RDFS. The architecture breaks down into three main layers:

- RDFS – The RDF Schema defines the underlying ontology and is used to validate the RDF Templates.
- RDF Templates – Defines the type of wiki pages that can be created and authored by the end user.
- Wiki Content – The wiki content is what the end user gets to edit. Templates and all that is required is the filling of the template fields only.

OntoWiki[7][11] is a form-based semantic wiki application. OntoWiki is more of a collaborative ontology editor rather than a Semantic wiki. It does not have the familiar wiki interface of entering natural language text informally to represent a concept but supports several collaborative features and also allows the installation of plug-ins. The OntoWiki implementation conceptualizes each page as a resource, storing triple statements in an RDF store.

The basis of the implementation is the Powl[8] framework.

SMW[9] [9] (Semantic MediaWiki) is an extension to Mediawiki, enabling it to become a Semantic wiki. It enhances WikiML annotations to allow users to include relations and properties to wiki pages. Semantic Mediawiki conceptualizes pages as concepts [14] and stores the semantic data in Mediawiki's MySQL database, but it can also be exported as RDF. Further presentation including semantic search and inline queries is detailed later.

KnowWE[10] is an extension of the JSPWiki adding semantic functionality to it. Its parsing engines and problem-solvers also build on the d3web project[11]. Each page is a concept in the controlled ontology. Semantics are included in the wiki markup and it provides three alternatives: explicit knowledge markup, semantic annotation, segmented text. Text is annotated by ontological content. Additionally, problem-solving knowledge within the text can be made explicit using, for example, rules or set-covering models. The authoring interface attempts usability improvement by using AJAX to provide suggestions. Additionally, knowledge exchange via OWL ontologies is provided.

Tiki Wiki CMS Groupware[12] is one of the most feature-rich CMS packages and allows defining some semantic relationships between wiki pages. User interface could be personalized using themes, based on Smarty templates markup[13].

Knoodl[14] is a web-based, collaborative ontology editor. Each resource is ontology and has its own web-page which is half structured content from the ontology and half unstructured content in the form of wikitext. Content in Knoodl is organized into Communities, which can be created by any user. Communities have a role-based permissions model. Ontologies can be imported and exported as OWL files, with or without the associated wikitext. Its underlying RDF store is Mulgara.

Comparison

In order to put what each Semantic wiki has to offer into perspective two dimensions were devised to compare the Semantic and Wiki aspects: *user perspective* and *knowledge expressivity*.

User perspective (x axis): This dimension refers to how much technical skill the user needs to have in order to use the wiki and contribute to the ontology. It divides the scale into a number of broad categories of users:

- Everyday user – A user familiar with the use of specific applications, without administrative, modeling or programming skills.
- Power user – A user familiar with the use of a wider range of applications, administration of their own computer, but with no modeling or programming skills.
- Professional user – An advanced user familiar with applications, administration, and with modeling and programming knowledge but no knowledge of Semantic Web technologies (such as OWL and RDF).
- Ontologist – Ontology expert user who knows how to use Semantic Web technologies, particularly ontologies.

Knowledge expressivity (y axis): This dimension refers to how expressive the final ontology is. The expressivity scale is

[4] http://attempto.ifi.uzh.ch/acewiki
[5] http://attempto.ifi.uzh.ch
[6] http://www.kiwi-project.eu/
[7] http://ontowiki.eu/
[8] http://powl.sourceforge.net/overview.php
[9] http://semantic-mediawiki.org/
[10] http://semanticweb.org/wiki/KnowWE
[11] http://www.d3web.de/
[12] http://doc.tiki.org/Semantic
[13] http://www.smarty.net/
[14] http://knoodl.com/ui/home.html

divided into broad levels of expressivity and formalism.

- Simple taxonomy – A taxonomic classification of the wiki pages with no added semantics about relations between pages and concepts.
- Relations between concepts – Concepts within the wiki are linked to each other using semantic annotations.
- OWL Lite level – Better formality than simple relations with some restrictions similar to the functionality offered by OWL Lite.
- OWL DL, OWL Full level – functionality similar to that of OWL DL and OWL Full, allowing for more restrictions and expressive relations with property characteristics such as transitive, functional, disjoint, etc.

Each Semantic wiki was approximately positioned in the graph of the two dimensions explained above, as shown in following figure, using the available information found online for each one [*figure 1*].

