# **Collaborative Semantic Annotation of Digitized Old Documents**

Mohamed El Ouaazizi, Noureddine Chenfour Laboratory of Informatics and New Technologies of Information and Communication Faculty of Sciences Dhar El Mahraz Sidi Mohamed Ben Abdellah University Fez, Morocco {elouaazizi\_med, chenfour}@yahoo.fr



**Keywords:** Digital Libraries, Digitized Documents, Semantic Web Technologies, Collaborative Semantic Annotation, Data Interoperability

Received: 18 June 2012, Revised 29 July 2012, Accepted 3 August 2012

© 2012 DLINE. All rights reserved

#### 1. Introduction

The architecture of the current web is based on the use of hypertext in order to provide online documents with the best presentation possible and easily exploitable by the humanuser. However, machine help remains limited to navigation between documents and their presentation for the users and it cannot access and uses their documentary content [1]. On the other hand, in the *Semantic Web* (SW) [2] the content is structured into a set of web resources identified as unique, interlinked, described with standardized metadata's, accessible and exploitable by various Semantic Web Agents. In this second context, we propose a new model of a platform that follows the Semantic Web Vision to online publishing digitized documents<sup>1</sup> that are enriched with a descriptive metadata. In addition to several features to assist and support the collaborative *Semantic Annotation* of these documents, the descriptive metadata are accessible and exploitable by various *Semantic Web Applications* (SWA). The interoperability is offered at the level of *Semantic Data*, unlike the case of '*classic libraries*' where only the data exposed by web services are accessible.

In the remainder of this paper, we will present the software architecture of our platform and we explain in detail the two aspects semantic and collaborative of our model.

<sup>1</sup>We are particularly interested by the digitized Arabic manuscripts, but with our platform, any digitized document can be used

For experimentation, we will take the manuscript of *Avicenna* "Al Qânûn fi At-Tibb" as an example to explain how to realize his description at each step of the *Semantic Annotation Process* and how this description will be used by *Semantic Web Applications*. Finally, we will give a conclusion, proposing some new features to improve our model.

# 2. Our Platform Architecture

The main objective of our platform is the online publishing digitized old documents and particularly the Arabic manuscripts of our cultural heritage, while providing all necessary means to access and explores their documentary content. We will follow the *Semantic Web Approach* [3] and we will consider that all items on our platform (*Users, Collections, Documents, Images, Fragments, Repositories ...*) as resources identified by their **Uniform Resource Identifiers (URI)** and can be described by a system of standardized metadata's. It is on this principle that is based the entire software conception of our platform "*Figure* 1".

To provide a collaborative work environment, simple and efficient for the *Semantic Annotation* of digitized documents, we took into account several constraints. The most important constraint is related to the nature of the documents that we have chosen for online publishing as image format and to allow access and use of their documentary content not only by human user, but also by software agents. We have defined a flexible structure of all resources and of their descriptive metadata that has allowed us to obtain a good quality of the implementation of different features and an easy management of all resources manipulated on our platform.







### 3. Storage and Access to Data

# 3.1 Documentary and ontological resources

The digitized documents are added to the platform as *collections of images*<sup>2</sup> stored in different folders. Each image is accessible via its *Universal Resource Locator* (*URL*) and can be viewed by a human user. However its documentary content cannot be exploited by a machine, because there is a significant semantic gap between its visual appearance and the information significance that it contains. To reduce this semantic gap, we use *Semantic Annotation* [4] to enrich the images with semantic metadata that are defined in shared vocabularies. All necessary vocabularies are first imported and stored as *RDF files* on our platform to ensure their availability.

### 3.2 The descriptive data of the resources

In the literature, there are two types of annotations [5]: (1) the annotation as a new free data attached to a documentary content. This type of annotation is used to add to a documentary content a note for clarification, an comment or other, to be reused by the human user [6], (2) the annotation as a *standardized description* of resources. This type of annotation attaches shared metadata to resources that are identified by theirs URIs. The goal is to be reused by different software agents.

The **Resource Description Framework**<sup>3</sup> (RDF) is a data model for describing the resources by using the second type of annotation. In this model, a resource is anything identified with a URI. It can be described with a set of **RDF triples** <*Resource*, *Property, Value>*. The used properties are also identified by theirs URIs and defined in controlled vocabularies. The value of a property can be either a *literal* or *another resource*. This set of *RDF* triples, represents a formal description of the resource and it can be used by different Semantic Web Agents.

