KiWi
Collaborating the Semantic Wiki Way[*]

Stephanie Stroka, Szabolcs Grünwald, Thomas Kurz, Mihai Radulescu,
Sebastian Schaffert, Rolf Sint

Salzburg Research Forschungsgesellschaft
Jakob Haringer Str. 5/II, 5020 Salzburg, Austria
<*firstname*>.<*lastname*>@salzburgresearch.at

Abstract. Since the last years the way people use the web has radically changed. The world wide web turned into one great knowledge base, fed by people who share their knowledge with the world. Wikipedia is the most familiar case example and wikis in general became one of the most applied social applications on the web. They provide easy and efficient platforms for collaboration and documentation, but introduce one major problem: chaos.
The intention of the 'KiWi' project (the name originally meant 'Knowledge in a Wiki'), is not to completely take back the chaos and introduce hard structure schemas, but rather to enhance content with semi-structure and semantics. KiWi is an EU project in the 7th Framework Programme and aims to result in a social-semantically enhanced platform, which easily allows to build extensions for the Semantic Web on top of it [4]. KiWi is build with the JBoss Seam Framework[1] and licensed under the New BSD license. In this paper, we would like to present the main features of the KiWi framework and motivate the Open Source community to use and further develop the project.

1 Chaotic, but Structured

KiWi follows the wiki philosophy by allowing multiple user content creation and modification in an unstructured way. Users may create and share their knowledge in a creative, but chaotic way. Authors are not restricted to a specific schema, which is close to how people communicate with each other.

The drawback of chaos is that content might get lost, like on a chaotically organized desk, where a bunch of hand-written documents have been spread all over the place. In a wiki, the only way to find information is to do a fulltext search over the whole content or to browse articles via links. A well-structured wiki, e.g. Wikipedia, demands that moderators attend to structure content by arranging it into categories by hand. In the end, information in a Wiki-like environment

[*] The research leading to these results is part of the project "KiWi - Knowledge in a Wiki" and has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 211932.

[1] http://seamframework.org/

is currently put into context with the help of human reasoning (by reading and interpreting its content).

To support the (semi-)automatic interpretation, and therefore the classification, searching and automatic reasoning over content, the data in KiWi is enhanced with semantics. This is, on the one hand, realized through social tagging (folksonomy[2]) and, on the other hand, by adding metadata in the Resource Description Framework (RDF) format [3]. Each content item is stored with these collaboratively created metadata, which enables to extend data with semi-structured information, e.g. typification, semantic linking, subclassing, etc, as can be seen in Figure 1. The advantage of RDF in contrast to storing such metadata in a relational database or in XML is obvious: Data can be interpreted by every application, since the structure is based on one or more underlying ontologies. and we are (still) not restricted to a certain schema, because the RDF implementation is intentionally allowed to be incomplete and/or inconsistent. This is the way the web works.

