

Supporting Collaborative Learning Activities with a Digital Library and Annotations

Tiago Rios da Rocha¹, Roberto Willrich¹, Renato Fileto¹, and Saïd Tazi^{2,3}

¹ Dept. of Computer Science, UFSC, Brazil

{tiagoriosrocha, willrich, fileto}@inf.ufsc.br

² LAAS-CNRS, Toulouse, F-31077, France

tazi@laas.fr

³ Université Toulouse; UT1, UPS, INSA, INP, ISAE; LAAS, Toulouse, France

Abstract. Digital Libraries (DLs) usually provide facilities for browsing and searching their collections, and can enhance noticeably learning activities. The integration of an annotation tool with a DL can foster knowledge exchange between instructors and learners. It is important that an annotation system for DLs should be easily integrated with existing DLs. This paper presents an annotation system, called DLNotes, which can be easily embedded in DLs in order to enable free-text and ontology-based annotations. DLNotes also supports supervised annotation activities and allows discussion threads to be associated with each annotation, what is particularly important for e-learning.

Keywords: Digital Libraries, free-text annotations, ontology-based annotations.

1 Introduction

Digital Libraries (DLs) are systems that provide facilities for organizing digital collections. They enable the creation of metadata catalogs that facilitate the information discovery from the DLs' collections. However, the basic facilities of a DL are not sufficient in the e-learning context. The editing and annotation of documents and the interactions among users need to be integrated [1]. An annotation can serve several objectives in the e-learning context [2]: procedural signaling for future attention, place marking, aiding memory, problem-working, interpretation, and tracing progress through difficult narrative or reflection about the material circumstances of reading. Moreover, annotation facilities embedded in a DL allow users to collectively build knowledge, enabling teaching and learning actors to communicate, and facilitate pedagogical evaluation.

Annotations fall into two broad classes: free-text and semantic annotations. Free-text annotations associate a selected portion of a document to a free-text providing additional information about the marked passage. Semantic annotations or ontology-based annotations associate terms occurring in a portion of a document to their respective semantic descriptions, which can be contained in an ontology or in a Knowledge Base (KB). These annotations help to make the knowledge explicit in a formal and machine-processable way, enabling semantic-based access [3]. Several

works have shown that this is a promising technology for e-learning [4]. However, most annotation systems deal with either free-text or semantic annotations, while both are relevant for e-learning.

This paper presents DLNotes, an annotation system for DLs allowing the creation and sharing of knowledge about the DLs contents for e-learning activities. By using DLNotes, instructors and learners can freely enrich the DL contents, by associating relevant parts of documents with both free-text and semantic annotations. The instances and relationships generated by the DLNotes semantic annotation process are stored and related in KBs that can be used for semantic browsing and semantic information retrieval. DLNotes also supports collaboration thanks to public annotations and discussion threads associated with annotations. Discussion threads are mechanisms particularly important to allow communication between learners and instructors about annotated subjects. Moreover, by using DLNotes instructors can propose tasks to their learners, such as the collaborative construction of sets of associated annotations. The execution of these tasks can be guided and evaluated by the instructors.

The remainder of this paper is organized as follows. Section 2 presents an overview of related work. Section 3 describes the proposed annotation schema and the DLNotes' architecture. Section 4 shows DL-Notes functionalities in use with the Brazilian Literature Digital Library [5]. Finally, section 5 presents the conclusions and future works.

2 Related Work

We consider that an annotation system for DLs should meet the following requirements: (i) both free-text and semantic annotations must be supported; (ii) an access policy must be associated with annotations, that are allowed to be private or public; and (iii) the communication and discussion between instructors and learners must be supported. In this section we discuss related work, outlining the main features of some relevant annotation systems. None of the analyzed annotation systems satisfy all the requirements enumerated above simultaneously.

There are several free-text annotation systems. Some of these systems adopt or extend the annotation schema proposed by the W3C Annotea project [6]. Annotea defines an annotation schema based on the Resource Description Framework (RDF) [7] and a protocol for publishing, retrieving and managing annotations. RDF is a W3C standard language for representing information about resources in the Web. The annotation systems based on Annotea, such as Amaya [8], allow only simple free-text annotations.