The descriptive data associated with different resources manipulated on our platform, are stored as *RDF graph* in a data repository. A *Relational Database Management System (RDBMS)* cannot be used to manage semi structured data, because we can't anticipate the creation of a suitable scheme for the storage of all different properties that will be associated with different resources. In this case, the use of data repositories proves to be necessary. We use three data repositories:

- Parameter: to store all platform and users profiles parameters
- Annotation: to store the finalized annotations associated with different annotated documents.
- Session: to temporarily store the annotations manipulated during the Annotation Sessions.

# 3.3 Data interoperability

A data repository can be seen as a table with three columns for storing RDF triples. With this generic structure, *Semantic Data Interoperability* is now possible. They can be extracted and used by various *Semantic Web Applications*.

The *Simple Protocol and RDF Query Language*<sup>4</sup> (*SPARQL*) is a *language* standardized by *W3C* for querying *RDF data* stored in repositories. Most servers of RDF repositories (*Triple Store*) implement a web service called *SPARQL endpoint*, that allows a *Semantic Web Agent* to access through a *SPARQL query* to a set of RDF data, returning the result in various formats (*HTML*, *XML*, *RDF/XML*, *RDF/JSON*, *N3*, *Turtle* ...).

Our platform provides interoperability of *Semantic Data* stored in the repository *Annotation*. Data access is not restricted as in the case of '*classic libraries*' where only the data exposed by web services are accessible "*Figure* 2".

### 4. Semantic Aspect

In this section, we will explain how to use the RDF data model in the *Semantic Annotation Process* of digitized documents. To provide a rich and precise description of the documents and their documentary content we use three types of annotations. In

<sup>2</sup> We want to preserve the documentary content and form of the original documents

<sup>4</sup> The SPARQL protocol is based on WSDL that standardizes the description of Web services and their access

<sup>&</sup>lt;sup>3</sup> All Semantic Web Technologies used in this paper are standardized by W3C (See http://www.w3.org/)

our model, the description of a document is not only a simple bibliographic notice but also an explicit description of its documentary content.

### 4.1 Administrative description of documents

The *Super-Admin*<sup>5</sup> of our platform supplies our platform with new digitized documents. Each document is stored in a particular folder as a collection of images.

A document is a resource who's its URI corresponds to the domain *URL* of our platform followed by the address of its storage folder. The *Super-Admin* proceeds to its description by using the *Dublin Core Terms (DC)*. This vocabulary is widely used on the *Semantic Web* to describe bibliographic resources. It is simple and sufficiently rich to provide a *summary description "Figure 3"* of any document. In addition, the *Super-Admin* can use other vocabularies in order to add more descriptions about the document, for example with the *Creative Commons (CC)* vocabulary he can specify the license of its use.

```
<!- Document description ->
<foaf:Document rdf:about= "http://localhost:8080/smdl/docId">
  <dc:title xml:lang="ar">Al-Qânûn fî At-Tibb</dc:title>
  <dc:title xml:lang="fr">Le Canon de la Médecine</dc:title>
  <dc:creator rdf:resource="http://localhost:8080/smdl/autId"/>
  <dc:description xml:lang="fr">
   En cinq volumes : (vol.1)Anatomie du corps humain,
   (vol.2) Pharmacologie et règles de la médecine expérimentale,
   (vol.3)Pathologie, (vol.4)Fièvres, (vol.5)Pharmacologie.
  </dc:description>
  <dc:type>Medecine</dc:type>
  <dc:language>ara</dc:language>
  <dc:date>1100</dc:date>
  <str:isClassifiedIn rdf:resource="http://localhost:8080/smdl/colId"/ >
  <dc:riqnhts
   resource="http://creativecommons.org/licenses/by-nc-nd/3.0/"/>
</foaf:Document>
<!- Author description ->
<foaf:Person rdf:about="http://localhost:8080/smdl/autId">
  <owl:sameAs rdf:resource="http://fr.dbpedia.org/resource/Avicenne"/>
  <foaf:name>Ali IBN SINA</foaf:name>
  <foaf:familyName xml:lang="en">Avicenna</foaf:familyName>
  <foaf:foaf:primaryTopic
  rdf:resource="http://fr.wikipedia.org/wiki/Avicenne"/>
</ foaf:Person >
```

Figure 3. Description example of a document and its author

# 4.2 Semantic Annotation of documentary content

This phase is the most delicate and the most important. It requires that the annotators have a good understanding of documentary content and also of the domain of knowledge approached by the document. In this phase, the annotators will provide a *detailed description* "*Figure 7*" of documentary content on the basis of their competence and their understanding of the main ideas which is contained in the document. The quality of the annotation depends strongly on the capacity of the annotators to identify in the documentary content, the fundamental concepts and the existential semantic relationships between these concepts which constitute the cognitive foundation of document.