It is useful to see the assumptions about knowledge expertise taken by the Semantic Wikis, and to see them in context with a more traditional MW as well known and most used wiki, which was added only for comparing purpose.

Wikis on the left side of the graph restrict user input and require little user knowledge (such as AceWiki and KawaWiki). However they require users with ontology expertise to set up, monitor and help grow the wiki.

The wikis on the right of the graph (KiwiWiki, KnowWE and to some degree OntoWiki) although quite expressive are hard to use by non-expert users, as they require knowledge of Semantic Web technologies or elaborate special syntax. Knoodl is at the right extremity and requires ontologist's skills for creating ontologies.

Wikis in the middle of the graph try to balance out expressivity and required knowledge, such as SMW and Tiki Wiki CMS Groupware.

From the methodological point of view in adding semantic to wikis, two types of semantic wikis can be distinguished:

- text-centered approaches that enrich classical wiki environments with semantic annotations (Semantic MediaWiki);
- logic-centered approaches use semantic wikis as a form of online ontology editors (OntoWiki, Knoodl).

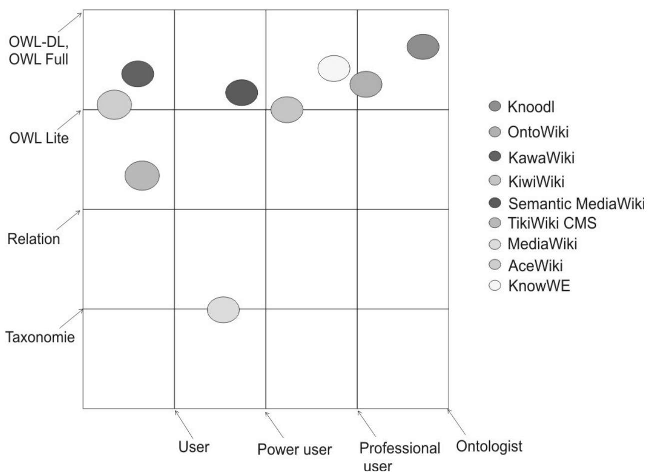

Figure 1. The semantic wikis comparator.

The dynamics of wiki's projects, reflected onto semanticweb.org5 couldn't make possible a long-term analysis and prediction of further development of today alive wikis. We observed in time, only few of them still remain on the list. The well-known and mostly used is Semantic MediaWiki.

This comparison between semantic wikis reveals the motivation for choosing MediaWiki and Semantic MediaWiki as underlying technology to address interoperability in heterogeneous environments that need to exchange and share knowledge:

- *user perspective*: they are text-centered and easy to learn and use;
- *knowledge expressivity perspective*: represent a balanced solution between expressivity and required knowledge.

V. SEMANTIC MEDIAWIKI AS SEMANTIC SOCIAL SOFTWARE

Semantic MediaWiki [9] is a semantically enhanced Wiki engine that enables users to annotate the wiki's contents with explicit, machine-readable information. It supports adding structured and semantically annotated information into wikis using a specific syntax. Using this semantic data, SMW addresses core problems of today's wikis:

- *Content consistency:* The same information often occurs on many pages.
- *Knowledge Access:* Finding and comparing information from different pages is a challenging and time-consuming task in large wikis.
- *Knowledge reuse:* The rigid, text-based content of classical wikis can only be used by reading pages in a browser or similar application.

SMW is based on a simple and unobtrusive mechanism for semantic annotation. Users provide special markup within a page's wiki-text, and SMW unambiguously maps those annotations into a formal description using the OWL DL ontology language.

Dynamically generated content are possible by embedding queries into the wiki pages. Import of external data from existing ontologies like FOAF, SIOC or Doublin Core is now possible by mapping wiki-annotations to elements from these vocabularies.

To enable external reuse, formal descriptions for one or more articles can be obtained via a web interface in OWL/RDF format. Since SMW strictly adheres to the OWL DL standard, the exported information can be reused in a variety of tools.

Semantic properties added to several wiki templates enriching page's content with structured information, better re-usable, searchable, and that is, allowing inference.

