

Metadata for WICRI, a Network of Semantic Wikis for Communities in Research and Innovation

Jacques Ducloy
DRRT Lorraine, France
Jacques.Ducloy@loria.fr

Thierry Daunois
Univ. Lorraine, France
Thierry.Daunois@inpl-nancy.fr

Muriel Foulonneau
CRP Tudor, Luxembourg
Muriel.foulonneau@tudor.lu

Alice Hermann
INSA Rennes, France
Alice.Hermann@irisa.fr

Jean-Charles Lamirel
Univ. Lorraine, France
Jean-charles.lamirel@loria.fr

Stéphane Sire
EPFL, Switzerland
Stephane.sire@epfl.ch

Jean-Pierre Thomesse
Univ. Lorraine, France
Jean-pierre.thomesse@loria.fr

Christine Vanoirbeek
EPFL, Switzerland
Christine.vanoirbeek@epfl.ch

Abstract

This paper introduces metadata issues in the framework of the WICRI project, a network of semantic wikis for communities in research and innovation, in which a wiki can be related to an institution, a research field or a regional entity. Metadata and semantic items play the strategic role to handle the quality and the consistency of the network, that must deal with the “wiki way of working” in which a metadata specialist and a scientist can work altogether, at the same time, on the same pages. Some first experiments of designing metadata are presented. A wiki, encyclopedia of metadata, is proposed, and related technical issues are discussed.

Keywords: network of wikis; Semantic MediaWiki; metadata encyclopedia, e-Science; CRIS.

1. Introduction

Since 1995, when Ward Cunningham launched WikiWikiWeb, wikis are playing an increasing role in scientific information systems. Yet, since most experiments in this field are tiny, the need for metadata is often limited. Wikipedia, with its 3.000.000 articles, already showed that a large wiki requires the use of specifically developed ways of dealing with metadata, as is illustrated by January 2010's statistics¹, showing 259.000 templates and 552.000 categories.

Thus, the question becomes even more accurate when somebody aims at dealing with a network of wikis. For now 2 years, we are facing the reality of what is the effective effort required for building up a large network of semantic wikis, and we are also experiencing the need for metadata. This paper aims at reporting about our experiment with WICRI (Wikis for Communities in Research and Innovation) network, which contains about sixty wikis, but we try to design as would be required for a network hosting thousands of wikis. Furthermore, semantic wikis introduce a new generation of metadata, allowing a knowledge modeling in a RDF framework that is interesting to consider.

In this paper², we will first introduce the WICRI network; then we will present the initial technical choices we started with. We will discuss the next issues in two ways: a contributor facing the writing of metadata, and new services for helping him in this task.

¹ < <http://stats.wikimedia.org/EN/TablesWikipediaEN.htm#namespaces> >

² This article is written as we have done for DC 2006 (Ducloy et al. 2006) with a “public collaborative version”: < http://maquettewicri.loria.fr/en.artist/index.php5?title=DC_2010_WICRI_paper >.

2. WICRI, a Network of Wikis for Research and Innovation

Wikipedia proves that wikis can help in building and disseminating a common knowledge at a very large scale. Yet it is not a sufficient answer to research needs, but given that still provides some guidance in building up WICRI. Wikipedia's validity is questioned by many academic institutions, making it necessary for WICRI to deal with the question of transparency of contributions and validity assessment. Wikipedia's practices (and especially the fact that Wikipedia's contributors must display information validated by external references) are not compliant with publishing policies of research activities: this is why some scientific committees will be required in order to control or moderate WICRI.

The WICRI network accepts two main classes of wikis. The role of the WICRI community is to design common **wikis**. On its own side, an entity can open an **institutional wiki**, and manage it its own way, with specific rules, differing from the rules of the common wikis. From a practical point of view, a common wiki has an identifier with *Wicri/* as first part, i.e. *Wicri/Lorraine*. *Wicri/Water(fr)* refers to the French component of a multilingual family of wikis, and *Wicri/Water(en)* to the English one. At the present time WICRI network contains almost 30 common wikis (on a regional framework such as *Wicri/Lorraine* or thematic, like *Wicri/Water*). *Wicri/Ticri* is a thematic wiki related to "Information & Communication Technology" (a Dublin Core portal is included). Several wikis are providing a global consistency. For instance, *Wicri/Wicri*, on which each topic and links to detailed pages in other wikis must appear, gives a global view of the network.