On the *Semantic Web*, there are several well-formalized ontologyies ,which can be imported and used in the *Semantic Annotation Process*. In case we can not find an ontology which models the knowledge's provided by the document, we can create one that

<sup>5</sup> See the section "*Types of Users*" at the bottom of this paper

Journal of Information & Systems Management Volume 2 Number 3 September 2012 133

meets the needs of the annotation. Thecreation of ontology [7] is a creative process whose quality depends solely on the expertise of the designer in the domain of knowledge and his capacity to model this domain. Our platform provides assistance for the creation all the components of ontology.

```
<str:TextArea rdf:about= "http://localhost:8080/smdl/docId/frgId">
    <dc:title>Gingembre</dc:title>
    <dc:subject
    resource="http://fr.dbpedia.org/page/Catégorie:Plante_médicinale"/>
    <str:content>
    <!- Textual content of fragment->
    </str:content>
    <oul>
        coul:sameAs rdf:resource="http://fr.dbpedia.org/resource/Gingembre"/>
        <foaf:primaryTopic
        rdf:resource="http://fr.wikipedia.org/wiki/Gingembre"/>
        </str:TextArea >
```

Figure 4. Detailed description example of a documentary fragment



Figure 5. Structural ontology (Prefix str)

The *Semantic Annotation* is an operation which consists of instantiating the *documentary fragments* on the concepts and on the relationships of ontology [8]. The set of instances is a formal and explicit description of the documentary contents. This

*knowledge base* can be exploited by different software agents and inference engines<sup>6</sup> which can realize various tasks, such as targeted research, extraction and integration of information, etc....

#### 4.3 Structural description

We have conceived a structural ontology "Figure 4" to describe the logical structure of the digitized documents that we publish online. The conception of this ontology is based on the fact that a digitized document is a collection of images where each image is a representation of a document page that can include important documentary fragments. This ontology has been created by extending of the **FOAF** ontology and by the using the terms of **DC** vocabulary. We have extended the **FOAF** ontology with new concepts and specialized of other concepts. By following a content analysis of manuscripts, we have determined that a documentary fragment is an sub class of class *foaf:Image* that can be present in various forms (*Text Area, Table, Graphic, Drop Cap, Signet, Illustration ....*)

During an *Annotation Session*, an annotator analyses the images of document and identifies important documentary fragments. In order to define any fragment as a resource, the annotator uses a *graphical tool* "*Figure 5*" which is provided by the annotation help service of our platform to delimit the image area that corresponds to fragment. Automatically, a resource is created as *str:Fragment* type with a set of necessary *structural properties* "*Figure 6*" needed to characterize the fragment, such as its relative position in the image, its size etc. .... Also, a semantic relationship *str:isPartOf* is added automatically to this resource to specify the document that contains this fragment.