VI. MEDIAWIKI AND SMW DATA MODEL

MediaWiki [4] provides an object oriented document storage model of medium granularity (titled chapters called "pages", rather than whole works). The storage model is akin in many aspects to the currently developed "nosql" database management systems[15] (predating these developments, however, MediaWiki typically uses mysql).

Thereby, knowledge is created and maintained in an

[15] http://nosql-database.org/

evolutionary way through a web–based interface by the community itself. In general, wikis are comparable with content–management systems by building, creating, and sharing knowledge (i.e., text, figure, and multimedia), but Wikis provide a much simpler and more open interface of the engineering process.

MediaWiki engine

The layered architecture of MediaWiki engine is built on LAMP architecture (PHP on server side and Javascript on client side, with MySQL as underlying database). The "core" of MediaWiki completed after the installation and configuration, supplying the basic functionality and wiki-like behavior. It allows pages creation and editing, categorization and interlinking, review, maintenance functions, user's management and statistics. From information structure point of view, MediaWiki provides a template system and the possibility of page categorization.

From the vibrant community of developer's perspective formed around emergent wiki, MediaWiki functional layers imply also layered wiki development:

- Base software (the core) and a complet API set provided for external programmatically access to basic MediaWiki functionalities;
- Extensions – add extra functionalities to MediaWiki's core: enhanced edit functions, enhanced user interfaces and semantic functionalities. Usually extensions are developed in PHP, using hook mechanism for exploiting MediaWiki events.
- Templates – allow wiki patterns creation aiming to mitigate wiki information inconsistency and a flexible schema development.
- Content creation – using the output of previous layers and effectively creation of MediaWiki instance, aka, a wiki site.

This layered structure of collaboration in MediaWiki's development induces layered human participation in development: core software developers, extension developers, template creators and content creators. From Social Software perspective, the „social object" in MediaWiki is MediaWiki itself [5].

From OOP perspective, creation of content instantiates MediaWiki software and creates a wiki site.

The basic unit in a MediaWiki (MW) site is a wiki page. Into wiki, the information could be kept at informal, semi-formal or formal level, depending on the wiki's community of interest and the collaborators implied in content creation.

MediaWiki provides several mechanisms for creating structure within the wiki. At this level data in wiki is structured, semi structured, or unstructured. By default, there is no required structure - but that tends to make the content very difficult to navigate. So sites using MediaWiki often build their own structure. The mechanisms are:

Namespaces. Pages in MediaWiki are divided into namespaces. Namespaces provided by the storage model allow re-using the basic model for 1st-class content objects as well as building objects used in hypertext inclusion. Examples for the latter are media items (binary plus metadata) in the "File" namespace or programming blocks and rich text fragments in the "Template" namespace [4].

Categories. Categories (pages in the Category namespace) can be used to automatically group pages. This mechanisms provides for slightly schema development into taxonomies.. Categories supply class definition from OO paradigm and the pages which belong to are instances.

Templates. MediaWiki provides a namespace devoted to transclusion, called Template. Articles in the Template namespace, known as templates, are designed to be transcluded into other articles. This mechanism is used for avoid inconsistency in wiki. This is the common way for structuring data.

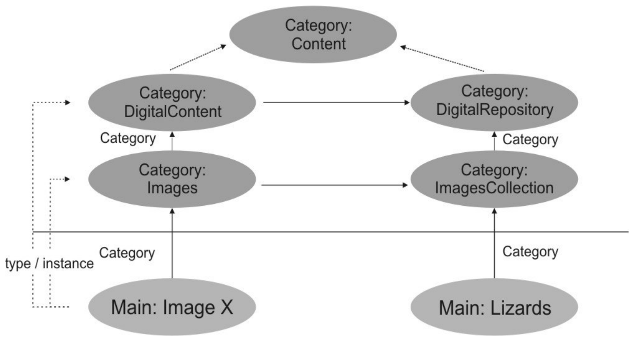

Figure 2. MediaWiki category mechanism of class inclusions

MediaWiki templates mechanism

The template model provides for flexible schema development. Each template defines a class with freely definable attributes (equivalent to an "entity type"), instances of which can be freely embedded into other objects. Template instances can be hierarchically nested.

Wiki content templating mechanisms which have been available since the first wiki clones under the name of seeding pages. The key idea of this work is indeed expressed by the intuition:

wiki template = information pattern: (content) templates

are used in wikis to give authors skeletons (or patterns) to be filled with precise information. Common use cases of templates are coherent with this intuition:

- information and navigation boxes;
- seeding pages are used to foster the creation of pages which will describe similar items.