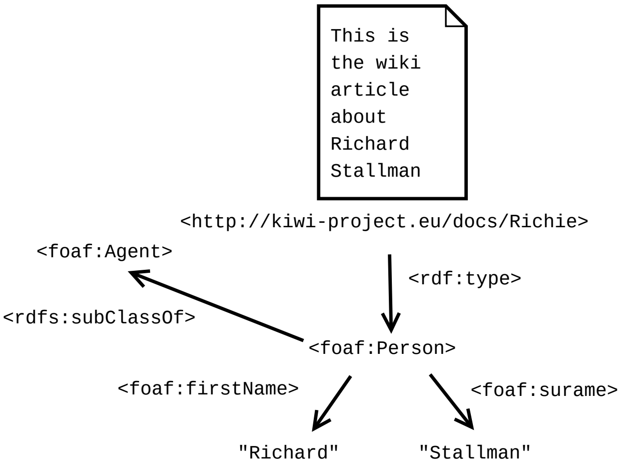


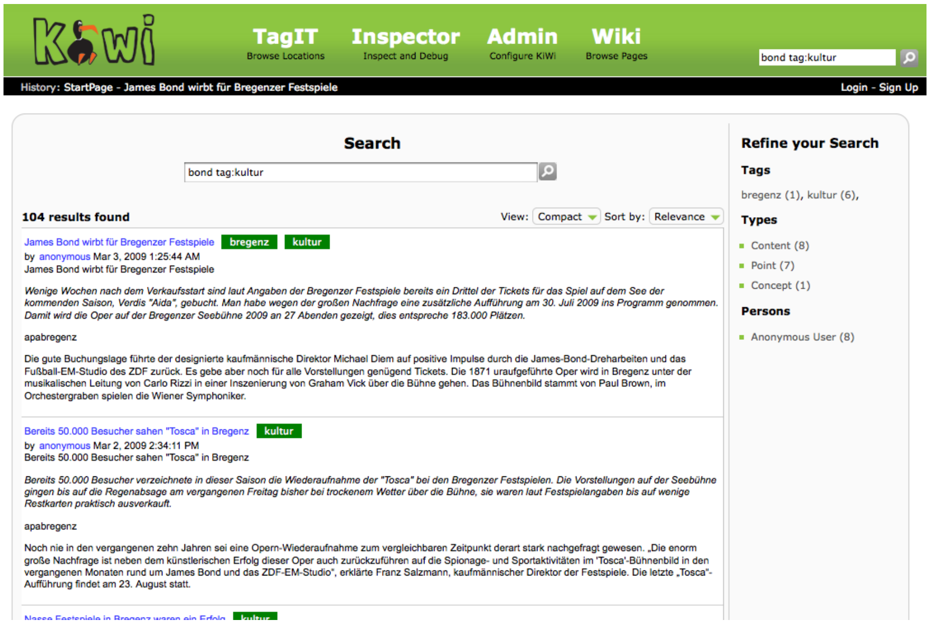
Fig. 1: Enhancing wiki content with RDF metadata

Semantically enhanced data can be used for machine reasoning (creating more metadata from an existing dataset or warning against found inconsistencies), information retrieval and extraction, faceted search, and more. Figure 2 gives an example of KiWi‘s search interface. On the right hand side of the interface, the user is able to refine his search. An alternative search framework based on KiWi can be used with the Keyword Query language (KWQL), which is currently implemented within the project. One may, for example, use the query `ci(tag(name:Java) link(target:ci(title:XML)))` to query for a

[2] `http://en.wikipedia.org/wiki/Folksonomy`

ContentItem (ci), which is tagged with `Java` and links to another ContentItem entitled `XML` [2].

Faceted search is not completely new, but past applications, which exclusively build on relational databases just allow to refine the search result with predefined facets. In KiWi, on the other hand, one can just create a new ontology scheme, load it into the application and annotate documents with a type from the ontology. The search can then be refined to just return results for documents which implement this type.



Search

bond tag:kultur

104 results found — View: Compact — Sort by: Relevance

James Bond wirbt für Bregenzer Festspiele — bregenz — kultur
by anonymous Mar 3, 2009 1:25:44 AM
James Bond wirbt für Bregenzer Festspiele

Wenige Wochen nach dem Verkaufsstart sind laut Angaben der Bregenzer Festspiele bereits ein Drittel der Tickets für das Spiel auf dem See der kommenden Saison, Verdis "Aida", gebucht. Man habe wegen der großen Nachfrage eine zusätzliche Aufführung am 30. Juli 2009 ins Programm genommen. Damit wird die Oper auf der Bregenzer Seebühne 2009 an 27 Abenden gezeigt, dies entspreche 183.000 Plätzen.

apabregenz

Die gute Buchungslage führte der designierte kaufmännische Direktor Michael Diem auf positive Impulse durch die James-Bond-Dreharbeiten und das Fußball-EM-Studio des ZDF zurück. Es gebe aber noch für alle Vorstellungen genügend Tickets. Die 1871 uraufgeführte Oper wird in Bregenz unter der musikalischen Leitung von Carlo Rizzi in einer Inszenierung von Graham Vick über die Bühne gehen. Das Bühnenbild stammt von Paul Brown, im Orchestergraben spielen die Wiener Symphoniker.