Collaborative Semantic Annatation of Digitized Old Documents	User
DOC App. Document Ontology Core Application	
	Documents
محمد المحمد المحمد المحمد المحمد المعالية المحمد المحمد المحمد الواحد الواحد الواحد المحمد المحمد المحمد المحم التعديم عليت الرائحة والتي لمن المحمد المحمد والمحمد المحمد المحمد الواحد الواحد المحمد المحمد المحمد المحمد ويستعمل المحمد الرائحة والمحمد المحمد ال التعرين المحمد المحم التعرين المحمد المحم المحمد المحم المحمد المحمد المحمد المحمد المحم المحمد المحمد المحمد المحمد المحمد المحم المحمد المحم المحمد المحم المحم المحم المحم المحم المحم المحم المحمد المحم المحمد المحم المحم المحم المحم المحم	<ul> <li>http://localhost.8080/smdl/dec321</li> <li>http://localhost.8080/smdl/dec1d(Sessive)</li> <li>decrements = "Al-Qimin &amp; Ar Tibb"</li> <li>decrements = http://localhost.8080/smdl/smtld</li> <li>decrements = "Madacine"</li> </ul>
	Fragments
بارس با المتعد وبمرد طريقة تصابط المدينة حد رياس لاكة الرطوية لشنية عبد في الاقسال والحواس في حرارت الرقم ولا على المعمد من المعيد من حوالية الالمدينة في اعصا عنون مدين محد اللغة حراري العق المسرعين رقيقة التصابط وجف التق على المعامل في جريد في المعام العامي مع ميران والعلق في العقب العلي بنايج حجار اللغة العربي الرطوية منذ و شريا في المعام العامي و بوانس والعن والعقب العلي المجيد عليه الذار العربي بل محد عنها من الرطوية المعام العامي في مع والات موالي الرائس ويقع من محيو العام في المعام العامي في العام الرطوية من محيو والعام في المعام العامي في العام و بلغم من محيو العراج و بلغم من محيو العراج	<ul> <li>Imp://localhost.8080/smdl/docld/hgl23</li> <li>Imp://localhost.8080/smdl/docld/hgl23</li> <li>Imp://localhost.8080/smdl/docld/hgl23</li> <li>Imp://localhost.8080/smdl/docs/TextArea</li> <li>Acombject = http://f.dbpedia.org/page/Catégorie:Plante_médici</li> </ul>
	Pedichaman's Table -

Figure 6. Our Web interface for Semantic Annotation of documentary content

<sup>6</sup> For example, an inference system can deduct that a resource of type str:TextArea is a sub type of str:Fragment class

Journal of Information & Systems Management Volume 2 Number 3 September 2012

```
<TextArea rdf:about="http://localhost:8080/smdl/docId/frgId">
<str:isPartOf rdf:resource="http://localhost:8080/smdl/docId"/>
<str:pageNumber>166</str:pageNumber>
<str:posX>10</str:posX>
<str:posY>70</str:posY>
<str:width>580</str:width>
<str:height>280</str:height>
</TextArea >
```

Figure 7. Structural properties of a documentary fragment

# 5. Collaborative Aspect

In this section, we will explain the role played by each registered user and how he collaborate with other users in an Annotation Session to enrich the available documents on our platform with a descriptive metadata of their documentary content.

# 5.1 The users registration and authentication

To benefit from the services offered by our platform, at first, the user must be registered at our platform by providing a set of necessary information's about his personal profile "*Figure* 8". Other optional properties about his skills can be added too. After registration, the user will receive an email containing its own account that enables him to connect to the platform.

Any logged user<sup>7</sup> to our platform can explore all the available annotations in the two repositories *Parameter* and *Annotation*. For example, he can explore the annotations associated with a document, the user profiles that have participated at least once to an *Annotation Session* or consult the documentary resources that meet its exact needs, etc...

```
<foaf:Person rdf:about="http://localhost:8080/smdl/usrId">
<foaf:name>M. EL OUAAZIZI</foaf:name>
<foaf:mbox rdf:resource="mailto:elouaazizi_med@yahoo.fr"/>
<foaf:phone rdf:resource="tel:+212660700600"/>
<foaf:img rdf:resource="http://localhost:8080/smdl/images/usrId.jpg"/>
</foaf:Person>
```

# Figure 8. An example of description of a user

```
<rdf:RDF xmlns:...>
<foaf:PersonalProfileDocument rdf:about="#me">
<foaf:maker rdf:resource="#me">
<foaf:primaryTopic rdf:resource="http://website.net/usrDesc"/>
</foaf:PersonalProfileDocument>
<foaf:Person rdf :ID="me">
<!- -
Description of user "me"
- ->
</foaf:Person>
</rdf:RDF>
```

# Figure 9. Personal profile document of a user

The user profiles are stored into the repository Parameter. They are created by using the *Friend Of A Friend (FOAF)* vocabulary. A user is a resource of type *foaf:Person*, identified by a URI that is defined as unique on our platform. A user profile can be enhanced by associating other properties to resource such as *foaf:givenName, foaf:img, foaf:phone*, etc....

<sup>7</sup> A user connects to our platform is a user of type str:Guest

Journal of Information & Systems Management Volume 2 Number 3 September 2012

If a user already has an *RDF profile document* that describe his personal profile, all the descriptions included in this document can be imported into the repository *Parameter* and used for his description on our platform.