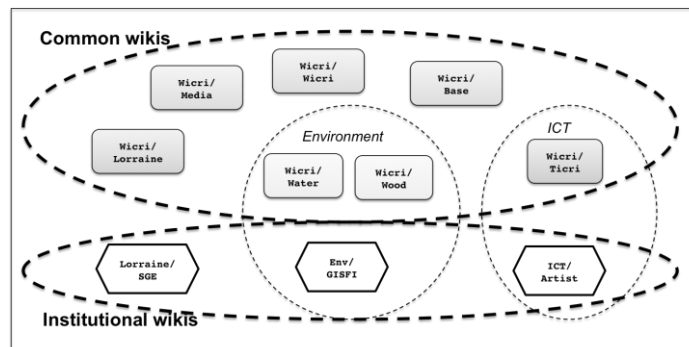


FIG. 1. The current WICRI network (a subset)

Wikis host editorial texts (scientific articles or surveys). They handle also a lot of items, such as program committees, laboratories, or funding programs, which belong to the CRIS model³. Jeffery's (2007) expected goal is to merge CRIS with open archives for generating an e-Science infrastructure. WICRI wants to go one step further in order to obtain a highly understandable CRIS while using editorial facilities of wikis for bringing a human readable summary. In this perspective, semantic wikis could provide a technical basis for implementing a CRIS as skeleton. Thus, MediaWiki was chosen as the engine of the WICRI network. Fully compatible with Wikipedia, it opens the way to the use of Semantic MediaWiki (Krötzsch, 2007).

³ A Current Research Information System, commonly known as "CRIS", is any information tool dedicated to provide access to and disseminate research information, such as People, Projects, Organizations, Results (publications, patents and products), Facilities, and Equipment (EuroCRIS, 2009). The European Commission supports the CRIS approach, through the CERIF (Common European Research Information Format) recommendation (<http://www.euroCRIS.org>). This way of working is spreading worldwide and, for instance, at the USDA (United States Department of Agriculture (<http://cwf.uvm.edu/cris/>)).

3. Writing a Networked Hypertext with Formulas and Metadata

Most content management systems, designed “before blogs and wikis”, clearly set apart editing contents and programming or managing metadata. Scientists are accustomed to writing mainly short and isolated papers, while Digital Libraries are reduced to storing isolated papers in archives or various databases. The question of setting up a global consistency of a knowledge domain is just not addressed. On a wiki, any actor can handle the whole process, from programming to writing contents, at any time, on any page. Authors may write pages and associated metadata in the same temporality: they can set up a “human brain designed” hypertext. Writing a scientific, readable and networked hypertext means handling consistently scientific objects and knowledge, and submitting given information in different ways and contexts, for different audiences.

3.1 Semantic Wikis for Scientific Objects

Scientists and engineers commonly use Technical objects (formulas, 3D images, etc.). Yet, the initial version of MediaWiki is quite poor for handling such objects and requires some specific extensions (for instance LaTeX⁴). The Proteopedia project (Hodis, 2008) is going one step forward, tackling the management of molecular items such as protein, RNA or DNA⁵. A contributor can use “green links” that interact with a Java applet (jmol). Generalizing this way of doing requires more complete XML support, with contributors having a good understanding of markup language. In such a context, handling syntax of metadata or semantic items would be quite easy. The difficulties would come from designing a global knowledge in a collective way⁶. About Semantic MediaWiki in science, some applications deal with CRIS approaches; for instance, semanticweb.org provides a semantic model for scientific events. Another set aims at building or curating ontology. But, until now, we have not found wikis that use ontologies in order to handle scientific objects with an editorial purpose. Moreover, Semantic MediaWiki does not represent a universal solution. For instance, SWiM (Lange, 2008), a semantic wiki for mathematical knowledge management⁷, has an interesting method of handling mathematical formulas.

3.2 Different Writing in Different Contexts for Different Audiences

In the WICRI context, most information should be developed on different wikis. For instance, each research project with several partners must be cited and commented in the regional wiki of each partner, as well as in all relevant thematic wikis. Table 1 shows different ways of managing the relationships between an ICT conference held in Lorraine, and its committee members⁸.

⁴ It requires installing LaTeX environment close to the operating system, which is a quite complex task.

