

Towards a Semantic Wiki-Based Japanese Biodictionary

Hendry Muljadi, Hideaki Takeda, Shoko Kawamoto, Satoshi Kobayashi,
and Asao Fujiyama

National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo, Japan
{hendry, takeda, skawamot, satoshi-k, afujiyam}@nii.ac.jp
<http://www.nii.ac.jp>

Abstract. This paper describes an on-going project to develop and maintain a web-based Japanese Biodictionary within a Semantic Wiki environment. For the development of the dictionary, MediaWiki is extended to enable the writing of labeled links that represent RDF triples. The extension enables Semantic Wiki to provide not only collaborative environment for experts in various biology fields to create and edit the dictionary, but also the navigation support to manage relations between terms. Using a simple wiki syntax, people can develop and maintain the dictionary visually and easily.

1 Introduction

Developing a web-based Japanese Biodictionary is obviously a hard and time-consuming task. The project for the development of a web-based Japanese Biodictionary has done a survey over 77 Japanese textbooks in various biology fields, which include textbooks for high-school, undergraduate and graduate students. There are more than 30,000 different terms extracted for the web-based dictionary. However, writing the description of each term is another thing to do. It will cost a lot of money to pay for the copyright if we use the description written on the textbooks or other dictionaries. On the other hand, as the web-based Japanese Biodictionary is also targeting high-school students as its viewers, it is necessary to provide an easy-to-understand description of the terms as well as the relations between terms [1].

Learning from the development of Wikipedia¹, we consider that collaboration on the web is perhaps the best solution for the development of the dictionary. By allowing experts in various fields of biology to collaborate, the dictionary may well be developed. However, as we implemented the MediaWiki² in its original form, we found that users cannot describe the relation between terms flexibly. Allowing the MediaWiki to write labeled links can solve this problem.

This paper presents the development of a Semantic Wiki-based Japanese Biodictionary. Section 2 presents the features of the proposed Semantic Wiki, which is

¹ <http://en.wikipedia.org/wiki/Wikipedia>

² <http://www.mediawiki.org>

called MewKISS. Section 3 presents the current state of the Semantic Wiki-based Japanese Biodictionary. The conclusion of this paper is presented in section 4.

2 Semantic Wiki

2.1 Extending MediaWiki

MediaWiki is a very useful tool for collaborative content management. MediaWiki also has a category management function which allows users to create class and sub-class relation between Wiki pages under the namespace (“Category”), and class and instance relation between Wiki pages under the namespace (“Category”) and common Wiki pages. However, MediaWiki in its original form does not allow users to create relation between pages flexibly. Our research has enabled MediaWiki to write labeled links [2]. Wiki syntax to write the labeled link is `[[term:target_page|property]]`. The labeled link relations will be displayed on the Wiki pages as follows.

1. On the source_page: `-> property -> target_page`
2. On the target_page: `<- property <- source_page`
3. On the property: `source_page -> target_page`

Fig.1 shows an example of the writing of labeled link on a Wiki page. Fig.2 shows how the labeled link relations will be displayed on the Wiki pages.



Fig. 1. Writing the labeled links on a Wiki page



Fig. 2. Displaying the labeled link relations on Wiki pages

2.2 Features of the Proposed MewKISS

MewKISS is an abbreviation for MediaWiki with Simple Semantics. The word “KISS” is written with full capital letters to stress that the proposed Semantic Wiki is

developed by considering the idea of the KISS principle³. MewKISS emphasizes the user-friendliness of the Semantic Wiki engine. It is developed to allow non-technical users to write and edit metadata according to RDF statements easily, and leaves the more technical aspects to external applications.

Fig. 3 shows the overall structure of the MewKISS⁴. The features of the MewKISS can be summarized as follows.

1. A collaborative lightweight metadata management. Enabling MediaWiki to write labeled links with simple syntax allows users to create and manage relations between Wiki pages easily and flexibly. The writing of labeled links allows users to write and edit RDF triples even though the users have no knowledge about it.
2. Navigation support. Displaying labeled links allows users to navigate the relation between Wiki pages visually.
3. Mapping to other Semantic Web application. MewKISS handles only simple RDF statements. By converting the RDF triples, which are stored in the new table of MewKISS, into XML-encoded RDF data format, the RDF triples can be exported to RDF database such as Sesame⁵.
4. An integrated content and metadata management. By extending MediaWiki, the Semantic Wiki has the benefit of having all the functions available in MediaWiki as a collaborative content management system. Thus, the Semantic Wiki can be used as a collaborative and integrated content and metadata management system.