### 5.2 Types of users

On our platform, we distinguish four types of users based on the role played by each of them: *Guest, Annotator, Admin* and *Super-Admin*:

• *Guest*: Is a user who is authorized to consult all available information's on the platform, such as the documents, semantic annotations and profiles of annotators.

• Annotator: Is a user who participates in at least one Annotation Session. In addition to the offered possibilities for Guest user type, he can create and update all the annotations that he has associated to different documents during the Annotation Sessions. He can also consult the annotations added by other annotators who participate at the same sessions of annotation.

• *Admin*: Is an annotator who manages a *Annotation Session*. He has the right to choose the annotations created by annotators of group that he manages and add them to its annotation context. He also manages his annotators group, for example, it can eradicate the annotators considered incompetent for the annotation of the documents.

• *Super-Admin*: Is the platform administrator. He has full access to the all features and to all data on our platform (*Collections, Documents, Annotations, Vocabularies, ...*)

### **5.3 Semantic Annotation Process**

By using the workflows "*Figure* 10", we describe the followed steps on our platform to realize the *Semantic Annotation* of digitized documents in a collaborative manner.



Figure 10. Semantic Annotation Process workflows

Journal of Information & Systems Management Volume 2 Number 3 September 2012 137

At first, the *Super-Admin* adds to our platform the new digitized documents as collections of images, then he use *DC* vocabulary to provide a summary description of these documents.

Any registered user can explore the administrative description<sup>8</sup> associated with each of the available documents. If a user is interested by the annotation of a document that is not yet annotated, he can prescribe his annotation service. The property *str:hasAnnotator* is added to document resource with the URI of the annotator as value.

Once an *annotators group*<sup>9</sup> is formed, the annotators of group can designate a director of annotation session by a simple vote. Each annotator proposes two annotators to be a director on the basis of their personal profiles. This resulted in the automatic addition to the document of a property with value is the URI of annotator selected as director of Annotation Session. In case two or more annotators obtain the same result of vote , the competent director is determined automatically by comparing the properties associated with these annotators such as the number of times they managed annotations sessions, the number of times they participated in *Annotation Sessions*, the date of registration to the platform....

When the director of the *Annotation Session* is selected, he can start an *Annotation Session*. This is reflected by adding to the document resource a *str:startSession* property with the date and time of the system<sup>10</sup> as a value, and by creating (if not yet created) of the temporary data repository Session to store annotations that will be associated with the document resource during this *Annotation Session*.

Annotator who participates in a *Annotation Session*, filed its annotations in its own context, which is reserved in the data repository *Session*. He has full access to his own context. He can add new annotations and can update his existing annotations. He can also explore the annotations deposited by the other annotators who participate in the same *Annotation Session*. Annotators can also exchange their views about the document and its annotations through remarks, comments, criticisms, etc... . These exchanges are recorded and may serve as trace for this *Annotation Session*.

The director of *Annotation Session* has full access to all data of an Annotation Session. When he remarks that the annotation of documentary content have reached a sufficient level of acceptable quality, he can close the *Annotation Session*. The property *str:endSession* is added to the document resource with the date and time of system as value. The temporary descriptive annotations of the documentary content of the document are transferred from the repository Session to finalized annotation repository *Annotation*. The new annotations become accessible on the web.

# 6. Experimentation and Results

Our platform offers the interoperability at the level of *Semantic Data*. A *Semantic Web Application* has full access to the *Semantic Data* available in the *Annotation* reposito and can extract these data by using the SPARQL query language. Our SPARQL endpoint located at the *URL "http://localhost:8080/smdl/sparql*" return the result in the XML format as is defined in the *SPARQL* result language.

We show through three example of *SPARQL* query how to find a particuler manuscript, after, how to find in this manuscript the documentary fragments of a particular type and at the end we will querying our SPARQL endpoint to get all the *Semantic Data* associated with a fragment.

In figure "*Figure* 10", we have (1) a SPARQL query for searching all the documents classified in a collection that has a property *str:ddc* with value is equal 610 and (2) the returned *SPARQL* result.

In the second example "*Figure* 11", we have chosen a *URI* of one document returned in the previous result to locate in this document all resources fragment that have the resource "*http://fr.dbpedia.org/page/Catégorie:Plante\_médicinale*" as value of property *dc:subject*.