⁵ < <http://proteopedia.org/wiki/index.php> >

⁶ For instance, a comparison between several wikis of the MediaWiki Foundation shows a multipurpose utilization of 3 classification schemes about life species. Here is an example about Acer on:

- Wikipedia Species: < <http://species.wikimedia.org/wiki/Acer>> ,
- Wikipedia (en): < <http://en.wikipedia.org/w/index.php?title=Maple&oldid=345810808> >
- Wikipédia (fr): < <http://fr.wikipedia.org/wiki/%C3%89rable> >
- Wikimedia Commons: < <http://commons.wikimedia.org/wiki/Category:Acer> >

⁷ < <http://wiki.openmath.org/> >

⁸ The call for papers is duplicated on Wicri/Ticri and Wicri/Lorraine. The event model of semanticweb.org is used with properties “Has PC member” and “Has OC member”. Paul Dupont, working in Lorraine, is always qualified with the property “PC member”. On Wicri/Lorraine John Smith is only linked to Wicri/Ticri with an interwiki link, because he has no author page on Wicri/Lorraine.

TABLE 1: A part of a page relative to a conference happening in Nancy.

The Committee as it appears on every pages.	<p>Program Committee</p> <ul style="list-style-type: none"> • <u>Paul Dupont</u>, Nancy (Fr) • <u>John Smith</u>, London (UK) <p>Organizing Committee</p> <ul style="list-style-type: none"> • <u>Jean Durand</u>, Nancy (Fr)
As it would be coded in a thematic (i.e. Ticri) wiki. <i>PC members are qualified by properties. OC members have only interwiki links</i>	<pre> ==Program Committee== * [[Has PC member::Paul Dupont]], Nancy (Fr) * [[Has PC member::John Smith]], London (UK) ==Organizing Committee== * [[wicri-lor.fr:Jean Durand Jean Durand]], Nancy (Fr) </pre>
As it would be coded in a regional (Lorraine) wiki. <i>Only, local PC or OC members are qualified by properties.</i>	<pre> ==Program Committee== * [[Has PC member::Paul Dupont]], Nancy (Fr) * [[ticri.en:John Smith John Smith]], London (UK) ==Organizing Committee== * [[Has OC member::Jean Durand]], Nancy (Fr) </pre>

We give below an example directly related to DCMI. Pittsburgh appears at least on 3 wikis. On *Wicri/Ticri*⁹, this city is linked to DC 2010 and the corresponding page speaks about main activities related to information science. On *Wicri/Water*, the content deals with confluence of Allegheny and Monongahela rivers to form the Ohio River. On *Wicri/Wicri*, the page gives general facts and links to the other pages. Even if these pages are related to the same topic, they display clearly distinct contents. As a last example, the article “*Qu’est-ce qu’une bibliothèque numérique, au juste?*” (Lagoze, 2005) is integrated in the Artist wiki in which it was first translated¹⁰. As a reference paper, a copy has been done in *Wicri/Ticri* where anchors and links are quite different than those existing on *Artist*. Since this paper's introduction could get a very large audience, this part is exclusively displayed on *Wicri/Wicri*.

3.3 Managing Network Consistency

For the WICRI project, a critical issue is to manage network consistency. Here follows an example that implies a large set of pages about geographic items such as countries, towns, etc.

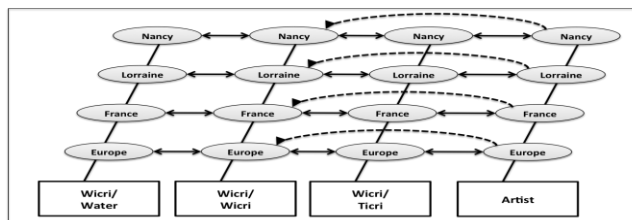


FIG. 2. Interlinks between geographic items

When a new city appears on a given wiki, the contributor should theoretically keep the connectivity of the networked hypertext. Fig. 2 gives an example with Nancy in an institutional wiki (*Artist*). The Nancy related page on *Artist* must be linked with Lorraine, France and Europe pages on the same wiki (these pages must eventually be created). It must also be linked to Nancy page on *Wicri/Ticri*, *Wicri/Wicri*, and so on. In a multilingual context, this graph must be duplicated while taking care of translation¹¹. To be understandable by readers, this consistency needs to be explained by text. Automatic tools could provide an initial building, but contributors must also be employed to write explanations. Thus, managing network consistency with metadata is a cooperative task involving both human contributors and computers.