A number of systems allow manual or semi-automatic creation of semantic annotations. Manual creation of annotations have a high cost, is prone to error, and usually require domain expertise [9]. Despite these drawbacks, we believe that manual annotation is interesting for e-learning, allowing the users to collaboratively annotate the DL collection and share knowledge. Moreover, it can be a learning activity by itself in some domains such as literature or scientific text analysis.

SHOE [10] and Ontobroker [11] allow the annotation of Web pages with ontological metadata using HTML extensions. The drawback of these systems is that they change the original document that is not allowed in DLs.

Some semantic annotation systems are oriented to learning. MemoNote [12] is a semantic annotation system for teachers. SABRE [13] is another annotation system oriented to learning. It allows the authors to semantically annotate their documents with their pedagogical intentions, with the objective of facilitating the reuse of annotated documents. However, these systems do not allow learners to create annotations and discussion threads during the learning process.

Vannotea [14] and COHSE [15] are examples of annotation systems supporting both free-text and semantic annotations. Vannotea extends the Annotea schema to annotate objects of any media type, where annotations can be free-text, files, URLs or terms from a controlled vocabulary (*e.g.*, WordNet) or ontology. The COHSE system also extends the Annotea schema to support three types of annotations: textual annotations, link annotations, and semantic annotations.

Various free-text annotation systems allow public and private (including group) annotations. Conversely, semantic annotations systems generally allow either public or private annotation. One exception is Vannotea. However, its authors do not detail the management of public and private semantic annotations.

What differs DLNotes from previous systems is that it supports free-text and semantic annotations, including public and private ones. Moreover, two other functionalities of DLNotes are particularly important for learning: (i) discussion threads about the annotated content, and (ii) the annotation activity concept, as detailed in the following. In addition, DLNotes can be adapted to different domains by changing the ontologies used for semantic annotation.

3 The DLNotes System

This section presents DLNotes, our free-text and semantic annotation system for DLs. The main purpose of this system is to support learning by offering functionalities to enrich the DL collection with annotations inserted by instructors and learners. Concerning free-text annotation, instructors may annotate the document to ask questions, to make suggestions or to discuss with learners. Learners may annotate text to answer questions posed by the instructor or discuss his answers with other learners or with the instructor. Annotations can support the learners' critical thinking and argumentation activities. The annotation process is characterized by interaction and exchange of experience, supported by shared public annotations and discussion threads about these annotations. Semantic annotations are made by both instructors and learners. They are used to identify concepts and relations of the studied matter that render a conceptualization of the studied texts. The system allows identifying text's features and helps in the analysis and comprehension processes.

3.1 Annotation Schema

DLNotes adopts an extension of the Annotea schema, satisfying the previously cited requirements. Some properties shown in Figure 1 are described below. The *rdf:type* property indicates the creator's intention when making the annotation. Its value can be *rdf:type Annotation* or some other specific type of annotation (Advice, Change, Comment, Example, Explanation, Question and SeeAlso). Each type is associated with an icon. The *annotates* property refers to the annotated document. The *context*