Bereits 50.000 Besucher sahen "Tosca" in Bregenz — kultur
by anonymous Mar 2, 2009 2:34:11 PM
Bereits 50.000 Besucher sahen "Tosca" in Bregenz

Bereits 50.000 Besucher verzeichnete in dieser Saison die Wiederaufnahme der "Tosca" bei den Bregenzer Festspielen. Die Vorstellungen auf der Seebühne gingen bis auf die Regenabsage am vergangenen Freitag bisher bei trockenem Wetter über die Bühne, sie waren laut Festspielangaben bis auf wenige Restkarten praktisch ausverkauft.

apabregenz

Noch nie in den vergangenen zehn Jahren sei eine Opern-Wiederaufnahme zum vergleichbaren Zeitpunkt derart stark nachgefragt gewesen. „Die enorm große Nachfrage ist neben dem künstlerischen Erfolg dieser Oper auch zurückzuführen auf die Spionage- und Sportaktivitäten im 'Tosca'-Bühnenbild in den vergangenen Monaten rund um James Bond und das ZDF-EM-Studio", erklärte Franz Salzmann, kaufmännischer Direktor der Festspiele. Die letzte „Tosca"-Aufführung findet am 23. August statt.

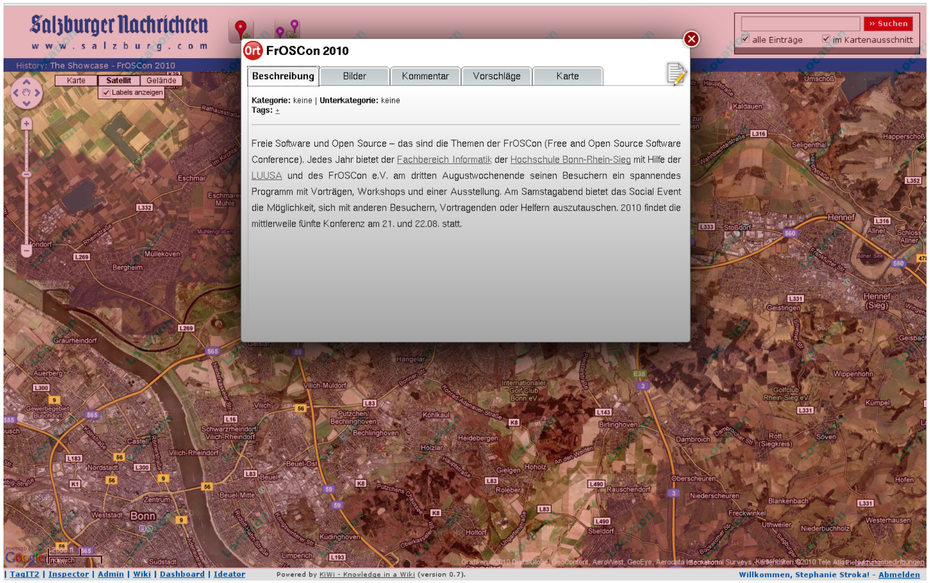
Nasse Festspiele in Bregenz waren ein Erfolg — kultur

Fig. 2: KiWi Search: generic, faceted search over KiWi content

2 One Content, but Versatile

KiWi is in fact just a semantic index store. We implemented some extensions to show the application cases of semantically enhanced data. TagIT (Figure 3), for example, is an extension on top of KiWi, which connects a Wiki article with a geographical point. Every content that is of the type `tagit:PointOfInterest` can be viewed in the TagIT application (additionally to the Wiki application view).

It is intended to let the system decide which view is used to display the content based on the system configuration. E.g. if a content is of the type `foaf:Person` it will be displayed as a user profile as it is common in social networks. It would be possible to show an *Add a Friend* widget only if the current content item is of type `foaf:Person`.