⁹ < <http://maquettewicri.loria.fr/fr.ticri/index.php5?title=Pittsburgh> >

¹⁰ < http://maquettewicri.loria.fr/fr.artist/index.php5?title=Ametist_0_Lagoze >

¹¹ For instance, the page “Lorraine” in French is related to "Lorraine (region)" in English.

4. Wicri/Metadata, a wiki for Authors and Contributors

Most wiki pages are written by human contributors, and not by computers. Computers could help but, *in fine*, contributors will make pages. In an OAI-PMH network, computer protocols share controlled metadata and provide consistency. In a wiki network, a contributor can write on many wikis and must interact with metadata that plays a crucial role. Here is an example: how to write: “DCMI announces that DC-2010 will be held in Pittsburgh” in a semantic wiki? While reading the manual, it seems to be easy; you just have to enter something like this:

```
[[organizer::DCMI]] announces that DC-2010 will be help in [[place::Pittsburgh]]
```

When pushing the “Save page” button, the relations and, if needed, the properties are created. Thus the true problem does not deal with syntax, but with semantics: how to choose and to name a property? For instance, about the role of the DCMI in DC conference, we could write: `organizer, has organizer, has global organizer, DC:contributor, etc.`¹². In WICRI, the problem that we have pointed out for `semanticweb.org` is distributed on a network. Thus the following aspects have also to be addressed. How do we know if a property exists in the semantic model of the wiki? How do we choose a new name for a new property to be consistent with the existing ones? In a multilingual family of wikis, how could metadata items be translated? We thus propose to set up a wiki, with an encyclopedic philosophy dealing with metadata. There are several wikis dedicated to metadata on the web, for instance, on the DCMI (Enoksson, 2008), but they are usually dedicated to specialists and, often related to a particular schema. Here, we want to be understood by a non-specialist¹³ who has to deal with many topics at the same time.

Metadata are related to a model (possibly expressed through an ontology) to represent the structure of the wiki and the properties of wiki resources. Each wiki can be created with several specific domain models (for instance the FAO World Reference Base (WRB) is used for soil resources, in `Wicri/UrbanSoils`) and several general models (for instance the research event model from `semanticweb.org`). Moreover, some concepts may exist in different languages. As a result, different wikis may use close or similar concepts using different models. A specific wiki, `Wicri/Base`, has been created to provide tools, including templates and metadata sets (e.g. Semantic Infobox) and metadata elements. But this wiki only deals with items that have reached a strong consensus; the `Wicri/metadata` wiki must help in building this consensus.

4.1 Representing General Research Resources

The major function of `Wicri/metadata` is to provide elements to define metadata related to general resources of scientific communication. It relates to CRIS as well as Research repositories. The representation of resources is bound to the general domain of research, including concepts which belong to CRIS, Knowledge Organization Systems used in the different research domains or created ad hoc (for instance, see Tifous, 2007), bibliographic formats such as MARC or the DCMI Scholarly Work Application Profile, datasets formatting models such as text formatting (TEI, etc.), survey datasets (DDI), educational formats such as LOM, persons (e.g. FOAF), etc. With this set of schemas, the same concept could appear several times, with several shades. `Wicri/Metadata` has to explain this kind of situation in order to design guidelines, or to support multilinguality (e. g. `Attribut:A pour ville` adapted from `Property:Has location city`).

¹² A look at `semanticweb.org` illustrates this difficulty. The “Property namespace” contains 773 pages with 768 real properties; 277 pages are classified as “wanted properties” (without explicitness). Looking for `DC:creator`, we have found several variants. The preferred term is “Has author” (frequency 99). The most used term is “Author” (1058). The expression “Written by” appears 35 times. At least, “Author of”, “Content author”, and “Creator” appear once. (Data collected on 4th, March 2010).

¹³ For instance, we avoid to link to pages containing a thousand lines of RDF/XML, as an explanation!

