Identity Matters: Constructing Social Identities through Ontology-based Metadata

Maria Cristina Pattuelli
University of North Carolina at Chapel Hill
Tel: +1 919 966 3589
Fax: +1 919 962 2697
pattm@email.unc.edu

Lisa R. Norberg
University of North Carolina at Chapel Hill
Tel: +1 919 843 2310
Fax: +1 919 962 2697
lnorberg@email.unc.edu

Abstract:
Students' understanding of history is often shaped by the social identity they bring to the classroom. Presenting history from the perspective of social identity can be an effective way to connect students to the past. Digital libraries are filled with primary sources that could be used to support this type of learning activity. Traditional descriptive metadata assigned to digital objects often fails to capture specific social identity aspects of the content as well as the relationships between objects that are relevant in history teaching and learning. This paper describes the potential of ontology-driven metadata that represent social, familial, spatial, and temporal relationships between learning objects to help students and faculty construct historical meaning.

Keywords:
Ontologies, Learning Objects, Semantic metadata, Educational aspects of Digital Libraries.

1. Introduction
While more and more digital libraries are offering instructional support for K-12 teachers with lesson plans, activities, and learning objects, the use and integration of primary sources into the college or university classroom is still hindered by faculty's ability to find and manage digital resources that support their individual teaching style (1). The importance of understanding the actual practices and information needs of digital library users has been stressed in the literature (2), but is often neglected in the design of learning objects.

2. Background
A collection of learning objects (LOs) is currently under development at the University of North Carolina at Chapel Hill (UNC) University Library to facilitate the use of the Library's digitized primary source materials in the classroom. Prior to the development of the LOs, a series of in-depth interviews was conducted with history faculty from UNC and other North Carolina colleges and universities to guide the design and development of learning objects based on historical primary sources. Faculty stressed

the need to find primary source materials that offer different perspectives on historical events that could be used to compare and contrast diverse points of view. They further revealed their desire to find materials on individuals whose lives students could relate to and who could help put historical events into a more meaningful context. The re-construction of personalized stories from the past is a powerful way to make students to connect to their community as well as to identify themselves in a broader historical context (3). "Young people's perspectives about the social world, like those of historians and teachers, are shaped by their identities as member of families, communities, regions, and nations, as well as by their affiliations with racial, ethnic, religious, and other groups" (4). Social identities exert a strong influence over how students interpret historical events as well as the ways in which they analyze the primary sources that document them.

Based on the feedback from faculty, the Library staff adopted the interpretation of LOs offered by Polsani: "A Learning Object is an independent and self-standing unit of learning content that is predisposed to reuse in multiple instructional contexts" (5). They began creating a series of learning objects that vary in granularity and emphasize flexibility and reusability. The greater challenge was to determine how to annotate the learning objects to facilitate the discovery and reconstruction of the social identities that existed within and between the primary source materials.

3. Limits of descriptive metadata

Search and retrieval based on full-text keyword search as well as on traditional descriptive metadata[1] often fail to identify aspects of the content that would help to reconstruct social identity. While subject metadata facilitate broad topical access to the learning objects, they provide little support for teachers attempting to put materials into context or to identify relevant relationships between objects. For example, a teacher looking for stories related to slavery might search for "slaves" or "slavery" and retrieve a list of excerpts from slave narratives that match their query. They might retrieve a narrative written by Harriet A. Jacobs (http://docsouth.unc.edu/jacobs/title.html) and further in the list of results another by John S. Jacobs (http://docsouth.unc.edu/neh/jjacobs/menu.html). Absent from the metadata, as well as from the full text of Harriet Jacobs' narrative is any clue that she and John Jacobs were brother and sister. Their separation and life experiences as slaves to different masters could provide a thought-provoking contrast for students and faculty to discuss. Conventional approaches to indexing and retrieval offer limited ways of expressing relationships between indexing terms beyond the typical thesaurus-like structure.

4. An ontology-based perspective

This paper describes the potential of an ontology-driven approach to the annotation of historical LOs for enhancing the discovery and retrieval of the objects according to the perspectives faculty have identified as relevant for their teaching.

Ontologies are models of organized knowledge that can help improve the efficiency of information services including search and retrieval. Ontologies formally define the knowledge of a domain and can be deployed as annotation tools that support rich

---

[1] The learning objects are annotated with a Dublin Core application profile that includes IEEE LOM metadata elements.

semantic descriptions of web content. Metadata based on ontologies are given a well-defined and explicit semantics that can be computationally processed for more sophisticated functionalities in information retrieval and knowledge management applications. For example, searches can be performed not only against attribute values, but also against relationships. Some level of automated reasoning can also be performed depending on the capabilities enabled by the knowledge representation technique and the functionality of the ontology management system.

A domain-specific ontology representing the content matter of the collection of learning objects related to North Carolina history is being designed to enhance indexing and retrieving of the objects. The ontology will provide the knowledge framework to semantically ground and relate metadata to be deployed for indexing the objects. Leveraging semantic metadata and their relationships is a way to facilitate discovery and retrieval of the objects based on the social, familial, spatial, and temporal relationships that shape social identity.

For example, by explicitly and formally defining relationships that represent family ties such as `brother-of` and `sister-of`, the ontology facilitates the identification of materials such as the slave narrative of the siblings Harriet and John Jacobs that they may want to compare for instruction purposes.

Within an historical context such as the one provided by the learning objects collection, teachers are most likely to need to retrieve resources according to temporal dimensions. Expressing temporal aspects of content is open to a variety of different modeling choices (e.g., measures of duration, calendar dates, frequencies, and concatenation of temporal intervals) (6).