Modeling with templates

Study state of the art wiki templating mechanisms in order to understand if and how they can be used for the task of fostering the production of semantically rich web pages. In doing so we identify two alternatives templating models which can be found in state of the art wiki engines: functional templating (in which templates are invoked by name and passed parameters) and creational templating (in which templates are simply copied as new content at the beginning of a page revision history).

Unlike "real'' CMSs, where templates are treated as special resources, wikis consider templates and instances at the same level, and allow users to manage both kinds in (almost) the same way. Two kinds of pages are involved in template processing [7]:

- template (e.g. MetadataTemplate): the master page on which other pages can be based;
- template instance (e.g. AlgaePage): the result page obtained by applying a master page to a target one.

In functional templates, several features have been highlighted:

- strong and permanent connection between a template and its instances;
- different source code of templates and instances;
- template instances is the non - linearity of the markup with respect to its rendered form.

In templates, data are structured per-se. Despite of this aspect, data embedded into template instances are not easy to be extracted and reused. To accomplish this task, the literature describes heuristic methodologies.

One of these methods is used by DBpedia[16] project which is a community effort to extract structured information from Wikipedia and to make this information accessible on the web.

Many Wikipedia pages contain templates often for layout purposes, but still approximately a quarter to one third of the Wikipedia pages already today contain valuable structured information for querying and machine interpretation. A special type of templates is infoboxes, aiming at generating consistently-formatted boxes for certain content in articles describing instances of a specific type. To reveal the semantics encoded in templates, [8] developed an extraction algorithm operating in five stages.

The gained experience has been valorized by authors into a set of guideline and suggestions for wiki templates built. By following these guidelines is not only good for semantic extraction, but usually also improves the corresponding template in general, i.e. it becomes more convenient to use by article authors.

In [14], authors mention the hypothetical role of templates in semantic rules definition to knowledge representation in MediaWiki. They pointed out that it is possible to switch on or off the entailment rules on individual pages, or a group of pages using a template-based strategy. In general, this approach will render sound, but not necessarily complete inference results.

SEMANTIC MEDIAWIKI DATA MODEL

Semantic MediaWiki (SMW) is based on a simple and unobtrusive mechanism for semantic annotation. Users provide special markup within a page's wiki-text, and SMW unambiguously maps those annotations into a formal description using the OWL DL ontology language.

Semantic MediaWiki is in use on over 216 public active wikis around the world, in addition to an unknown number of private wikis.

Annotations Syntax

MW has a scripting language to describe contents of wiki pages, which is extended by SMW with the following three basic sets of semantic annotations:

Class and Property Definitions: SMW reuses the "Category" namespace to define classes. For example, a wiki page with the name "Category:Images" is intended to represent the class of all images. Likewise, a new namespace called "Property" is introduced to define concept properties. Through this approach, SMW supports both binary and n-ary properties.

Axioms: SMW allows for the declaration of subclass and subproperty relations. For example, on the page "Category:Images", the annotation "[[Category:DigitalContent]]" would indicate that Images is a subclass of DigitalContent. SMW also supports equality relations between two wiki pages, or between two classes, via the redirection mechanism of MediaWiki.

SMW	RDF
Category:x	rdfs:Class
Property:x	rdf:Property
[[Category:x]] (on category page)	rdfs:subClassOf
[[Category:x]]	rdf:type
[[Subproperty Of::x]]	rdfs:subPropertyOf
Property:x	rdf:Property
#REDIRECT [[x]] #REDIRECT [[Category:x]] #REDIRECT [[Property:x]]	owl:sameAs, owl:equivalentClassOf, owl:equivalentPropertyOf (non-transitive)

Figure 3. Correspondence between SMW-ML (version 1.4.2) and RDFS/OWL

Instance Assertions: SMW allows a page to be declared as an instance of a class or the subject of an RDF triple. For example, if the wiki page "Galium odoratum" contains the annotation "[[Category:Images]] [[Type::StillImage]]", this means that Galium odoratum is a digital resource and he has the type "StillImage". It is also possible to create an instance of an n-ary property. For instance, on the wiki page "Galium odoratum" the annotation "[[General KeyWords::Galium odoratum; Tracheophyta]]" would mean that the ScientificName is Galium odoratum and having botanical Subject Category: Tracheophyta.