4.2 Ensuring Interoperability with other Semantic Applications

An interesting strategy is to find a “kernel ontology” that can be used without major adaptations. In this case, only the extensions have to be explained in *Wicri/Metadata*. This ensures interoperability with other semantic applications. WICRI operates like this for the model of conferences, starting from *OpenResearch.org* and explaining local adaptations. This approach is generalized for describing scientific contents. WICRI is thus exploiting *Eurovoc*¹⁴ as a general ontology, which should be completed by specialized ones, for example *WRB*. Some repositories, such as *OntologyPattern*¹⁵ or *Watson*¹⁶ can be used for discovering domain ontologies. However, metadata editors will still have to search specifically for existing properties and sometimes they may find close but not exactly similar properties. This raises the issue of defining the relationships between concepts defined in different models.

4.3 The Wiki as a Metadata Registry?

Until now, WICRI has chosen to define redirects (i.e. *owl:sameAs* relations) with concepts from ontology repositories. However, in this case, the strict equivalence of two concepts is limited. Ontology mapping requires richer relations to be encoded, such as *SKOS* mapping properties *skos:exactMatch*, *skos:closeMatch*. Moreover, collaborative ontology mapping mechanisms (Correndo, 2008) should be available to the network so that any contributor who creates a new metadata concept or identifies a new relation should be able to enrich the system.

This should result in a wiki-based metadata registry for the WICRI network, with some specificity. The wiki architecture allows the expression of a mix between structured and unstructured content. Scientific concepts might not be presented only with traditional definitions, but also using scientific literature, guidelines etc. This is particularly important in a multilingual context, as we observed in the WICRI network as well as in other collaborative scientific platforms. A review of concepts used to describe resources in the field of education (Sarre, 2010) demonstrates that many concepts proposed as metadata for this domain are not fully specified. There are metadata schemas, as well as concepts, defined only in journal articles. It should therefore be possible to add concepts, even outside the scope of a proper ontology. In addition, semantic wikis include some intelligence, which can be useful to make inferences on the existing, or potential, relations between the concepts used in the network. The wiki network is thus not only an interface to a CRIS and research repositories, it also makes research content and scientific communication a building block of the semantic Web by providing dereferencable resources and reasoning mechanisms through a decentralized and collaborative environment.

5. Metadata for Computers

5.1 Networks and Distributed Wiki Applications

The “wiki way” puts the contributor in the heart of the metadata handling. So, what could be the role of the computer? One strong issue is replication management (Rahhal, 2009). We have identified 5 classes of replication cases.

1. **Wiki replication.** A whole wiki could be duplicated in a P2P network of wikis with a distributed replication mechanism. This feature is useful for technical reasons (strategic wiki as *Wicri/Wicri*) or sometimes for political ones (a wiki providing visibility for several institutions). But, it does not matter with editorial replications, or metadata.

¹⁴ < <http://europa.eu/eurovoc/> >

¹⁵ < <http://ontologydesignpatterns.org> >

¹⁶ < <http://kmi-web05.open.ac.uk/WatsonWUI/> >

2. **Page replication.** A page (or a set of pages) is replicated on several wikis. This kind of facility begins to be available for invariant pages, such as templates related to semantic models. For instance, in DSMW¹⁷, this mechanism is driven by metadata (semantic properties).

3. **Paragraph replication.** Until now, we have not found an extension of SMW extending the previous mechanism at the paragraph level. This need is ubiquitous in WICRI network. A palliative method, creating templates for each paragraph, might work, but a human contributor could not really use it (for instance, this latter will need one page for each bibliographic reference).

4. **Paragraph replication, with transformations.** In many cases, the previous mechanisms could not be applied because the paragraph must be transformed while replicating. For instance, for editorial reasons, requirements for handling organization committees can be different in a regional wiki (with semantic links for local members) and in a thematic wiki (no links).

5. **Replication of sets of several pages.** Such an example was given before (geographic items).

Due to this large amount of problems, we have left out fully automated systems, and think instead about "computer assisted hypertext writing".

5.2 Handling WICRI Network Consistency

For handling network consistency, a first approach consists of extending facilities that are soon provided on a simple wiki. Consequently, we are testing bots that use an XML schema.

```
<wicri>
  <wiki prefix="wicri.fr" type="public" title="Wicri (en)"
    server="http://maquettewicri.loria.fr" path="/en.wicri/index.php5?">
  <recentChanges title="Special:Recent_changes"/>... </wiki>... </wicri>
```

FIG 3. XML description of WICRI network, used for piloting bots for providing a consolidation of facilities, such as "RecentChanges" at a network level.

