

DokuSIOC – SIOC Plugin for DokuWiki

Michael Haschke

(Freelancing Webworker, Germany
michael.haschke@eye48.com)

Sebastian Dietzold

(Universität Leipzig, Institut für Informatik, Germany
dietzold@informatik.uni-leipzig.de)

Abstract DokuWiki is a popular and widely used wiki engine, keeping all the data directly in the file system. DokuSIOC integrates the SIOC ontology within DokuWiki and it provides meta data about the wiki documents, using RDF/XML views.

Key Words: Semantic Web DokuWiki Plugin SIOC

1 Introduction

The Semantic Web extends the original Web with machine processable semantics [Berners-Lee et al., 2001] and it provides technologies and standards to integrate, relate and combine machine readable data. In contrast to the original Web, the Semantic Web describes not only interlinked documents but rather interoperable data which can „be shared and reused across application, enterprise, and community boundaries.“¹ Semantic annotations as well as data interoperability are good reasons to enhance existing applications in order to connect them with the Semantic Web and get their data out of their silos.

DokuWiki is a wiki [Leuf and Cunningham, 2001] engine, „mainly aimed at creating documentation of any kind. It is targeted at developer teams, workgroups and small companies. All data is stored in plain text files – no database is required.“² In DokuWiki, wiki pages can have sub pages, in this case a wiki page is a namespace as well.

DokuWiki is widely used³, known in literature [Badger, 2006] and even nominated for the Top 10 wiki engines at Cunningham’s WikiWikiWeb⁴. In addition

¹ <http://www.w3.org/2001/sw/>

² <http://dokuwiki.org/>

³ Indicated by some popularity statistics, taken from different websites. On WikiMatrix (<http://www.wikimatrix.org/statistic/>), DokuWiki is listed first with 7306 detail views (MediaWiki: 3518) and is on the second place in comparison views: 37670 (44674). At Ohloh (<http://www.ohloh.net/tags/wiki>) it is listed straight after MediaWiki with 123 users (489) for projects tagged with „wiki“ and is ranked first with a popularity score of appr. 2499 (801) at Freshmeat’s site for wiki projects (<http://freshmeat.net/tags/wiki>). (2009-08-05)

⁴ <http://c2.com/cgi/wiki?TopTenWikiEngines> (2009-08-05)

to this, DokuWiki is the technical foundation for the research project ICKE 2.0 (Integrated Collaboration & Knowledge Environment, Fuchs-Kittowski and Hüttemann [2009]), so we assume a broad user base with many installations and target content for our triplification project.

The SIOC (Semantically-Interlinked Online Communities) Core Ontology specifies the main concepts and properties for describing information from online communities, e.g. wikis and weblogs, on the Semantic Web [Berrueta et al., 2007]. As other vocabularies, SIOC is extensible and can be combined with other models as well.

In Section 2, I discuss how DokuWiki and its users benefit from the possibility to export data using the SIOC model. The implementation with DokuWiki is described in Section 3 and then in Section 4, I will do a conclusion and an outlook what steps could be the next.

2 Benefits

Leaving the closed data silo and providing interoperable data are the main benefits of using Semantic Web technologies, offering more possibilities to aggregate content – beside Atom and RSS – is a nice side effect. To connect user spaces and content in DokuWiki instances from currently unconnected user profiles is another basic benefit.

2.1 Overcoming closed data silos

As always we are discussing the technological openness, not the undermining of the rights management. Beside the DokuWiki view of the wiki pages via browser you can currently access the page content via file system or raw text export. These possibilities are lacking semantically defined meta data, e.g. creator, contributors and relations between wiki pages, to name a few. Other applications cannot process this wiki data without a specific importer. The SIOC plugin opens the wiki instance to third party applications via Linked Data [Berners-Lee, 2006] and the usage of the well defined SIOC vocabulary. A consuming agent is now able to process the data without any knowledge of DokuWikis internal structure.

2.2 Content aggregation

Currently, DokuWiki provides RSS and Atom feeds. The feed content depends on the configuration created by the administrator of the DokuWiki instance. It could be the current version of recently added pages or maybe only the differences between revisions. With the SIOC plugin it is possible to start at any point and aggregate the content part by part, including page relations and revision

links. As an example figure 1 shows aggregated content in the semantic data wiki OntoWiki[Auer et al., 2006], which was fetched by OntoWiki's Linked Data wrapper.

