Semantics-driven Recommendations in Cross-Media Museum Applications

Natalia Stash[1], Lora Aroyo[2], Yiwen Wang[1], Lloyd Rutledge[3] and Peter Gorgels[4]

[1] Technische Universiteit Eindhoven
{n.v.stash, y.wang}@tue.nl
[2] Vrij Universiteit Amsterdam
l.m.aroyo@cs.vu.nl
[3] Open Universiteit Nederland
Lloyd.Rutledge@ou.nl
[4] Rijksmuseum Amsterdam, The Netherlands
P.Gorgels@rijksmuseum.nl

Abstract. In this paper we present the CHIP demonstrator aimed at helping users to explore the Rijksmuseum Amsterdam collection both online and inside the museum. Cultural heritage data from various external sources is integrated to provide an enriched semantic knowledge structure. The resulting RDF/OWL graph is the basis for CHIP main functionality for recommendations, search and personalized interaction.

1 Introduction

Different types of people visit museums — school children, tourists, art experts, students — and all of them have different goals and interests in art and would need a different route and a story inside the museum. The main focus of the CHIP project is to provide a personalized experience to the museum visitors on the museum Web site as well as in the museum. We built several CHIP demonstrators that (a) allow users to find out what they like and dislike in the museum collection, (b) allow being their own curators by selecting artworks and topics they want to see in a museum tour, and (c) show how mobile tour guide can help guiding users inside the museum [1].

The main question that we focus on in CHIP is whether we can use semantic metadata of cultural heritage to improve personalized access through multiple devices, such as PCs, PDAs or mobile phones. The personalized access and navigation that we provide is based on a semantically enriched data from the digital database ARIA[5] of the Rijksmuseum Amsterdam. The online CHIP demonstrators [2] and a tutorial with a brief walk-through of the personalization functionality can be found at http://www.chip-project.org/demo.

[5] http://rijksmuseum.nl/aria/

2 The CHIP Data Approach

The main problem with museum thesauri is often the fact that they have a rather flat and non-hierarchical structure. For example, there are missing intuitive relationships between topics like "mythology" and "Zeus". In order to be able to explore interesting relationships between artworks we often need more semantic knowledge. Thus, the use of semantics is an important instrument in our approach to enhance the interaction with museum collections. We use external cultural heritage structures to enrich the existing Rijksmuseum collection metadata. For the CHIP demonstrators we applied a general strategy, as explained in the steps below, that supports such class of semantics-based applications in cultural heritage domain.

1. **Making museum metadata available in RDF/OWL**.
   In collaboration with the MultimediaN E-Culture project[6], we made the relevant metadata of the website-targeted digital ARIA database (a kiosk system for the museum) available in RDF/OWL. This collection contains 729 artworks, 486 themes, 690 encyclopedia keywords, 43 catalogue terms, and finally resulted in 47.329 triples. We also converted the curator-targeted AdLib database, which contains 16.156 artworks and resulted in 400.405 triples. The current CHIP demonstrator uses only the ARIA conversion, while preparations are made to include also the AdLib RDF/OWL.
2. **Making relevant external vocabularies available in RDF/OWL**.
   Next important step was to select the relevant external cultural heritage sources and convert them in RDF/OWL. In the CHIP demonstrator we use the RDF/OWL conversion of the three Getty thesauri[7] as provided by the E-Culture project, i.e. the list of Geographical Names (TGN, 425.517 triples), the Unified List of Artists Names (ULAN, 1.896.936 triples) and the Art and Architecture Thesaurus (AAT, 1.249.162 triples). Next to this we use the RDF/OWL conversion of IconClass[8], containing 24.349 triples, as provided by the STITCH[9] project.
3. **Aligning and enriching vocabularies/metadata**.
   In this step we did the following manual alignments of the ARIA vocabulary to the Getty and IconClass concepts: 2.825 mappings of subject themes and artists styles to 283 different AAT concepts, 485 mappings of artists names to 263 ULAN concepts, 507 mappings of concepts for places to 69 TGN concepts, 503 mappings of subject themes to 178 IconClass concepts. In order to introduce the *"style"* concept in the Rijksmuseum collection (and thus allow users to search for artworks in a particular style, e.g. *"Baroque"*) we mapped the Getty `aat:style` metadata to the artists metadata in ARIA collection. Further we mapped the Rijksmuseum "Location/Period" metadata to `aat:period` and `tgn:location`.

[6] http://e-culture.multimedian.nl/
[7] http://www.getty.edu/research/conducting_research/vocabularies/
[8] http://www.iconclass.nl/libertas/ic?style=index.xsl
[9] http://www.cs.vu.nl/STITCH/

We also did a manual alignment of the AdLib vocabulary to the Getty thesauri, resulting in mappings to 534 AAT concepts and 3.846 ULAN concepts.