Interpretations of the Semantic Annotations

From the functionality point of view, this mechanism of semantic annotations presents three facets [Figure 5]:

- *Hypertext Link Functionality* inherited from MW as linked pages;
- *Knowledge Representation* Functionality into semantic net;
- *Data Base Functionality*, allowing n-tuples representation of properties SQL-like style.

Data Types for Property Values

The semantic annotations, named semantic properties in SMW are typed. SMW use an auto-describing mechanism for property assigned type. Thereby, every semantic property has the own wiki page. Exemplifying this, property "Provider" wiki page might contain the annotation:

[[has type::Page]]

where "has type" property is recognized by SMW as special property. Current property types supported are: String, Number, Annotation URI, Boolean, Email, Text, Geographic Coordinates etc.

---

16 http://dbpedia.org

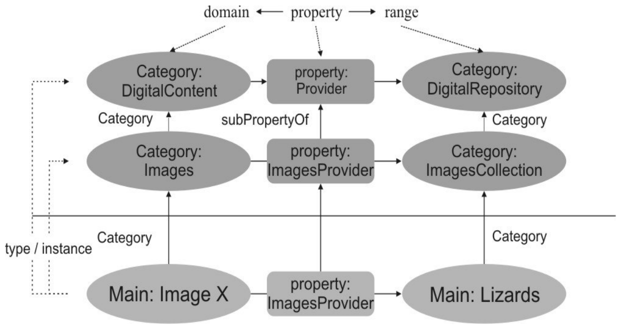

Figure 4. Semantic MediaWiki annotation functional model.

Data Aggregation Semantic Forms
Templates themselves are wiki pages in a special namespace, representing master-pages with. Adopting templates is useful for structured user input. An aggregation of properties into templates and then, an aggregation of templates represent semantic forms.

Examples:

Property: CollectionName | hasType: String
Property: BestQualityURI | hasType: URL
Property: Type| hasType: String

Template:
{{Metadata | CollectionName = | BestQualityURI = | Type = }}

Form:
{{#forminput:Form-MetadataEntry
{{{for Metadata
{| class="formtable" | Collection Name: | {{{field | CollectionName}}}
…
{{{end template}}}
}}

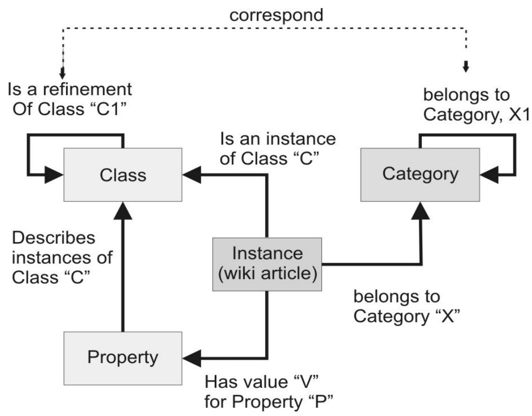

Figure 5. Semantic MediaWiki Ontology.

A SMW ontology intuitive diagram is depicted in the [*figure 5*].

For information retrieval, SMW uses inline queries. Semantic properties and inline queries, along with semantic forms and templates represents a powerful tool that can support modeling for different project's needs.

V. RELATED WORKS AND USE CASES

Related works are discussing other dimensions when they are comparing semantic wikis. In [15], the authors have pointed out perspectives like requirements engineering for semantic wikis. With respect to [16] three dimensions can be used to categorize a wiki. These dimensions are induced by the operation of semantic: authoring, knowledge representation and user presentation. We choose user perspective and knowledge perspective because they are carrying the mainly features in a semantic collaborative platform. The authors of [20] have conducted a survey with the emphasis on the usability and mentioned the differences between MW and SMW. From the non-technical skilled user perspective, they concluded that learning curve for MW is lower than for SMW.

Many scientific projects employ MW and SMW as underlying technologies to support research in different domains. To enlist several use-cases: usage in eLearning environments (Aposdle)[17], infrastructure support for agile IT development [21], distributed military Coalition Planning [21], e-Government infrastructures [19], e-Health, Bio-Informatics[18] and Virtual Research Environments in Humanities, like "Semantic MediaWiki for Collaborative Corpora Analysis" (SMW-CorA)[19].