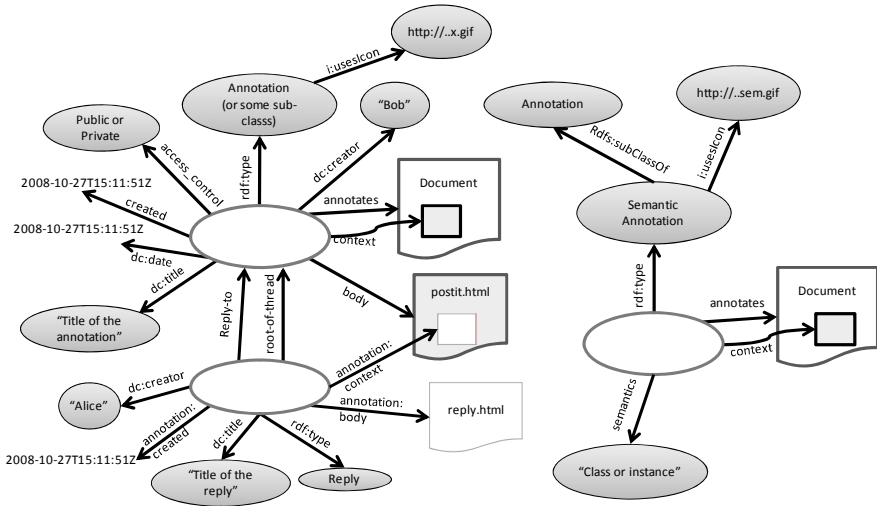


Fig. 1. Annotation Schema of DLNotes

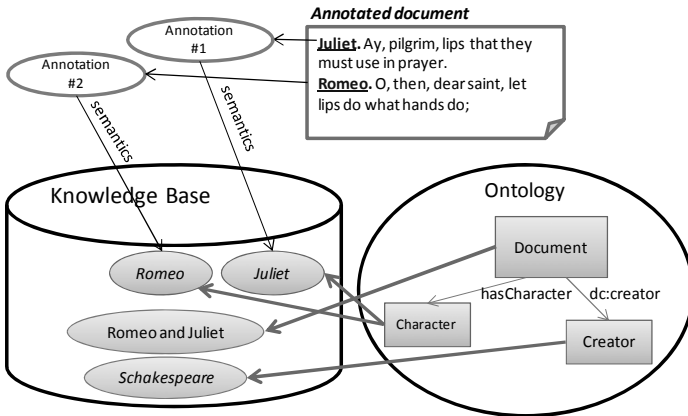


Fig. 2. Semantic Annotation with DLNotes

property specifies to the actual place of the annotation within the document with XPointer. The *body* property contains the content of the annotation. Annotea uses multiple RDF schemas, including Dublin Core (DC) [16], a set of standard metadata largely used in DLs. The descriptive properties *dc:title*, *dc:date* and *dc:creator* are taken from DC. The *created* property defines the date and time on which the annotation was created. Two properties keep tracks of discussion threads: *reply-to* defines which annotation or reply was the previous one in the discussion thread; and *root-of-thread* is the first annotation in the thread.

We have extended the Annotea schema by including new properties and a new annotation type. The new *access_control* (*Public/Private*) property allows defining

public and private annotations (group annotations are private annotations with more than one creator). Moreover, we defined a new annotation type, called *Semantic Annotation* and the property *semantics*. The latest associates a context with a class or instance defined in an ontology or KB via an URI.

In DLNotes, one or more ontologies and some KBs provide a vocabulary, specifying concepts, entities (concept instances), and their semantic relations. DLNotes allows the user to associate document fragments with ontological concepts or instances they refer to. If a new term or named entity is identified by the user, he/she can create a new instance in the respective KB when inserting the annotation. For instance, in Figure 2, the user can annotate the named entities Juliet and Romeo by creating two instances of the Character concept described in the domain ontology.

3.2 DLNotes Architecture

There are three categories of Web annotation systems [17]: server-based, where only documents on a specific server can be annotated; proxy-based, where the annotated documents are accessed through a proxy; and extension (or plug-in) for a Web browser, which requires the installation of a specific browser or plug-in. DLNotes is aimed to annotate HTML documents that make up a collection made available by a DL server. Therefore, we have adopted the server-based approach. This approach does not require any software installation or browser configuration. Moreover, the server-based approach simplifies the annotation sharing between the users of the DL. Figure 3 illustrates the overall architecture of the DLNotes system. Its main components are described in the following subsections.

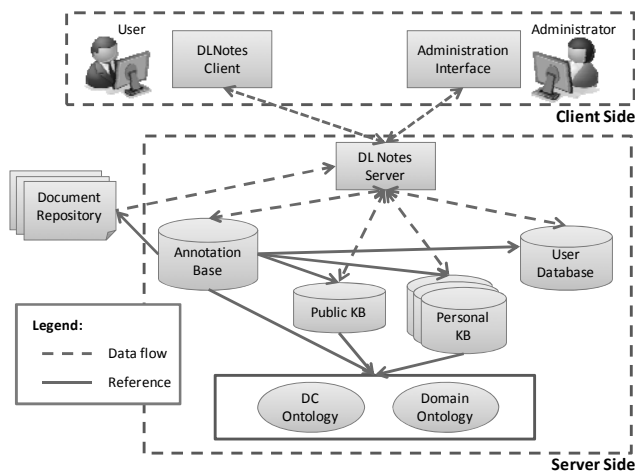


Fig. 3. Overall architecture of the DLNotes System

Dublin Core and Domain Ontologies. DLNotes uses Dublin Core (DC) and Domain Ontologies, all represented in RDFS (RDF Schema), in the annotation process. The Domain Ontologies specify the domain specific concepts and relationships handled by DLNotes in order to create semantic annotations. These ontologies can be set by the

administrator using the Administration interface, providing a way to customize DLNotes to specific domains. The DC Ontology is a predefined ontology with the DC elements. It is used to map some DL's metadata repository into individuals in the public KB.

