EASY CREATION OF SEMANTICS-ENHANCED DIGITAL ARTWORK COLLECTIONS

Darina Dicheva, Christo Dichev, Cristan Brown,
Dhanya Veetil, William Thomas

Abstract: In this paper we propose an approach for cost-effective employing of semantic technologies to improve the efficiency of searching and browsing of digital artwork collections. It is based on a semi-automatic creation of a Topic Map-based virtual art gallery portal by using existing Topic Maps tools. Such a 'cheap' solution could enable small art museums or art-related educational programs that lack sufficient funding for software development and publication infrastructure to take advantage of the emerging semantic technologies. The proposed approach has been used for creating the WSSU Diggs Gallery Portal.

Keywords: Cultural Heritage, Semantic Web, Topic Maps, Personalization.

ACM Classification Keywords: H.3 Information Storage and Retrieval - Systems and Software (User profiles and alert services)

Conference: The paper is selected from International Conference "Intelligent Information and Engineering Systems" INFOS 2008, Varna, Bulgaria, June-July 2008

Introduction

In response to the increasing concern for preserving cultural heritage and making it universally accessible, numerous virtual museum projects have been initiated worldwide. Most of these projects are focused on problems related to using content management systems to enter or convert, catalogue, and index digital content. The content is typically indexed with a mix of local and industry-standard metadata, such as Dublin Core, Categories for the Description of Works of Art, etc.

The potential of the emerging Semantic Web technologies to enhance the efficiency of the search within virtual museum collections and make them interoperable has also attracted the attention of researchers and practitioners in the area. Several projects have been recently investigating how the Semantic Web technologies can be deployed to provide better indexing and search support within large virtual museum collections, including the Dutch Continuous Access to Cultural Heritage (CATCH) Program [CATCH url] that includes the Cultural Heritage Information Personalization Project (CHIP) [CHIP url], Charting the Information Landscape Employing Context Information (CHOICE) Project, etc., the Finish Museum-Finland and Culture-Sampo Projects [Hyv¨onen et al, 2005], the Danish Knowledge Management in Museums Project, etc. These large-scale projects aim to employ semantically richer knowledge representation scheme than traditional metadata schemas in order to support enhanced semantic search and browsing, reasoning, and personalized access to the collections. They typically use W3C Semantic Web languages RDF and OWL to semantically annotate the targeted museum collections, and develop ad hoc frameworks requiring considerable programmers' expertise and resources. It is not easy to reuse the created platforms for new museums/digital art collections.

In this paper we propose an approach for cost-effective employing of semantic technologies to improve the use of digital artwork collections. It highly reduces the involved software development being based on a semi-automatic creation of a Topic Map-based art gallery portal (through the use of our existing Topic Maps tools). Such a 'cheap' solution could enable small art museums or art-related educational programs that lack sufficient funding for software development and publication infrastructure to take advantage of the emerging semantic technologies. We have bed tested our proposed approach in creating the WSSU Diggs Gallery Portal.

The paper is organized as follows. We start with a discussion of the advantages of using the Topic Map technology for developing semantics-navigated Web portals, which made the development of the generic Web portal possible. Section 3 describes our approach to creating a semantics-enhanced virtual art gallery portal. In

Section 4 we discuss the search, navigation, and personalization in the Diggs Gallery Portal, and we conclude in Section 5.

Topic Maps-Driven Semantic Web Portals

The ISO Semantic Web standard Topic Maps (TM) [Biezunski et al, 2000] provides a very suitable technology for semantic structuring and annotation of digital artworks collections and can contribute significantly for improving the efficiency of their searching and browsing. The following backs our proposal:

- Topic Maps offer a standards-based solution for knowledge integration, or Seamless Knowledge [Pepper, 2004]. The Topic Maps data model supports naturally the design of information spaces that offer ontologically rich representations of information, based on heterogeneous information sources, in an integrated fashion.
- The Topic Maps model is designed to be understandable for both machines and humans. This allows for easy and quick creation of Topic Map-based end user interfaces. It also allows the use of some free general purpose topic map tools (such as TM4L) for creating, maintaining, and searching TM-based art collections.
- Topic Maps provide a perfect model for web portals; this allows for quick and easy modification, adaptation and reuse of portals.
    - By implementing the very structure of the portal, a topic map actually drives it: navigating the portal is navigating the topic map.
    - Topic maps support a subject-centric content organization that allows very efficient construction of “topic pages” that contain all the available information in the collection about a particular subject.
- The Topic Maps model supports elegantly both semantic navigation and search:
    - The semantic aspect of Topic Maps is covered with associations. Differently from RDF statements, they are not triples (<subject, predicate, object>), but bind together an arbitrary number of topics. Each of the participating topics plays a specific role. Thus associations support naturally semantic navigation in topic maps.
    - Topic Maps-based collections can offer contextual support to users during their search and query, by helping them to express correctly their information needs so as to find quickly needed information. This is especially important when users don’t know what they are looking for, as well as for small collections where a standard keyword-based search will often fail to find matching objects.
- The fundamental “merge” feature of Topic Maps allows for easy and effective merge of existing heterogeneous information resources while maintaining their meaningful structure. The possibility for merging topics can support:
    - Flexible and efficient re-using and extending of existing collections.
    - Collaborative content authoring.
- Topic Maps can support exchangeability, reusability, and interoperability of information through the identity mechanism built in their model. This is a distributed and democratic mechanism for assigning unique, global identifiers (URIs) known as Published Subject Identifiers, to subjects of common interest, thus making it possible to know when topics subjects are the same.
- The TM concept of scope (context within which a characteristic of a topic may be considered to be true) allows supporting naturally different views or perspectives on a collection. A common use of scope is for providing localized names for topics or defining contexts.