Faculty suggested that when dealing with historical events, broad sequential relationships, of the type `prior-to/after`, are important and should be included in the ontology model so they can eventually be exploited in the querying process (e.g., What were the conditions of slave life `before` and `after` emancipation.)

Were slaves from northern states treated differently from slaves in southern states? Formal representation of spatial relationships may help to discover objects that answer this question. For example, a geographical `part-of` relationship will allow aggregating all the northern and the southern United States. Leveraging computationally such relationship will facilitate the discovery of content that can be contrasted from different geo-political viewpoints.

Another hypothetical use of the ontology is the very reconstruction of an historical figure's social identity. Take the example of Wilson Caldwell (http://www.lib.unc.edu/ncc/testing/slavery), born into slavery, the son of November Caldwell, slave of North Carolina Governor Joseph Caldwell, and Rosa Burgess, slave of UNC President David Swain. As the property of his mother's slave owner, Wilson was given the name Swain. Following Emancipation he took his father's surname Caldwell who had then become President of UNC. Caldwell's personal identity is the direct result of a socio-political process that the ontology can make explicit and also accessible to reasoning (see Figure 1).
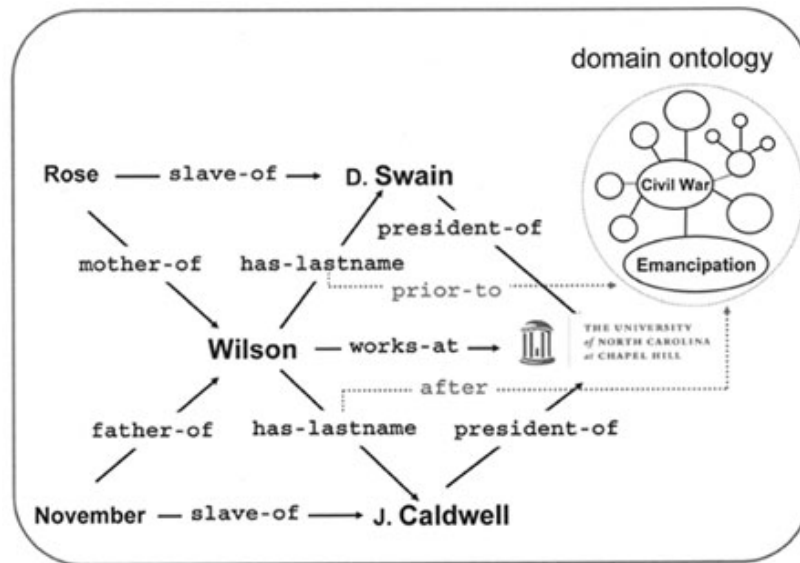
Figure 1. Social identity of Wilson Swain Caldwell.

While ontological relations support the aggregation of objects according to familial and social linkages, additional knowledge may be inferred by exploiting the semantics of the ontology. In a working system, an implementation must be able to make statements and reason on class properties and draw interesting inferences on social aspects of Wilson Swain Caldwell and beyond. For instance, we can deduct that Wilson Swain and Wilson Caldwell are the same person; that slaves do not have family names, but inherit the surnames of their owners; that UNC employed slaves, and that at least two presidents of UNC were slaveholders.

5. Conclusions and future work

In this paper we offer an approach to addressing some of the information seeking needs and teaching objectives that have emerged through focus groups and interviews with higher education teachers of history and related disciplines. We suggest that semantic metadata grounded into a domain-specific ontology could be used to support the description and retrieval of a collection of digital learning objects and to facilitate their use in the classroom.

Based on the feedback gathered from the teachers, a basic core of geospatial and temporal relationships have been identified. They address the topology of space (e.g., `part-of`) and the sequence of events in time (e.g., `prior-to/after`) that would be helpful to support the aggregation and comparison of content. The range of choices in representing geospatial and temporal semantics is wide and will be further explored and assessed in terms of expressivity as well as implementation feasibility as the project advances.

Personal relationships, including familial relationships, are also offered as a way to reconstruct the social identity of people from the past and support educational activities that make connections between students' life and the lives of those who lived centuries before. The next step in this research will be to model these relationships as a part of the development of the North Carolina History Ontology and evaluate implementation

options offered by knowledge representation formalisms and development tools currently available for knowledge-based systems and semantic web applications.

References:

1. D. Harley, et al. *Use and users of digital resources: A focus on undergraduate education in the humanities and social sciences*. Berkley: Center for Studies in Higher Education, 2006.

2. C. L. Borgman et al. Comparing faculty information seeking in teaching and research: Implications for the design of digital libraries. *JASIS&T,* 2005, vol. 56, n. 6, pp. 636-657.

3. G. W. Clarke and J. K. Lee. The promise of digital history in the teaching of local history. *The Clearing House*, 2004, vol. 78 n. 2, pp. 84-87.

4. T. Epstein and J. Shiller. Perspective matters: Social identity and the teaching and learning of national history. *Social Education*, 2005, vol. 69, n. 4, p. 201.

5. P. R. Polsani. Use and abuse of reusable learning objects. *Journal of Digital Information*, 2003, vol. 3, n. 4, article n. 164, 2003-02-19.

6. J. R. Hobbs and F. Pan. An ontology of time for the semantic web. In *ACM Transactions on Asian Language Information Processing (TALIP)*, 2004, vol. 3, n. 1, pp. 66-85.