SMARTMUSEUM Knowledge Exchange Platform for Cross-European Cultural Content Integration and Mobile Publication

Tomi Kauppinen[1] (tomi.kauppinen@tkk.fi), Tuukka Ruotsalo[1] (tuukka.ruotsalo@tkk.fi)
Frédéric Weis[2] (frederic.weis@irisa.fr), Sylvain Roche[2] (sylvain.roche@irisa.fr),
Marco Berni[3] (marco@imss.fi.it), Eetu Mäkelä[1](eetu.makela@tkk.fi),
Nima Dokoohaki[4] (nimad@kth.se), Eero Hyvönen[1] (eero.hyvonen@tkk.fi)
Semantic Computing Research Group (SeCo)[1]
Laboratory of Media Technology
Helsinki University of Technology (TKK), Finland
National Institute for Research in Computer Science and Control INRIA[2]
Institute and Museum of the History of Science (IMSS), Florence, Italy[3]
Royal Institute of Technology in Stockholm (KTH)[4]

Abstract

European museums and other cultural institutions host rich collections that have ability to attract EU citizens and tourists. Cultural objects, e.g. paintings, in these collections are related in many ways and in many cases they refer to same underlying concepts, people and places. The Cultural Heritage Knowledge Exchange Platform, SMARTMUSEUM requires that these collections are interoperable over cultural and language barriers, and provides a mobile publication channel for collections.
Keywords: semantics, mobile phones, tourism, personalization

1. Introduction

It has been argued that museums should publish their activities, collections, services, and products in cooperation with cultural tourism agencies (Mulrenin, 2002). The Cultural Heritage Knowledge Exchange Platform, SMARTMUSEUM, is a platform for innovative services enhancing on-site personalised access to digital cultural heritage through adaptive and privacy preserving user profiling. Using knowledge bases, global digital libraries and visitors' experiential knowledge, the platform makes possible the creation of innovative multilingual services for increasing interaction between visitors and cultural heritage objects in a future smart museum environment, taking full benefit of digitized cultural information. In this paper we present components needed to realize a knowledge exchange platform for SMARTMUSEUM needs. The annotation framework enables managing of semantic annotations of museum collections. The SmartMuseum recommendation web services use relations defined by ontologies together with user profile and contextual information to provide search and recommendation for a user of SMARTMUSEUM. These web services can be used via mobile and web interfaces.

Proceedings: *Cultural Heritage on line. Empowering users: an active role for user communities*

2. The Content Architecture of SMARTMUSEUM

The vast majority of museums hold a legacy database with cataloguing cards for owned objects. In IMSS case (Institute and Museum of the History of Science[1]) the database contains also some ontological data. The ending result needed from the SMARTMUSEUM application is a set of RDF[2] triples complying to an ontological schema. A clear vision of the target ontology is strictly mandatory especially in a commercial system. Since content providers can have different items and different objectives, small adjustments to the target ontology can be requested for a single implementation, but in general there should be global target ontologies available that all organizations may use.

Next, we will analyze some representative cases about how to map existing metadata to the SMARTMUSEUM ontology. Since not all the required concepts and instances were available in existing ontologies, an integration and an upgrade methodology had to be established to minimize the manual intervention of mapping. In the IMSS case the preexistent database holds information and links among information. For this reason IMSS added SMARTMUSEUM specific ontological annotations to the existing database. This implied a semi-automatic mapping and a manual addition of external references to Getty vocabularies [3]. In the second phase IMSS realized an automatic extractor that used XML-based configuration about what and how to extract from the dataset. The outcome of this process is directly usable by the SMARTMUSEUM system.