In the context of virtual museums, the Topic Maps model can support subject-centric architecture that is able to accommodate not only digitized images and text, but equally their subjects, themes and movements, their

authors, as well as the people and places, objects etc. mentioned or depicted in those images (or texts). Because of the generality of the Topic Maps paradigm, the conceptual structure can be extended as needed, e.g. to include extra classification schemes such as particular categorization of sculpture artifacts, or to provide more specific types of relationships between images. The subject-centric principles, which are in the center of the Topic Maps paradigm, could have a positive impact on art collections organization. From an authoring perspective, the subject-centric Topic Maps model assumes that the author begins with the subject itself (reified by a particular topic) and extends it with associations and occurrences. Unlike the resource-centric view, in which resources are surrounded by metadata, in the subject-centric view subjects are surrounded by data. A topic map aggregates information to provide binding points from which everything known about a given subject can be reached. This approach affects both the collection creation and information retrieval.

The above features of the Topic Maps technology prove its appropriateness for creating efficient topic map-driven Web portals that support intuitive navigation. Such portals create dynamically the frame structures and content of the rendered webpage from an underlying topic map. Taking an advantage of this semantic technology we have created a generic topic map-driven portal along with a tool for automatic generation of specific portals from the generic one. Our generic portal provides the interface, the topic filtering functionality, and the topic map-directed browsing functionality. The latter supports two different patterns of browsing based on the structure of the underlying topic map. By providing a specific description of the desired presentation categories (indexes) and a specific topic map (an XTM file) to the tool, an author may generate a specific portal from the generic model. Separate portal versions may be generated by specifying different categories that should be explicated by the portal (depending on the application, these could be papers, authors, journals, conferences, or as in this case artworks, artists, exhibitions, etc). The specified categories must be resource or topic types in the corresponding topic map.

Our approach to rapid development of semantics-enhanced artworks portals was to develop such a portal by using our previously developed Topic Maps-related tools: (1) TM4L Editor - a general topic map editor [Dicheva & Dichev, 2006], (2) a generic topic map-driven portal, and (3) a tool for automatic generation of specific portals from the generic one.

As a proof of concept we have created a semantic web portal for the WSSU Diggs gallery [Diggs Gallery url].

The Virtual Diggs Gallery

Diggs Gallery is a small art gallery but offers one of the largest exhibition spaces dedicated to the arts of Africa and the African Diaspora in North Carolina. Exhibitions, publications and programs address a broad range of artistic expression, with special concentration on African-American and regional art. With the creation of the virtual Diggs gallery, our goal was from one side to investigate the appropriateness and easiness of applying the TM technology for creating software for cultural heritage projects, based on Topic Map-based semantic structuring and annotation of digital artworks collections, and from another, to increase the accessibility of the Diggs Gallery artifacts to the general public.

According to our approach, the development of the Diggs Gallery Portal involved:

1. Creating an appropriately structured topic map that contains the semantic annotations of the Diggs Gallery artworks by using our TM4L Editor;
2. Automatic creation of a Web Portal as an instance of our generic Topic Map-based Portal that was slightly adapted to meet the requirements of the Diggs Gallery Portal;
3. Changing the design of the main web page of the portal (optional).

With regard to (1), we designed a topic map model of an artwork collection and using this model created a topic map in the TM4L Editor. Our goal was to organize the collection in a way that helps to highlight the meanings and relationships that exist between the artifacts in the collection and across the collection as a whole. This goal provided the strategic factors forming our design requirement of the model. Fig. 1 displays a screenshot of the Diggs topic map, as created in the TM4L Editor.