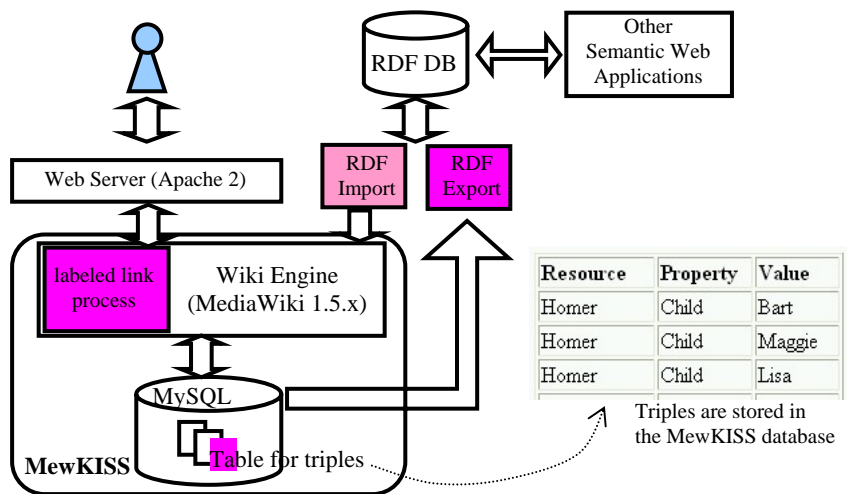


Fig. 3. Overall structure of the proposed Semantic Wiki

³ http://en.wikipedia.org/wiki/KISS_principle

⁴ Our current prototype is accessible from <http://semanticwiki.jp/>

⁵ <http://openrdf.org/>

3 Semantic Wiki-Based Japanese Biodictionary

3.1 Current State of the Semantic Wiki-Based Japanese Biodictionary

Currently the prototype system contains more than 4,000 terms. There are 5 RDF properties used in the biodictionary: is-a, part-of, synonym, English and English synonym. The first two properties, is-a and part-of, are used to represent class and sub-class relations. Synonym is used to link a Japanese biology term with its synonym Japanese term. English is used to link a term with its direct translation's English term, while English synonym is used to link a term with its English term's synonym. Using our proposed Semantic Wiki, other properties, such as kind-of, may also be easily used.

Category management function is also used for categorization and also to list all the terms in the dictionary.

Fig. 4(a) shows a category page that is used to list all the terms in the biodictionary. Fig. 4(b) shows the editing page of a term. Fig.4(c) shows the Wiki page of a term.



Fig. 4. Wiki pages in the Japanese Biodictionary

3.2 Future Works

The current Semantic Wiki-based Japanese Biodictionary allows users to collaboratively create and edit biology terms as well as their relations to other terms. By converting the stored RDF triples into XML-encoded RDF data format, RDF triples can be exported to Sesame. However, further work needs to be done to enable users to export the RDF triples to Sesame directly from the MewKISS environment. Further works also need to be done to allow simple and complex queries directly from the MewKISS environment, and on how the semantic mapping to other Semantic Web applications may benefit the biodictionary.

4 Conclusion

For the development of a Semantic Wiki-based Japanese Biodictionary, MediaWiki is extended to enable the writing of labeled links that represent RDF triples. MewKISS provides a collaborative, easy-to-use and integrated content and metadata management system. In the MewKISS environment, users can write and edit RDF triples, even though they have no knowledge about it. RDF triples stored in the table of MewKISS can be converted into XML-encoded RDF data format and exported to RDF database such as Sesame. Thus, MewKISS may well serve as a bridge between non-technical users and Semantic Web technology.

Developing the Japanese Biodictionary within a Semantic Wiki environment does allow experts in various biology fields to create and manage content of the dictionary as well as the relation between terms easily and visually.

References

1. Kobayashi, S., Kawamoto, S., Mizuta, Y., Demiya, M.S., Muljadi, H., Suzuki, S., Abe, T., Araki, J., Shirai, Y., Ito, T., Kondo, T., Kitamoto, A., Miyazaki, S., Gojobori, T., Sugawara, H., Takeda, H., Fujiyama, A.: The New Generation Bioportal: the Development of a Web Site for Biology Education. In Proc. of the 80th Domestic Conf. of the Society of Biological Sciences Education of Japan (2006) 27 (in Japanese)
2. Muljadi, H., Takeda, H.: Semantic Wiki as an Integrated Content and Metadata Management System. In Poster & Demonstration Proc. of the 4th Intl. Semantic Web Conf. (2005) PID 44