In future, we intend to use specialized tools in interaction with the wiki network. For instance, defining a geographical ontology is not really secure as soon as the items are handled simultaneously by administrators, or bots, but also by non-specialist contributors. Thus a better way uses external tools, like Protégé, to handle consistency. Several methods designing ontology in a cooperative way (Tudorache, 2008) are promising.

For Human-computer interaction, the semantic forms of SMW might be useful in some cases (for instance for entering a bibliographic record), but appear insufficient with editorial constraints. Using XML editors, for instance, XTiger (Sire, 2010) seems promising, but requires a better handling of XML objects by MediaWiki. Hence, implementing some requirements such as "structuring a wiki page in TEI" or "templates with list as parameters" is an issue, which must be planned for long term. However, in a more short term, we could expect to better help a contributor in discovering resources when he writes a new page.

5.3 Enriching the Wiki Network through the use of Web Data

Web mining represents an important challenge for enhancing the dynamicity, the flexibility and the scope of a wiki network. Hence, on the one hand, this process is mandatory for assisting the upcoming contributors with elaborated and reliable redaction guidelines during the network construction phase. On the other hand, it is also decisive for supplying end-users with external information whose added value is to maintain significant relationships with the semantic context of the wiki network.

On the end-user's side, the goal of querying the web is both to complete as well as to enrich the information on a given topic as soon as this latter has been formerly furnished to the user by the

¹⁷ Distributed Semantic Media Wiki < <http://m3p.gforge.inria.fr/pmwiki/pmwiki.php> >

wiki network semantic context. The wiki network can thus be considered as a structured information support for intelligently querying and mining the Web where clustering processes can be used (Lamirel, 2006). On the author's side, relevant semantic roles that should be included in the wiki context can be selected through looking up a large amount of unstructured Web data. In such a case, one can also rely on the help of clustering processes in combination with the use of wiki network metadata and external annotation sources, in order to organize the querying results in a suitable way with the final goal of facilitating the author's decisions.

In our case, an important task is to find out the main actors and the salient institutions of a domain. This implies highlighting their various potential roles in the given domain, as well as to characterize the nature of their relationships in the social networks associated to their disciplines. This kind of information can only be obtained by a large scale querying process, stacking a sufficient amount of information to bring out reliable hypotheses and conclusions. It thus led us to consider intelligent and guided access to external data through the use of existing wiki metadata. A main challenge is thus to be able to isolate wiki strategic information as authors or institution names in a flow of unformatted data. This approach relies itself on the global domain of automated techniques for labeling named entities. The current statistical systems that could be used in this context need to exploit a great quantity of pre-annotated data to learn all the possible forms of the named entities. In this case, it is thus necessary to label a corpus, which will serve as training tool. Since this task is quite unaffordable with limited human resources, recent initiatives such as DBpedia (Bizer, 2009) or Yago (Suchanek, 2009) seek to provide likely semantic corpora to help in designing labeling tools. In the same spirit, some semantic ontologies, such as NLGbAse,¹⁸ are largely directed towards labeling. In our own case, the WICRI network itself can also play the role of a particularly rich database for mining reliable labeling information.

6. Conclusion

About 18 months ago, we launched the WICRI initiative, in order to highlight the central role that can be played by wikis in the research and innovation communities. The need for an approach based on a network of wikis was apparent and was confirmed by the further setting up of environmental oriented wikis. Doing so, we have been and we are still facing many difficulties, to which our answers are sometimes only partial and unsatisfactory. Yet, the networked way of managing feels superior, though harder to set up, than isolated services.

The quality and consistency of the network are correlated with the quality of its metadata. To improve it from a technical point of view, Semantic MediaWiki allows skipping a step, in a data-centric approach. A wiki works as a light structured CMS, "a set of html pages", which could be boosted by RDF annotations or micro formats. Our feeling now is that a better handling of XML by a wiki is a key issue.

A wiki is also a cooperative place where specialists work altogether. Remembering that many of them are using strong formalisms (such as LaTeX), their training and education open an immediate way of improving qualities of metadata and semantic models in the WICRI network.