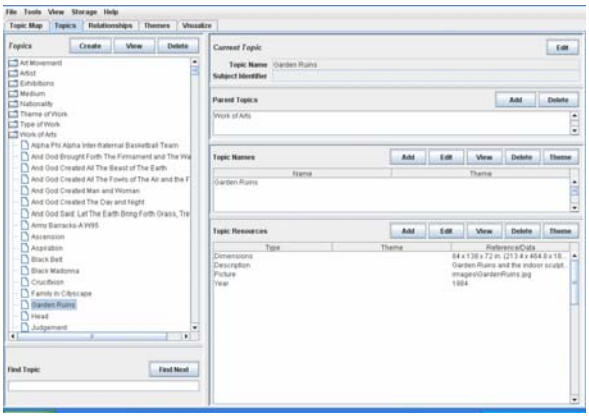

Figure 1. Screenshot from the TM4L Editor used for creating the Diggs topic map.

The topic map model contains the following topics and relationships:

- Topics
  - Art Works
    - Topic name: title of the art work
    - Occurrence types: Year, Dimensions (in inches and centimeters), Narratives / Description, Inscription, Gift from
  - Type of work (e.g. Public Art, Painting, Drawing, Print)
  - Medium (e.g. Acrylic on canvas, Oil on canvas, Granite, Steel, Stone, etc)
  - Theme of Work (e.g. Animals, Still Life, Portrait, Family, Religious, etc.)
  - Collection
  - Exhibition (will be used for constructing the Exhibition History)
  - Art movement (e.g. Renaissance, Impressionism, Graffiti Art, etc.)
  - Artist
    - Topic name: Artist's full name
    - Short (variant name)
    - Occurrence types: Year of Birth, Year of Death, Biography, References (for the biography), Nationality
- Relationships
  - Artist-Work (with roles Artist and Artwork)
  - Type-Work (with roles Type of artwork and Artwork)
  - Medium-Work (with roles Medium and Artwork)
  - Collection-Work (with roles Collection and Artwork)
  - Exhibition-Work (with roles Exhibition and Artwork)
  - ArtMovement-Work (with roles Art movement and Artwork)
  - Theme-Work (with roles Theme of artwork and Artwork)
  - Artist-Nationality (with roles Artist and Nationality)

With regards to (2), we have used our program for automatic generation of a Web portal from our generic Topic Map-based portal. To match the particular art gallery portal requirements this process was followed by minor adaptation involving replacement of the names of the artworks upon rendering with the actual image (as shown in Fig. 2), or with an icon representing the artwork (as shown in Fig. 5).

Figure 2. Diggs Gallery Portal: An artwork view.

Figure 3. Editing functionality in the Diggs Gallery Portal.

Figure 4. Searching Diggs Gallery Portal.

Figure 5. Diggs Gallery Portal: A personalized entry page for a registered user.

Although the generated Diggs Gallery Portal was fully functional and usable, we decided to extend further its functionality so that authorized end users have a publication channel , i.e., can add new artworks or edit the existing information directly from the portal, without a need to download the topic map, edit it in TM4L and upload it again on the site. Thus we implemented an extension to the portal allowing direct authoring of the underlying topic map. The editing menu is displayed only to registered users that are appropriately authorized. Figure 3 displays a screenshot from the Diggs Gallery interface showing the editing menu.

Search, Exploration, and Personalization in the Diggs Gallery Portal

As we already mentioned, the Topic Maps model supports elegantly both semantic navigation and search. Dubbed "the GPS of the information universe", topic maps are destined to provide powerful new ways of navigating large and interconnected corpora – their connectivity can support intuitive navigation and topical search. This is the feature that made possible the proposed topic map-driven generic portal, since it ensured seamless navigation using the topic map connectivity without the need of defining specialized domain-dependent queries to the portal back-end. The resulting portal provides versatile facilities for searching, browsing, and navigating around the art collection. Semantic search based on ontological concepts and semantic browsing allow dynamic presentation of semantic associations between search objects to the end-user as recommendation links with explicit explanations. Thus a portal visitor can navigate and explore the collection in a way dictated by their interests. For example, when seeing an artwork, if the visitor clicks on the Art Movement value for this work, say it is Harlem Renaissance, they will be presented with all artworks of the Harlem Renaissance movement. If they click on the Artist, the artist page will be displayed, etc.

In addition to the topic map-based navigation, for accessing the content of the collection we currently support the two dominant interface types for searching and browsing arts collections: keyword based search, and searching based on grouping and semantic relationships to the other objects in the collection. We have implemented multi-faceted search over type of work, medium, and theme of work. Figure 4 displays a screen shot from the Diggs Gallery Portal illustrating this functionality.