FrOSCon 2010

Kategorie: keine | **Unterkategorie:** keine
Tags: +

Freie Software und Open Source – das sind die Themen der FrOSCon (Free and Open Source Software Conference). Jedes Jahr bietet der Fachbereich Informatik der Hochschule Bonn-Rhein-Sieg mit Hilfe der LUUSA und des FrOSCon e.V. am dritten Augustwochenende seinen Besuchern ein spannendes Programm mit Vorträgen, Workshops und einer Ausstellung. Am Samstagabend bietet das Social Event die Möglichkeit, sich mit anderen Besuchern, Vortragenden oder Helfern auszutauschen. 2010 findet die mittlerweile fünfte Konferenz am 21. und 22.08. statt.

Powered by KiWi - Knowledge in a Wiki (version 0.7).

Fig. 3: TagIT - KiWi GeoLocation Extension

Every extension stores the textual or media content, the title, the creation and modification dates and the author in the same database and further data as RDF triples. KiWi‘s core system implements an abstraction layer, which is used to make the development of additional data entities easier by preventing that developers have to construct RDF triples by themselves. Listing 2 shows how an abstract RDF entity might look like. The interface that is shown in this listing is wrapped around a proxy object and will then store the field values as RDF triples with the specified field property, which is defined in the @RDF annotation.

```
@KiWiFacade
@RDFType({ Constants.NS_TAGIT+"PointOfInterest", Constants.NS_GEO+"Point" })
public interface PointOfInterestFacade extends ContentItemI {

        /**
         * The longitude of this point of interest. Maps to geo:long of the geo ontology
         * in the triple store.
         *
         * @return
         */
        @RDF(Constants.NS_GEO+"long")
        public double getLongitude();

        public void setLongitude(double longitude);

        /**
         * The latitude of this point of interest. Maps to geo:lat of the geo ontology
         * in the triple store.
         * @return
         */
        @RDF(Constants.NS_GEO+"lat")
        public double getLatitude();

        public void setLatitude(double latitude);

}
```

3 Open Data, to get more Privacy

One of the research topics of KiWi is to implement and use Linked Open Data (LOD) [3]. The ambition of LOD is to distribute and connect data in form of RDF triples on the net. One of the biggest projects which is implementing LOD is DBpedia[4], which extracts data from Wikipedia and provides them in form of RDF, and consequently generates one of the world biggest knowledge bases.

KiWi‘s objection is to contribute to the LOD cloud by providing (sorted) data, which has been created and annotated by its users and extended by its automatic reasoning service. So what’s the purpose of LOD? Futuristically speaking, the WWW will turn into the worlds biggest distributed database that can be queried for resources and used by agents [1], maybe also leading to an application-neutral web.

[3] http://linkeddata.org/

[4] http://dbpedia.org/

Currently, KiWi uses LOD in two specific ways: Thesaurus management and mapping of data from fully structured content to unstructured content. Thesaurus management is used to create and maintain its vocabularies. KiWi synchronizes its set of vocables via LOD with the data set of the Poolparty[5] thesaurus management. On the web, thesauri are mainly used to automatically understand and interpret unstructured content and to use it for information extraction (e.g. tag recommendation).

The meaning of mapping data from fully structured to unstructured content is to provide the user the possibility to discuss content that is loaded out of a relational database into a wiki in a way that it does not restrict the content that can be added. The even more interesting part is that the unstructured content is still RDF annotated and can thereby be updated into the relational database without any explicit maintenance.

Another use case of reusing existing data on the web is the one click authentication with FOAF+SSL[5], which enables client certificate authentication and importing of personal data into the platform and, hence, mainly reduces maintainability of different community applications. In this case, personal data, like first name, last name, birthdate, friends and whatever one is willing to put on the web is stored in a so-called FOAF[6] file, connecting to a unique web-id which represents the users identity. This seams scary on the first view, but the intention of FOAF+SSL is to allow users to decide which data can be published and which should, for example, just be provided in encrpyted format or not at all.