User Database. The user database maintains users' IDs, passwords and profiles. The user profile maintains his preferences and privileges. The preferences indicate if the user wants to see public and/or private annotations. We defined four privilege levels: administrator, instructor, specialist, and regular. The public annotations created by regular users are published only after an approval of the administrator. Specialists and instructors can publish public annotations without authorization. Instructors can access the private annotations of regular users for evaluation purposes.

Administration Interface. The administration interface is a Web page allowing the DL's administrator to manage the DLNotes system. The main functionalities provided by this interface are: importation of Domain Ontologies; importation of the DL's metadata to a KB; creation of annotation activities; the management of KBs, User Base and Annotation Bases; and the management of public annotations.

The instructor can also propose annotation activities to his learners. An annotation activity is defined by a unique identifier and a list of annotation types that the user is authorized to create within its activity. When a learner accesses a document within an annotation activity context, he may create only annotations of the authorized types. For instance, if the user accesses a document in a *FreeTextOnly* annotation activity, he/she can create only free-text annotations.

Using the import ontology functionality the administrator can load the Domain Ontology to be considered by the annotation system, adapting DLNotes to the DL's domain. Another important DLNotes functionality is the mapping of the DL's metadata into instances at the Public KB. Two options are supported: importation of DC metadata using a standalone application, or using the OAI-PMH protocol [18] to harvest the DC metadata to populate the Public KB.

The management of public annotations includes the acceptance of new public annotations posted by regular users. When accepting a public semantic annotation referring to an instance, that instance must be moved from the Personal KB to the Public KB, if it is not already there. The administrator is responsible for the consistency of the Public KB, and can remove or edit inconsistent annotations.

DLNotes Server. The DLNotes Server offers an Application Programming Interface (API) used by the Digital Library and by the DLNotes Client to pose annotation requests. The only method currently available is *startAnnotationSession*, which starts an annotation session and returns the HTML page coding the annotated document. Three arguments can be passed to this method: the URL of the document fragment to be annotated; the annotation activity; and the user authentication information. The latest is optional. If it is not informed, the DLNotes Client requests the user login. After the user authentication, the DLNotes Server verifies the user profile and dynamically includes annotations.

DLNotes Client. The DLNotes Client allows the user to select a fragment of a document and create an annotation associated to it. The right side of Figure 4 illustrates the creation of a free-text annotation. The user can select a portion of the

document (Juliet in Figure 4) and add the annotation (to describe Juliet). The first step is to select the option “free-text annotation” and define if the annotation is public or private one. After that the user can provide its title and contents. When the annotation is saved, DLClient changes the HTML code representing the document to include a link to the created annotation and it sends this annotation to the DLNotes server.