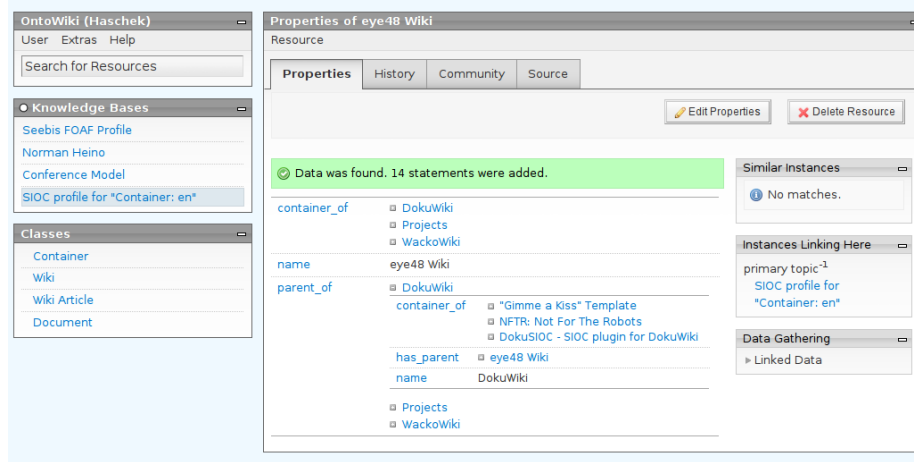


Figure 1: OntoWiki view on <http://eye48.com/dokuwiki/> after import of some resources via Linked Data wrapper

2.3 Declare your content

If a person contributes content to one or more than one DokuWiki instances, it is possible to link from the person's homepage to the various wiki pages. Those HTML links are untyped relations between web sites and need to be updated with the creation of every new wiki page. The SIOC plugin simplifies this process by providing SIOC profiles for user accounts. Now a person only needs to add the resource URIs of those accounts to the personal FOAF (Friend of a Friend, [Brickley and Miller, 2007]) profile to declare her own content. In particular, wiki pages could be retrieved automatically.

3 Implementation with DokuWiki

The DokuWiki architecture is extensible by action plugins, which can subscribe their methods as event listeners to DokuWiki events. The DokuWiki documen-

tation⁵ gives detailed information how to do it. In this section, we discuss the implementation of DokuSIOC⁶.

3.1 Creating a user space

DokuWiki handles internal user identifiers and profiles, e.g. to provide access control and rights management. It does not provide any URI for a user. In order to do create such a user space, DokuSIOC offers a way to configure a DokuWiki namespace, where user identifiers can be used as sub pages. This hook provides profile sites for user accounts, e.g. <http://eye48.com/dokuwiki/doku.php?id=user:haschek>.

3.2 Creating missing resources

A DokuWiki URI can stand for different resources. The profile example may describe the user (`sioc:User`), a wiki page (`sioc:WikiArticle`), a container for sub content (`sioc:Container`) and even the special wiki container (`sioc:Wiki`) if the user page is used as DokuWiki starting page. The SIOC plugin adds a type parameter to distinguish exactly between the resources, for example the URI for the user account is now <http://eye48.com/dokuwiki/doku.php?id=user:haschek&type=user>, an example for a wiki article is <http://eye48.com/dokuwiki/doku.php?id=en&type=post>.

3.3 SIOC documents and Linked Data

DokuWiki saves all content and its meta data to flat files. This kind of data store prevents the usage of already existing projects like Triplify [Auer et al., 2009] to publish content as RDF. DokuSIOC uses the DokuWiki export action to provide a meta document which describe the requested resource. The Linked Data approach is used to interlink between related resources, e.g. containers of wiki pages, backlinks and contributing users. Linked Data wrappers now can request further documents to access more data. To generate the XML serialisation of the SIOC documents, DokuSIOC uses the PHP exporter provided by the SIOC project⁷. The meta document for the article example is http://eye48.com/dokuwiki/doku.php?do=export_siocxml&id=en&type=post.