In general, to use FOAF+SSL one has to create a X.509 certificate from an RSA key-pair, which contains the web-id. The web-id is the primary resource of the FOAF file, which is extended with the users public key. Providing the public key for each web-id is useful for enabling digital signatures and public-key encryption on collaboratively created content and to make signed updates on the FOAF file. This might lead us one step further to the Web Of Trust[7].

4 Enterprise Use Cases and Open Source Community

The KiWi use cases are lead by our project partners (the old) SUN Microsystems (cz) and Logica (dk). The Logica use case is about knowledge sharing in the field of project management. The SUN use case, which is maybe more interesting for this target group, is to use KiWi as a platform to support the Netbeans development group in their daily work life.

The following example describes how KiWi supports the knowledge sharing process in the field of Software development. Contributions in form of article creation, modification, taggin, annotation or simply reading are calculated for each user in order to identify his level of expertise to a specific topic. Two persons, *Joe* and *Mary*, for example, are maybe both Software developers. *Joe* has just

[5] `http://semanticweb.org/wiki/PoolParty`

[6] `http://www.foaf-project.org/`

[7] `http://de.wikipedia.org/wiki/Web_of_Trust`

been employed and is working on the Netbeans user interface for the first time, while *Mary* has her expertise in the core of the Netbeans system. Mary is very active in the contribution to wiki articles about the Netbeans Core System. She writes and reads a lot of articles, leaves comments, tags and RDF annotations, which increases her so-called *Community-Equity-Value*.

Now *Joe* gets an error which hints to a tricky problem in the core system. Let's assume that Joe and Mary are physically separated. Then *Joe* could either google or ask his colleagues if they know someone who knows how to fix that, or he can use the KiWi system to find an answer to his problem. His first objection is probably to look up the set of content if a solution already exists. He searches for all ContentItems which are tagged with `Netbeans Core`, `Java` and `Bug`, and adds a part of his exception stack trace, to the query. If no correct search result was found, he might refine his search for `type:Person` and get *Mary* as the first result, since she is the one the system identifies as expert.

If the bug is not blocking his work, he might decide to just create a new Wiki article instead of directly contacting Mary. He describes his problem, tags it and waits for a response. Mary and other experts will receive the article in their recommendation box and start a discussion about how to solve *Joe*'s problem.

In a big company, like SUN (or Oracle) it makes sense to have such a platform which allows to collaborate without boundaries of geographical areas. In an Open Source community it would make even more sense to work with KiWi, because the developers mainly just interact over the internet. Knowledge sharing would become a lot easier, more effective and also more social interactive.

Finally, we would like to announce that KiWi also needs open source developers to maintain the project behind the project's end date and to improve KiWi's performance and code quality, since, as being a field research project, it has not been attached much importance to this topics. Further information about the project can be found under `http://www.kiwi-project.eu/` and `http://www.kiwi-community.eu/`. The source is available under `https://svn.salzburgresearch.at/svn/kiwi` and the showcase demonstrates the current KiWi version 0.8 `http://showcase.kiwi-project.eu`.

References

1. T. Berners-Lee and J. Hendler. The Semantic Web. *Scientific American*, 284(5):34–43, 2001.
2. F. Bry and K. Weiand. Flavors of KWQL, a Keyword Query Language for a Semantic Wiki. *SOFSEM 2010: Theory and Practice of Computer Science*, pages 247–258.
3. Graham Klyne, Jeremy J. Carroll, and Brian McBride. Resource Description Framework (RDF): Concepts and Abstract Syntax. `http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/` (22.05.2010), February 2004.
4. Sebastian Schaffert, Julia Eder, Szabolcs Grünwald, Thomas Kurz, Mihai Radulescu, Rolf Sint, and Stephanie Stroka. KiWi–a platform for semantic social software. In *Proceedings of the 4th Workshop on Semantic Wikis, European Semantic Web Conference*, 2009.

5. Henry Story, Bruno Harbulot, Ian Jacobi, and Mike Jones. FOAF+SSL: RESTful Authentication for the Social Web. In *Proceedings of the First Workshop on Trust and Privacy on the Social and Semantic Web (SPOT2009)*, 2009.