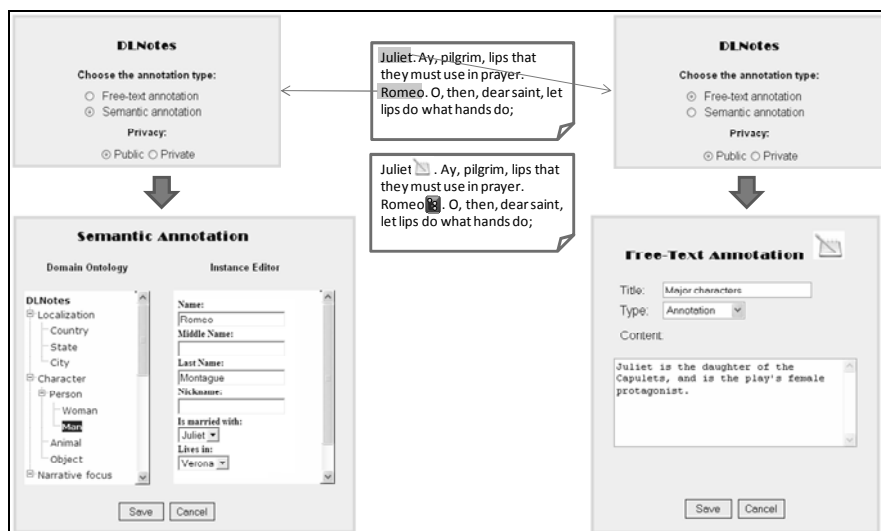


Fig. 4. Creating free-text and semantic annotations with DLNotes

The left side of Figure 4 illustrates the creation of a semantic annotation. When the user chooses the option “Semantic Annotation”, the DLNotes Client displays the Semantic Annotation window. Using this window the user can select one class and choose whether the annotation will be associated with a class or individual of that class. In the latest case, the user can select an existing instance in a KB or he can create the individual using the instance editor panel. During the instance creation process the user can create new relations with other entities on the KB. For instance, during the creation of an annotation for the *Character* Juliet, the user can specify that Juliet *isCousinOf* Tybalt (another instance of *Character*).

Annotation and Knowledge Bases. DLNotes stores all annotations in Annotation Base, which is separated from the documents themselves. These annotations are represented using the RDF annotation schema illustrated in Figure 1.

The Public KB maintains all instances and relationships obtained by the DL metadata importation, and those created by the users during the generation of public semantic annotations. This KB is the basis for the semantic browsing and querying. The Private KBs maintain instances and relationships created during the generation of private semantic annotations.

In DLNotes, any annotation can be the root of a discussion thread. When the user clicks in an annotation, the annotation content is displayed, accompanied by the existing discussion threads. The user can create a discussion thread or message, by

choosing the New Thread option associated with an annotation or the Reply option associated with a discussion thread.

3.3 Implementation and Integration of DLNotes in a DL

The DLNotes system is being implemented using the LAMP (Linux, Apache, MySQL, PHP) open source platform, AJAX (Asynchronous JavaScript And XML), and the RDF API for PHP (RAP) [19]. AJAX is used to implement all the interactivity with the user. It allows the user to interact with the DLNotes Client without the need to reload the entire Web page on each interaction. The RAP API allows parsing, querying and handling of RDF models.

An annotation system for DLs should be easily integrated with existing DLs. The integration of DLNotes with a DL requires changing how the user accesses the collection. The URL giving access to the documents should be changed to a URL calling the *startAnnotationSession* method. Moreover, the existing DL metadata repository must be taken into consideration. For that purpose, DLNotes provides the importation of this repository contents into the public KB.

4 A Case Study: DLNotes integrated with a Literature DL

DLNotes is being integrated with the Digital Library of Brazilian Literature [5], in order to demonstrate its use and conduct learning activities. The purpose of this DL is to make available a great collection of Brazilian and Portuguese literary texts of public domain, with information about authors, dates of publication, publishers, and literary genres. The available resources are used in several learning activities such as text comprehension and analysis at Santa Catarina Federal University (UFSC).

A small ontology was developed for testing DLNotes in the literature domain. This ontology defines basic concepts (such as character, protagonist, antagonist, narrator, and geographic place) and their relationships, extracted from literary texts, along with concepts related to learning activities using these texts.

The activity of literary analysis was chosen to demonstrate the possible uses of DLNotes to deploy annotations for e-learning purposes. This activity involves interpreting a literary work and arguing for a particular way of understanding it. The literary analysis in general evolves through several steps accomplished by the learner. Using the annotation activity concept of DLNotes, anyone can follow the guidelines to develop literary analysis. In each step, the user can access the literary work within an annotation activity, which limits the annotation types to be created.

In a “Character Identification” annotation activity, the learner can access the literary work to identify the characters and semantically annotate each character to identify its traits and relations (using *Character* class). Using free-text annotations, the user can write comments about each character.

In a “Figure of Speech” annotation activity, the learner can identify the figures of speech used by the author. It is done by selecting a paragraph and creating a semantic annotation associating this paragraph with a sub-class of *FigureOfSpeech* (e.g. *Antithesis*, *Apostrophe*, *RhetoricalQuestion*).

Another important activity is to create the plot, *i.e.* the arrangement of events and actions that make up a literary work. The user can identify some text and create semantic annotations to identify instances of the *Event* class or the *Action* class. Each one of these instances can be related with instances of *Character*, *Event* or *Action*. Semantic relations between instances of *Event* and *Action* can specify causal relationships (*e.g.*, an *Event* can cause an *Action*), conflict (struggle between opposing actions or events), and so on.