References

- Bizer, Christian; Lehmann, Jens; Kobilarov, Georgi; Auer, Soren; Becker, Christian; Cyganiak, Richard; Hellmann, Sebastian (September 2009). "DBpedia - A crystallization point for the Web of Data". *Web Semantics: Science, Services and Agents on the World Wide Web* 7 (3): 154-165. ISSN 1570-8268
- Correndo, G., Alani, H., & Smart, P. (2008). A community based approach for managing ontology alignments. In *The 7th International Semantic Web Conference* (p. 61). From <http://eprints.ecs.soton.ac.uk/16673/>
- Ducloy, Jacques, Yann Nicolas, Diane Le Hénaff, Muriel Foulonneau, Luc Grivel, Jean-Paul Ducasse. *Metadata towards an e-research cyberinfrastructure: the case of francophone PhD theses*. Proceedings of DC 2006, Manzanillo, Mexico, 2006. , from <http://dcpapers.dublincore.org/ojs/pubs/article/view/846>.

¹⁸ < <http://www.nlgbase.org/publi.html> >

- EuroCRIS (2009). Recording Research. Report for CRIS seminar September 2009. Retrieved February 10, 2010, from <http://www.eurocris.org/fileadmin/Upload/200909.pdf>.
- Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. Yago: a core of semantic knowledge. In Proceedings of the 16th international conference on World Wide Web (WWW '07). ACM, New York, NY, USA, 697-706. DOI=10.1145/1242572.1242667 <http://doi.acm.org/10.1145/1242572.1242667>.
- Hodis, Eran (2008), Jaime Prilusky, Eric Martz, Israel Silman, John Moulton and Joel L. Sussman. Proteopedia - a scientific 'wiki' bridging the rift between 3D structure and function of biomacromolecules, Genome Biology 2008, doi:10.1186/gb-2008-9-8-r121. From <http://genomebiology.com/2008/9/8/R121>.
- Jeffery, Keith (2007). Technical infrastructure and policy framework for maximising the benefits from research. Proc. of the 11th Int. Conf. on Electronic Publishing, Vienna, Austria 13 June 2007. Leslie Chan and Bob Martens. ISBN 978-3-85437-292-9, 2007, pp. 1-12, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.5044>.
- Krötzsch, Markus, Denny Vrandečić, Max Völkel, Heiko Haller, Rudi Studer (2007). Semantic Wikipedia. In: Journal of Web Semantics 5/2007, pp. 251–261. Elsevier 2007.
- Lagoze, Carl, Dean Krafft, Sandy Payette, and Susan Jesuroga. (2005, November). What is a digital library anyway, anymore? Beyond search and access in the NSDL. D-Lib Magazine, 11(11). Retrieved, January 10, 2007, from <http://www.dlib.org/dlib/november05/lagoze/11lagoze.html>.
- Lamirel, Jean-Charles (2006), and Shadi Al Shehabi. MultiSOM: a multiview neural model for accurately analyzing and mining complex data. In Proceedings of the 4th International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV), London, UK, July 2006.
- Lange, Christoph (2008). SWiM – a semantic wiki for mathematical knowledge management. In Sean Bechhofer and al, editors, ESWC, volume 5021 of Lecture Notes in Computer Science, pages 832–837. Springer, 2008.
- Sarre, S., Foulonneau, M. (2010) "Reusability in e-assessment : Towards a multifaceted approach for managing metadata of e-assessment resources", Fifth International Conference on Internet and Web Applications and Services.
- Sire, Stéphane (2010), Christine Vanoirbeek, Vincent Quint, Cécile Roisin. Authoring XML all the time, everywhere and by everyone. Proc. of XML Prague 2010, p. 125-149, Institute for Theoretical Computer Science, March 2010.
- Tifous, A., El Ghali, A., Dieng-Kuntz, R., Giboin, A., Christina, C., and Vidou, G. 2007. An ontology for supporting communities of practice. In Proceedings of the 4th international Conference on Knowledge Capture (Whistler, BC, Canada, October 28 - 31, 2007). D. Sleeman and K. Barker, Eds. K-CAP '07. ACM, New York, NY, 39-46. DOI=<http://doi.acm.org/10.1145/1298406.1298415>.
- Tudorache, Tania (2008), Natalya F. Noy, Samson Tu and Mark A Musen. Supporting collaborative ontology development in Protégé. In: Lecture Notes In Computer Science; Vol. 5318 archive Proceedings of the 7th International Conference on The Semantic Web.