# The Role of Ontologies in Semantic Digital Libraries*

Sebastian Ryszard Kruk[1], Bernhard Haslhofer[2], Piotr Piotrowski[1,3], Adam Westerski[1,3], Tomasz Woroniecki[1,3]

[1] DERI, NUI Galway, Ireland* <firstname.lastname@deri.org>
[2] ARC Seibersdorf research, Research Studios – DME, Vienna, Austria <bernhard.haslhofer@researchstudio.at>
[3] Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Gdansk Poland

Providing uniform access to the growing amount of information available in digital libraries is still an open issue. Ontologies can help us to deal with structural and semantic heterogeneities and by maintaining machine-interpretable metadata we can give users a uniform interface for searching over distributed digital assets. In this paper we present and discuss application areas where ontologies have successfully been used in existing semantic digital library systems.

## 1 Introduction

There is a growing need of libraries and cultural institutions to cooperate with each other and to expose cultural artifacts and digital contents to a broader audience using a common digital library system. Since metadata serve as basis for search and discovery services, it is essential to establish uniform access to metadata provided by the various involved institutions and to provide machine-support for the end users in their search for information.

We understand *Semantic Digital Libraries* as digital library systems that apply semantic technologies to achieve these goals. In such systems ontologies play a major role to cope with the problems caused by the structural differences of existing systems and the semantic differences of metadata standards. By analyzing two Semantic Digital Library projects (JeromeDL [1] and BRICKS [2]) we have identified three application areas of ontologies which will be discussed in the following sections.

## 2 Bibliographic Ontologies

Usually a digital library uses a certain metadata format for organizing its bibliographic descriptions. A very popular bibliographic description format, which originates from traditional cataloging card systems, is the MARC family with MARC 21 [3] as an example. Of course there are many more existing metadata standards like Dublin Core [4], BibTeX [5], etc. Unfortunately these standards were defined for human usage and interpretation and do not define the semantics of metadata fields in a machine interpretable way. Compared to traditional standards, semantic-aware metadata formats expressed in terms of an ontology offer the ability to use well-defined types and concepts and to infer implicit data from bibliographic descriptions.

One of the bibliographic ontologies being currently developed and already used in JeromeDL is the MarcOnt Ontology [6, 7]. It aims at combining different metadata standards that can describe various concepts on different levels of granularity. MarcOnt Ontology also takes advantage of other ontologies

[1,3] This work was supported by Science Foundation Ireland Grant No. SFI/02/CE1/I131 and by the Knowledge Web project (FP6 - 507482) and partially by KBN, Poland under grant No. 4T11C00525. The authors would like to acknowledge Stefan Decker, Henryk Krawczyk, John Breslin, the DERI Semantic Web Cluster and the Corrib.org working group for fruitful discussions.
[2] This work was supported by the Austrian Federal Ministry of Economics and Labour and partly funded by the European Comission under BRICKS (IST 507457)

like FOAF [8] to express the *author* concept or the Dewey Decimal Classification [9] for subject indices.

In BRICKS the approach for bibliographic ontologies is different: one of the main design goals was to support arbitrary metadata formats and to enable management of metadata that describe contents in various, domain specific ways. BRICKS relies on OWL-DL [10] for defining the structure and semantics of metadata descriptions and uses the mapping support provided by OWL for crosswalks and mappings between distinct formats [11].

# 3 Ontologies for Content Structures

Modern digital library systems not only store bibliographic metadata but also an electronic representation of the content itself. Depending on its type, content typically follows some structure which we can take into account in our metadata descriptions. For instance, in case of a book stored in a library we can decouple its structure into chapters, provide individual descriptions for each chapter and store information about relations between different chapters.

By including structural concepts in ontologies and using these concepts in metadata descriptions, we can provide a universal layer for metadata and content retrieval. Furthermore, this approach allows us to easily extend the structure description of resources with new concepts, without changing the underlying database schema or violating the integrity of existing data.

As earlier research shows [1], the application of ontologies for structural descriptions enables uniform access to structural and bibliographical information and in this way opens new search and discovery possibilities.

# 4 Community-aware Ontologies

Typically a library is an institution which serves certain user groups or communities. The individual user not only wants to increase her knowledge by using a library but also has some knowledge on certain topics to share. This aspect is often neglected by today's digital library systems because too often they simply represent a digital representation of their "real-world" counterpart. However, in a semantic digital library, besides storing contents and metadata, we can also keep track of its users, their interactions, and incorporate their knowledge into the system. The overall goal of this approach is to share knowledge within groups of users, so that each user can utilize and learn from the experience of other users. User-defined categories in combination with descriptions from a thesauri or controlled vocabularies (e.g. WordNet [12]) provide the necessary means for answering complex but quite common queries, like "What does the person X whom I value as an expert in mathematics consider interesting in subject of fractal art?

The FOAF ontology [13] provides a unified way for describing people. Among properties like *name*, *location* or *interests* FOAF allows to specify friendship relations. FOAFRealm [14] extends FOAF ontology with *level of friendship relation* and offers sharing bookmarks and catalogs among friends, thus provides a base for social semantic collaborative filtering [15].

Using publicly accepted ontologies and RDF [16] as a format for storing data helps us to keep information about users and their knowledge understandable across systems. In that way, we can provide interoperability with other community sites.

# 5 Conclusions

In this paper we have presented three application areas for using ontologies in semantic digital libraries: first we proposed to lift bibliographic metadata to a machine-interpretable semantic level by applying concepts defined in an ontology, second we applied ontologies to model the structural aspects of contents stored in digital library systems, and third we introduced ontologies as a mechanism to describe the knowledge of users and communities so that digital libraries can perform the step from static information to dynamic knowledge spaces. With the development of RDF, OWL, and SKOS [17] the W3C Semantic Web Activity has provided the necessary technologies and currently running projects like JeromeDL, and BRICKS are aiming in this direction.

# Bibliography

[1] Kruk, Sebastian Ryszard and Decker, Stefan and Zieborak, Lech: JeromeDL - Adding Semantic Web Technologies to Digital Libraries; 2005
[2] BRICKS Project: Building Resources for Integrated Cultural Knowledge Services (IST 507457); http://www.brickscommunity.org
[3] Marc21
[4] Oren Patashnik: Bibtexing; 1988
[5] DublinCore Initiative: Dublin Core Metadata Element Set, Version 1.1: Reference Description; 2003
[6] Marcin Synak: MarcOnt Ontology - Semantic MARC21 Description for L2L & L2C Communication; 2005
[7] Kruk, Sebastian Ryszard and Synak, Marcin and Zimmerman, Kerstin: MarcOnt - Integration Ontology for Bibliographic Description Formats; 2005
[8] Dodds, Leigh: An Introduction to FOAF; 2004
[9] Dewey, Melvin: A Classification and Subject Index for Cataloging and Arranging the Books; 2004
[10] OWL Web Ontology Language: http://www.w3.org/TR/owl-guide/
[11] Haslhofer, Bernhard: Metadata Management in a Heterogeneous Digital Library; 2005
[12] WordNet: http://wordnet.princeton.edu/
[13] FOAF Ontology: http://www.foaf-project.org/
[14] Kruk, Sebastian Ryszard: FOAF-Realm - control your friends' access to the resource; 2004
[15] Kruk, Sebastian Ryszard and Decker, Stefan: Semantic Social Collaborative Filtering with FOAFRealm; 2005
[16] RDF Resource Description Framework: http://www.w3.org/RDF/
[17] Simple Knowledge Organisation System (SKOS), http://www.w3.org/2004/02/skos/