5 Conclusions

This paper presents DLNotes, an annotation system designed to satisfy the basic requirements of a DL for e-learning. DLNotes adopts an extension of the Annotea Schema that allows both free-text and semantic annotations. DLNotes users (learners and instructors) can identify important passages in the text, create different interpretations, and make relationships of the contents with formal semantics described in an ontology. DLNotes supports annotation ownership, access control, and discussion threads. The users can share their annotations with others and engage in organized discussions supported by the system, so that there is an exchange of experiences and the possibility of building knowledge through discussion, reflection and decision-making.

DLNotes provides facilities for easing and speeding-up the understanding DLs' contents. It can be adapted to different domains of DLs just by changing the domain ontology. It adds to e-learning technology, by supporting the cycle of production and aggregation of knowledge.

The following developments are among the future works: (i) integration of DLNotes with Moodle, (ii) services for semantic searches based on the knowledge acquired by using DLNotes, (iii) a module for making inferences on the knowledge base, (iv) techniques for mapping ontologies in order to integrate knowledge and provide unified access to diverse DLs, and (v) treatment of inconsistent knowledge.

References

1. Agosti, M., et al.: DiLAS: a Digital Library Annotation Service. In: Proceedings of Annotation for Collaboration - A Workshop on Annotation Models, Tools and Practices (2005)
2. Marshall, C.: Annotation: From Paper Books to Digital Library. In: ACM International Conference on Digital Libraries (DL 1997) (1997)
3. Gómez-Pérez, A., Corcho, O.: Ontology Languages for the Semantic Web. IEEE Intelligent Systems 17(1), 54–60 (2002)
4. Stojanovic, L., Staab, S., Studer, R.: eLearning based on the Semantic Web. In: World Conference on the WWW and Internet (WebNet 2001), pp. 23–27 (2001)
5. Willrich, R. et al.: Adaptive Information Retrieval System applied to Digital Libraries (in Portuguese). In: XII Symposium on Multimedia and Web, pp. 165–173 (2006)
6. Annotea, <http://www.w3.org/2001/Annotea/> (accessed in March 2009)
7. RDF, <http://www.w3.org/RDF/> (accessed in March 2009)
8. Amaya Browser, <http://www.w3.org/Amaya/> (accessed in March 2009)

9. Uren, V., et al.: Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics: Science, Services and Agents on the World Wide Web* 4(1), 14–28 (2006)
10. Luke, S., Spector, L., Rager, D., Hendler, J.: *Ontology-based Web Agents*. In: *Proceedings of First International Conference on Autonomous Agents* (1997)
11. Decker, S., Erdmann, M., Fensel, D., Studer, R.: *Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information*. In: *Database Semantics: Semantic Issues in Multimedia Systems*, pp. 351–369. Kluwer Academic Publisher, Dordrecht (1999)
12. Azouaou, F., Desmoulins, C.: *MemoNote, a context-aware annotation tool for teachers*. In: *7th Int. Conf. on Information Technology Based Higher Education and Training (ITHET 2006)*, pp. 621–628 (2006)
13. Tazi, S., Al-Tawki, Y., Drira, K.: *Editing pedagogical intentions for document reuse*. In: *4th IEEE Technology Based Higher Education and Training*, pp.274–278 (2003)
14. Schroeter, R., Hunter, J., Guerin, J., Khan, I., Henderson, M.: *A Synchronous Multimedia Annotation System for Secure Collaboratories*. In: *Second IEEE International Conference on e-Science and Grid Computing (e-Science 2006)*, p. 41 (2006)
15. Bechhofer, S., Carr, L., Goble, C., Kampa, S., Miles-Board, T.: *The Semantics of Semantic Annotation*. In: *ODBASE: First Int. Conf. on Ontologies, DataBases, and Applications of Semantics for Large Scale Information Systems*, pp. 1152–1167 (2002)
16. Dublin Core, <http://www.dublincore.org> (accessed in December 2008)
17. Brush, A.J.B.: *Annotating Digital Documents for Asynchronous Collaboration*. Dept. of Computer Science and Engineering, Univ. of Washington. Technical Report 2-09-02 (2002)
18. Open Archives Initiative, <http://www.openarchives.org/> (accessed in March 2009)
19. RDF API, <http://sites.wiwiss.fu-berlin.de/suhl/bizer/rdfapi/> (accessed in March 2009)