In line with the best practices in the field, we have also extended the Diggs Gallery Portal with personalization functionality. Personalization targets improving museum websites usability by supporting user navigation and assisting users in finding information of interest. To this end, virtual museums automatically adapt the content presentation using user data stored in a user profile. This data typically include user interests, however, personal characteristics, such as age, gender, education, previous knowledge, etc. could support even better comprehension of the collection. Studies have shown that understanding is stimulated when the system uses concepts familiar to the users (considering their interests and knowledge level. User data is typically collected in two ways: (1) users fill in online forms to provide data, and (2) the system monitors users' activities to infer and record their preferences. In addition, [Rutledge et al, 2006] propose an interactive approach for determining user interests in a museum collection with the help of a dialog, which uses artefacts from the semantically annotated collection of the Rijksmuseum Amsterdam to elicit specific user's interests in artists, periods, genres and themes and uses these values to recommend relevant artefacts and related concepts from the museum collection.

Similarly to the leading large-scale projects in the area, the visitors of the Diggs Gallery Portal get personalized view on the collection depending of their user profiles. We have employed similar approaches for collecting user data. Upon registration a new user fills in a registration form, where they can provide some personal data (if they wish so). Our methods are non-intrusive - the system further builds users' profiles by tracking user behavior and mining their logs. As the visitors review art works in the portal (after logging into the system), their activities are tracked and the tracking data are stored in the system's database. Tracking allows finding out how much time a user has spent on reviewing a specific art work and how often he/she has visited the same art work. In addition, we are currently implementing an artwork rating feature that allows visitors to rate their interest in specific facets of artworks, which enhance system knowledge about the user. The user data is used for generating a personalized view of the art work collection for the user (or recommendation of artifacts from the collection). Once

a visitor logs into the system, it analyzes user's data and a personalized list of art works is generated to be displayed to that particular user (see Fig. 5).

The envisaged personalization will improve the Diggs Gallery website usability by supporting user navigation and assisting users in finding appropriate and interesting information.

Conclusion

This paper sketches our experience involving the efforts to improve user access and navigation to a (small-to-medium size) arts collections, by providing flexible exploration model and by personalizing the user view on the collection. The reported work has a twofold objective: On one hand, to investigate how the technologies and tools developed previously from us can be further expanded and repurposed to support semantically enhanced digital art work collection, and on the other hand, to explore the potentials of Topic Maps-driven digital art collections compared to other technologies in terms of low cost systems. As a conventional small art gallery the "proof of concept" repository is composed of a logically related collection of digital objects that can be used to provide various points of information access, giving the visitors a multi-way and multi-purpose navigation path through the collection. The presented experience confirms that Topic Maps provide a suitable model for building portals and other forms of Web-based information delivery.

Future work includes enhancements to content annotation, concept-based organization, and user interface, and use of the architecture as a platform for more powerful personalization and accelerated multipurpose classification.

Acknowledgement

This work is partly supported by the WSSU RIP 2007 Project "Semantic Annotation of Digital Cultural Heritage Collections and LSAMP.

Bibliography

[CATCH url] Continuous Access to Cultural Heritage Program, http://nwo.nl/nwohome.nsf/pages/NWOP

[CHIP url] Cultural Heritage Information Personalization Project (CHIP) http://www.chip-project.org/demo/

[Hyv¨onen et al, 2005].Hyv¨onen, E., M¨akela, E., Salminen, M., Valo, A., Viljanen, K., Saarela, S., Junnila, M, Kettula, S.: MuseumFinland – Finnish Museums on the Semantic Web. J. Web Semantics, 3(2): 224–241 (2005)

[Dicheva & Dichev, 2006] Dicheva, D., Dichev, C.: TM4L: Creating and Browsing Educational Topic Maps, British Journal of Educational Technology – BJET, 37(3): 391-404 (2006)

[Biezunski et al, 2000] Biezunski, M., Bryan, M., Newcomb, S.: ISO/IEC 13250:2000 Topic Maps: Information Technology, www.y12.doe.gov/sgml/sc34/document/0129.pdf

[Pepper, 2004] Pepper, S.: Seamless Knowledge: Spontaneous Knowledge Federation using Topic Maps, In: Proceedings of Extreme Markup Languages, Montréal, Québec (2004)

[Diggs Gallery url] WSSU DIGS Gallery, http://www.wssu.edu/WSSU/About/Administration/Information+Resources/Diggs+Gallery

Authors' Information

Darina Dicheva – Paul Fulton/Delta Sigma Theta Distinguished Professor of Computer Science, Winston-Salem State University, 601 Martin Luther King, Jr. Drive, Winston Salem, N.C. 27110; e-mail: dichevad@wssu.edu

Christo Dichev – Professor of Computer Science, Winston-Salem State University, 601 Martin Luther King, Jr. Drive, Winston Salem, N.C. 27110; e-mail: dichevc@wssu.edu

Cristan Brown, Dhanya Veetil, William Thomas – Students, Winston-Salem State University, 601 Martin Luther King, Jr. Drive, Winston Salem, N.C. 27110