4. **Using resulting RDF/OWL graph for building a combined (virtual and physical) user model**.
   While interacting with the CHIP demonstrator users can give ratings to artworks and topics using 5-star-rating-scale, e.g. *I hate it*, *I do not like it*, *It is OK*, *I like it* and *I like it very much*. Ratings are stored in the user model as *artwork or topic URI/rating value* pairs representing an overlay of the CHIP enriched data model [3]. In order to solve a cold start problem we have done first attempts in importing users' data in CHIP from iCITY[10].
5. **Using resulting RDF/OWL graph for recommendations, (semi)automatic generation of museum tours and search.**
   The ratings stored in the user model are used for generating recommendations of Rijksmuseum topics and artworks. A topic (e.g., "Rembrandt", "Portraiture") can be recommended to the user if (s)he rates positively artworks that have this topic or if (s)he rates positively semantically related topics (e.g., if the user rates positively "Birds" then a parent topic "Animals" is recommended). CHIP demonstrator finds artworks related to positively rated or highly recommended topics and recommends them to the user. Currently two tours are generated automatically for the user based on his/her user model — *Tour of Rijksmuseum Favorites* containing artworks positively rated by the user and *Tour of Recommended Artworks* containing 20 top recommended artworks. User can also create tours by using the semantic-search option integrated via the E-Culture open API. In this way the user is able to search not only for artworks and topics exactly matching the search term, but also for semantically related artworks in the Rijksmuseum collection. For example, if the user searches for "Rembrandt" (s)he gets not only artworks created by Rembrandt himself but also artworks created by Rembrandt's teachers and students, as well as artworks by artists in the same style as Rembrandt, etc..

3 CHIP Demonstrator Architecture

CHIP demo is based on a client-server architecture. Server side contains the following information:

- **Collection data** — the enriched Rijksmuseum collection maintained by Sesame Open RDF memory store and queried with SeRQL.
- **Users data** — RDF user models which are later transformed in XML to be downloaded to the mobile device and XML tours data.
- **Demo components**. The *Art Recommender* provides a user-rating-driven interactive interface that helps the user discover his/her interests in the Rijksmuseum collection by recommending relevant topics and artworks from the collection. The user can rate each recommended item as an explicit

[10] http://icity.di.unito.it/dsa/

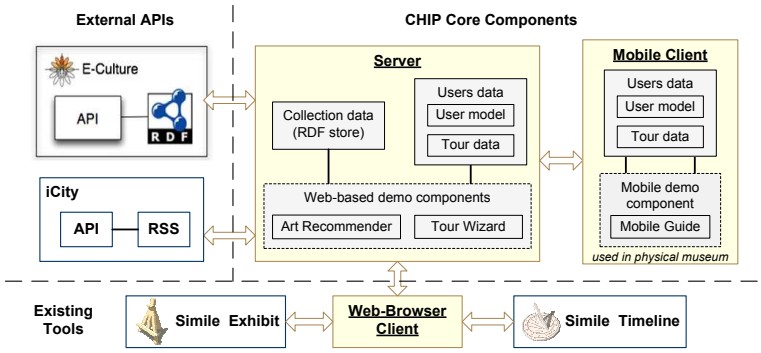

Fig. 1. CHIP Overall Architecture

feedback. The *Tour Wizard* helps the user build personalized museum tours and visualize them on a museum map and a historical timeline. The semantic search option is available in both tools. Art Recommender and Tour Wizard are realized as Java Sevlets and JSP pages with CSS and JavaScript. AJAX is involved for enhancing user experience. Simile Exhibit and Simile Timeline[11] javascript web applications are applied in the demo components for data presentation.

CHIP PDA client with MS Windows Mobile as operating system contains a standalone application *Mobile Guide*. It is RFID reader enabled and can work offline inside the museum and subsequently be synchronized with the server-side on demand. Windows Mobile application is written using the .NET framework in C#. openNetCF and fmodCE libraries are used to support audio features. User model and tour data are downloaded from the server and are being used during the tour. User model is updated when the user provides artworks ratings on a PDA. After the tour this user model can be synchronized with the one on the server.

During the demonstration we will present both online components Art Recommender and Tour Wizard and the mobile component Mobile Guide.

References

1. L. Aroyo, R. Brussee, P. Gorgels, L. Rutledge, N. Stash, and Y. Wang. Personalized Museum Experience: The Rijksmuseum Use Case. In *Museums and the Web Conference*, 2007.
2. L. Aroyo, N. Stash, Y. Wang, and P. Gorgels. CHIP Demonstrator: Semantics-driven Recommendations and Museum Tour Generation. In *ISWC Conference*, 2007.
3. Y. Wang, L. Aroyo, N. Stash, and L. Rutledge. Interactive User Modeling for Personalized Access to Museum Collections: The Rijksmuseum Case Study. In *UM Conference*, pages 385–389, 2007.

[11] http://simile.mit.edu/