Started from different perspectives as the previous works and also considering the comparative study we accomplish here, asserts:

- The learning curve for MW is lower than for SMW and the appropriate use-case are the environments where actors present different level of technical expertise and make use of external storage system for metadata repository;
- For heterogeneous data integration and middle level of user's technical expertise, who need intuitive use of semantic web technologies, the appropriate use-case requires SMW.

The article's authors first approach, making use of MW category mechanism and metadata attached, gives semantic dimension to common functional templates (rich templates) [5]. This heuristic methodology, first applied into Key To Nature[20] project, introduces the automatic workflow for metadata aggregation and ingestion [6]. Several steps have been developed for accomplish automatic metadata management and MediaWiki web-based repository layer:

- Data providers submit metadata trough wiki pages following the wiki instance template syntax (e.g. Metadata). The harvestable wiki page corresponds to a digital repository collection.
- Extract and select metadata from templates based on specific patterns.
- Parse each template and generate appropriate XML format. A URL derived from the title of the wiki page the template occurs in is used as feedback mechanism to data providers.
- Post-process object values to generate digital object model in a form of RDF datastream. At the repository level, the

[17] http://www.aposdle.tugraz.at/
[18] http://offene-naturfuehrer.de/
[19] http://km.aifb.kit.edu/projects/cora/index.php
[20] http://www.keytonature.eu

logical object maintained is the collection [6].

- Collections can be interpreted as classes subsuming instances described by pages in the corresponding category. Every collection belongs to a data provider. The category mechanism has been used to fulfill the feedback reports to data providers.

The second approach uses SMW technologies: wiki templates, semantic properties, RDF export feature and semantic in-line queries. The model implies a harvesting method which automatically collects the metadata generates by a web-based tool used by biologists, an identification keys editor (jKey editor). Single-access identification keys can be stored directly in wiki objects [5].

VI. CONCLUSIONS AND OUTLOOK

MediaWiki and Semantic MediaWiki, as sampling of explicit Mass Collaboration systems, are meant to share knowledge. Although the embedded knowledge inside wiki is easily accessed by humans, this is not the case when machines should re-use and retrieve wiki's information. The known solutions are as follows:

- Make use of templates;
- Enabling semantic wiki's information by adopting semantic technologies.

The article presents an up-to-date study of Semantic wikis with emphasis on Semantic MediaWiki. We chose SMW because present a balanced solution for enabling semantic in wikis. We present also a hybrid method for structuring wiki´s information which combines semantic templates along with automatic metadata repository workflow management.

From the underlying technology point of view, wikis are applications profiles for a common model, Semantic MediaWiki model.

At the end of the article, several use-cases and applications using SMW have been mentioned. We summarized the application profiles for MW and SMW detailed in other previous works:

- An online authoring and publishing platform [4]
- Usage of MediaWiki templates as white-space invariant for involving non-technical users in the whole process of digital content management [6]
- Usage of Semantic MediaWiki model for metadata management inside wikis [5].

Ongoing and future works focus on strategies for heterogeneous metadata integration using SMW as underlying technologies and having the aim to build a VRE in humanities. SMW has been demonstrated to be reliable solution for the metadata integration process from a digital library to SMW-CorA. This solution address research needs and mapping with Semantic Web technologies.

Another research direction is aiming to prove the usefulness and effectiveness of MW and SMW technologies for metadata repository management, starting from user perspective and considering the automatic feedback provided by metadata repository ingest tool. This task is carrying out by using SWM technologies as support.

Acknowledgment

This work has been carried out in the frame of the Social European Found – FSE – in the frame of POSDRU/6/1.5/S/5 project.

REFERENCES

[1] A. Doan, R. Ramakrishnan, A. Y. Halevy, *Crowdsourcing Systems on the World-Wide Web*, Communications of the ACM, 2011, doi :10.1145/1924421.1924442

[2] Karin Knorr Cetina. *The practice turn in contemporary theory*, (London 2001: Routledge).

[3] J. Breslin, Al. Passant, St. Deker. *The Social Semantic Web*, Springer, 2009.