<sup>&</sup>lt;sup>8</sup> The administrative description is returned to user in the HTML format

<sup>&</sup>lt;sup>9</sup> The number of annotators per group is defined by the property str:groupeSize in our structural ontology

<sup>&</sup>lt;sup>10</sup> The date and time of system is comply with the ISO-8601

```
<!- (1) SPARQL query ->
select ?x
from http://localhost:8080/smdl/sparql
where {
?x str:isClassifiedIn ?y.
?y str:ddc 610.
}
<!- (2) SPARQL result ->
<?xml version="1.0" encoding="UTF-8"?>
<sparql xmlns="http://www.w3.org/2005/sparql-results#">
 <head>
  <variable name="x" />
  <variable name="y" />
 </head>
<results>
 <result>
  <br/><binding name="x">
   <uri> http://localhost:8080/smdl/docId </uri>
  </binding>
 </result>
 <result>
  <br/><binding name="x">
   <!-Another result ->
  </binding>
  </result>
 </results>
</sparql>
```

Figure 11. SPARQL query example and the corresponding result

```
select ?x
where {
    ?x rdf:type str:Fragment;
    str:isPartOf
    http://localhost:8080/smdl/docId;
    dc:subject
    http://fr.dbpedia.org/page/Catégorie:Plante_médicinale.
}
```

Figure 12. SPARQL query example for searching a documentary fragments

```
select ?x ?y
where {http://localhost:8080/smdl/docId/frgId ?x ?y.}
```

```
Figure 13. SPARQL query for extracting all properties attached to a fragment
```

Finally, we can submit the following query "Figure 12" to our SPARQL endpoint to extract all properties and their corresponding values that are attached to one of returned fragments in the previous example.

### 7. Conclusion and Perspectives

To get out of the theoretical framework around the *Semantic Web*, we have proposed a new model of a platform for collaborative *Semantic Annotation* of digitized old documents and we have concretely shown how to implement a *Semantic Web Application* 

Journal of Information & Systems Management Volume 2 Number 3 September 2012 139

using several *Semantic Web Technologies* to online publishing of an important content that is totally informal. With *Resource Oriented Architecture (ROA)* and the *Semantic Web* possibilities, it has become possible to render an informal documentary content accessible and exploitable by various *Semantic Web Agents*. Our platform, that we want it to be simple and effective, can benefit from several features and improvements such as managing data security and storing of document in files of multiple-image format while managing extracting important documentary fragments by using their *URIs*.

### References

[1] EL, M., Ouaazizi, Chenfour, N., Benslimane, R. (2011). Evolution de l'acte communicatif sur le web. *In*: La Communication: Ses dimensions Linguistiques, Psychologiques, Éducatives et Techniques. p. 41-47, October.

[2] Michael Uschold. (2003). Where are the Semantics in the Semantic Web, *AI Artificial Inteligence Magazine*, 24 (3) 25-36, September.

[3] Roger Roberts. (2008). Introduction to Semantic Web technologies, In: AIDA informazioni, Italie, p.103-112.

[4] Guus Schreiber, Alia Amin, Lora Aroyo, Mark van Assema, Victor de Boer, Lynda Hardman, Michiel Hildebrand, Borys Omelayenko, Jacco van Osenbruggen, Anna Tordai, Jan Wielemaker, Bob Wielinga. (2008). Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture demonstrator, *Journal of Web Semantics*. 6 (4) 243-249, November

[5] Prie, Y., Garlatti, S. (2004). Méta-données et annotations dans le Web sémantique. *In*: I3 Information-Interaction-Intelligence, p. 45-68.

[6] El Ouaazizi, M. (2007). Développement d'une application JDI basé sur l'API JAI, Memory of DESS computer Engineering, p. 35-82, June.

[7] Michel Héon, Gilbert Paquette, Josianne Basque. (2009). Méthodologie assistée de conception d'une ontologie à partir d'une conceptualisation consensuelle semi–formelle, in acts of 20es Francophone Days of Knowledge Engineering, Knowledge and online communities, p. 61-72, May.

[8] Yiwen Wang, Natalia Stash, Lora Aroyo, Peter Gorgels, Lloyd Rutledge, Guus Schreiber. (2008). Recommendations based on semantically enriched museum collections, *Journal of Web Semantics*. 6 (4) 283-290, November.