3.4 Linking to RDF export

DokuSIOC provides two different ways to direct client applications to the meta data. Firstly, DokuSIOC adds a meta link to the header of the DokuWiki HTML

⁵ http://www.dokuwiki.org/devel:action_plugins

⁶ <http://eye48.com/go/dokusioc>

⁷ <http://wiki.sioc-project.org/index.php/PHPEXportAPI?version=1.01>

view, `<link type="application/rdf+xml" rel="meta" href="http://[...]/doku.php?do=export_siocxml&id=en&type=post" />`. Clients can follow this link for fetching the RDF/XML document with meta data about the currently shown HTML document. In addition to this, DokuSIOC supports content negotiation as discussed in [Ayers and Völkel, 2008]. If the client requests `http://eye48.com/dokuwiki/doku.php?id=en&type=post` as `application/rdf+xml`, the plugin will send a 303 (see other) HTTP status code and forwards to the location of the RDF export view.

4 Conclusion and outlook

In Section 1 we introduced DokuWiki and why it is worth to have an RDF export there. The main benefits for DokuWiki and its users has been described in Section 2, mainly it is about accessible data. Afterwards we presented an implementation as DokuWiki plugin and showed, how it works. Additionally the plugin offers the feature to ping `http://pingthesemanticweb.com/` when content is created or modified.

In a future version, we want to enhance important DokuWiki plugins like discussion and tagging in order to support RDF export too as well as to add more meta data to the SIOC profiles, e.g. by automatically linking DokuWiki articles with DBpedia [Auer et al., 2008] resources.

Finally we hope to convince the DokuWiki community and the project maintainers for packaging DokuSIOC in the default DokuWiki distribution.

References

- S. Auer, S. Dietzold, and T. Riechert. Ontowiki - a tool for social, semantic collaboration. In I. F. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. Aroyo, editors, *Proceedings of 5th International Semantic Web Conference 2006*, volume 4273 of *Lecture Notes in Computer Science*, pages 736–749. Springer, 2006. ISBN 3-540-49029-9. URL <http://dblp.uni-trier.de/db/conf/semweb/iswc2006.html#AuerDR06>.
- S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. Dbpedia: A nucleus for a web of open data. In *Proceedings of 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference (ISWC+ASWC 2007)*, pages 722–735, November 2008. doi: 10.1007/978-3-540-76298-0_52.
- S. Auer, S. Dietzold, J. Lehmann, S. Hellmann, and D. Aumueller. Triplify – lightweight linked data publication from relational databases. In *Proceedings of the 17th International Conference on World Wide Web, WWW 2009, Madrid, Spain, April 20-24, 2009*, pages 621–630, 2009. URL <http://www.sheridanprinting.com/09-www-cd35mxcg/docs/p621.pdf>.

- D. Ayers and M. Völkel. Cool uris for the semantic web. World Wide Web Consortium, Note NOTE-cooluris-20081203, December 2008. URL <http://www.w3.org/TR/2008/NOTE-cooluris-20081203/>.
- M. Badger. Dokuwiki - a practical open source knowledge base solution. *Enterprise Open Source Magazine*, 4(10), 2006. URL <http://opensource.sys-con.com/node/318853>.
- T. Berners-Lee. Linked data. World wide web design issues, July 2006. URL <http://www.w3.org/DesignIssues/LinkedData.html>.
- T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5):34–44, May 2001.
- D. Berrueta, D. Brickley, S. Decker, S. Fernández, C. Görn, A. Harth, T. Heath, K. Idehen, K. Kjernsmo, A. Miles, A. Passant, A. Polleres, L. Polo, and M. Sintek. Sioc core ontology specification. W3c member submission, W3C, June 2007. URL <http://www.w3.org/Submission/2007/SUBM-sioc-spec-20070612/>.
- D. Brickley and L. Miller. The Friend Of A Friend (FOAF) vocabulary specification, November 2007. URL <http://xmlns.com/foaf/spec/>.
- F. Fuchs-Kittowski and D. Hüttemann. Towards an integrated collaboration and knowledge environment for sme based on web 2.0 technologies - quality assurance in enterprise wikis. In K. Hinkelmann and H. Wache, editors, *Wissensmanagement*, volume 145 of *LNI*, pages 532–543. GI, 2009. ISBN 978-3-88579-239-0.
- B. Leuf and W. Cunningham. *The Wiki Way. Quick Collaboration on the Web*. Addison-Wesley, Boston, 2001.