[4] Hagedorn, G., Weber, G., Plank, A. Giurgiu, M. Homodi, A., C. Veja, and others,, '*An online authoring and publishing platform for field guides and identification tools'*, Proc. of Conf. Tools for Identifying Biodiversity: Progress and Problems, Paris, September 2010 – pp. 13-18

[5] C. Veja, G. Hagedorn, G. Weber, M. Giurgiu. *'Semantic MediaWiki Interoperability Framework from a Semantic Social Software Perspective'*, Proc. of 9th International Symposium on Electronics and Telecommunications ISETC 2010, Timisoara, November 2010 , pp. 403 – 406, ISBN: 978-1-4244-8460-7, IEEE DOI=10.1109/ISETC.2010.5679307.

[6] C. Veja, G. Hagedorn, G. Weber, M. Giurgiu 'Metadata Repository Management using the MediaWiki Interoperability Framework A Case Study:The KeyToNature Project', Proc. Of Int. Conf. eChallenges e-2010, Warsaw, Poland, IEEE, ISBN: 978-1-4244-8390-7, 2010.

[7] A. Di Iorio, F. Vitali, St. Zacchiroli, *'Wiki Content Templating'*, WWW 2008, April 21–25, 2008, Beijing, China. ACM 978-1-60558-085-2/08/04.

[8] S. Auer, J. Lehmann. *'What have Innsbruck and Leipzig in common? Extracting semantics from wiki content'*. Proc. 4th European Semantic Web Conference (ESWC), 2007.

[9] M. Krötzsch, D. Vrandecic, M. Völkel, H. Haller, R. Studer. '*Semantic Wikipedia'*. In Journal of Web Semantics 5/2007, pp. 251-261. Elsevier, 2007.

[10] S. Schaffert, J. Eder, S. Grünwald, T. Kurz, M. Radulescu, R. Sint, S. Stroka. '*KiWi – a platform for semantic social software*'. In Proc. of the 4th Workshop on Semantic Wikis, ESW C 2009, Hersonissos, Greece, 2009.

[11] S. Auer, S. Dietzold, Th. Riechert. '*OntoWiki- a tool for social, semantic collaboratin*', Proc. of the 5th Int. Semantic Web Conference, in Lecture Notes in Computer, Science, pp. 736-749. Springer, 2006.

[12] D. Vrandečić, M. Krötzsch, ‚*Semantic MediaWiki. In Davies, J et al. (eds.), Semantic Knowledge Management'*, Springer-Verlag Berlin Heidelberg 2009, pp. 171-179

[13] Leuf, B., and Cunningham, W. *The Wiki Way Quick Collaboration on the Web,* 2001.

[14] J. Bao, L. Ding, J. Hendler. *Knowledge Representation and Query in Semantic MediaWiki: A Formal Study.* Revision: August 12, 2009 http://tw.rpi.edu/wiki/TW-2008-42

[15] B. Hoenderboom, P. Liang. '*A Survey of Semantic Wikis for Requirements Engineering',* RUG-SEARCH-09-L03, 2008.

[16] Kousetti, Ch., Millard, I., Yvonne Howard. '*A Study of Ontology Convergence in a Semantic Wiki'*. Based Infrastructure. University of Southampton. UK, 2008.

[17] K. Kawamoto, Y. Kitamura, Y. Tijerino. '*KawaWiki: A Template-Based SemanticWiki Where End and Expert Users Collaborate', 2008.*

[18] T. Kuhn. '*How Controlled English can Improve Semantic Wiki'*. arXiv:0907.1245v1 [cs.HC] 7 Jul 2009.

[19] D. Vrandecic, Ch. Lange, M. Hausenblas, J. Bao, L. Ding, '*Semantics of Governmental Statistics Data'*, Web Science Conf. 2010, April 26-27, 2010, Raleigh, NC, USA.

[20] K. M. Hüner, B. Otto. ‚*The Effect of Using a Semantic Wiki for Metadata Management:A Controlled Experimen',* DOI: 978-0-7695-3450-3/09,IEEE, 2009.

[21] F. Kleiner, A. Abecker, 'Towards a Collaborative Semantic Wiki-based Approach to IT Service Managemen't, 2008.

[22] D. Braines, P. Smart, J. Bao, 'Emergent Capabilities for Collaborative Teams', in the Evolving Web Environment 3rd Annual Conference of ITA (ACITA), 2010.