3. SMARTMUSEUM Recommendation Service

3.1 Profile retrieval

The first phase in the recommendation is to retrieve a profile that matches the user's current context. Retrieving is done by mapping the user's location, determined by GPS to ontological concepts. This is done by expanding the query of the single coordinate point to cover a circular area within some radius *r*, say, 1000 meters. This is done in two steps. First, a simple bounding box is created where each of the edges of the box have a distance *r* from the user's location, i.e. the distance to the "sides" of the bounding box from the given location is *r*. In the second phase, places outside the radius *r* are further pruned away from the results. This results into an ontological resource representing the position information.
We use the likelihood of a context generating a certain triple. It can be observed from the relative frequencies of the profile entries. For example, if a user profile contains tags for triples about

[1] http://www.imss.fi.it/
[2] Resource Description Framework, http:1/www.w3.org/RDF/

Italian paintings in the context of Helsinki, say 10 times, and triples in Helsinki in total 20 times, the P (<sm:Painting, sm:manufacturedIn, place:Italy> | <rdf:Resource, sm:userLocation, place:Helsinki>) would be 10/20 = 0.5. Because we have the negative or positive votes for the triple we calculate the average of the votes of the triple in the given context and multiply it with the probability of the triple in the context. The contexts in which the observations are done can be very sparse. Therefore, we use Laplace (i.e. add one) smoothing [4] to shave a share of the probability mass to contexts for which no observations are available. In this way, we can observe some probability for every triple even if it has not been tagged in the specified context.

3.2 Recommendation retrieval

Recommendation retrieval is performed by using the query constructed from user profile and context [2]. Based on the earlier phases we have a set of profile triples that each has weight. Each triple may be expanded using query expansion to multiple triples, that each has the weight of the original triples. This is done by including all triple combinations having a Wu-Palmer value higher than a fixed constant. We have set this value to 0.85 by error and trial. As a result we have a set of triples each having a weight.

We can now define the retrieval as a two step matching procedure that utilizes the spatial constraints and a scoring function used to calculate the cosine similarity [6] in vector space model [5], where vectors are formed by using a triples times documents matrix. Further, we cluster the best 300 objects using independent component analysis (ICA). This makes it possible to reveal different viewpoints to the data and avoid over-specialization. For example, if the user has a very strong interest in Italian paintings and a light interest to scientific instruments and telescopes, a traditional retrieval system would only rank Italian paintings high. By clustering, we are able to build three interest clusters, one for Italian paintings, one for telescopes and one for other scientific instruments. Finally, items from each cluster can be included into the final recommendation list.

4. Mobile access to SMARTMUSEUM

Mobile access to SMARTMUSEUM is based on two scenarios: the inside scenario - user visits a museum and the outside scenario - user walks around the city, looking for (outdoor) points of interest (POI). In both scenarios the user is equipped with a PDA or a smartphone as a main device for user's positioning and for presenting recommendations and multimedia (A/V, Text To Speech) information. SMARTMUSEUM mobile access is based on five major features:

1) User is requested to enter the expected visit duration and a purpose of visit from a predefined list.
2) User can acquire contents describing recommended objects by clicking on URLs displayed by the mobile device. According to user's profile preferences, multimedia files or text-to-speech contents are automatically launched when they are available.
3) Each object has a unique URI, a URL is usually stored into a RFID tag for indoor scenario, or GPS coordinates are used for outdoor POIs. When looking at an interesting

indoor object, user scans the RFID tag attached to the object to get information on the mobile device. Each user browsing action is logged.

4) User has an opportunity to rank each page (physical objects and content pages).
5) After the end of visit ranking and log information is automatically sent into the SMARTMUSEUM profile server.

4.1 Mobile user positioning and interest monitoring

Mobile user monitoring is performed (1) for positioning and (2) for discovering user's interest in order to process statistical recommendation. Outdoor user positioning in SMARTMUSEUM context is used for determining location for nearby POI search and for determining location context for semantic recommendations. The SMARTMUSEUM solution supports GPS (WGS84 coordinates) and mobile network cell based outdoor localization. For indoor scenario the objects are equipped with 13,56MHz (ISO144443A) RFID tags used for content triggering and user positioning as well.

4.1.1 RFID reading

A unique solution was developed for RFID tag access. Existing solutions are mainly based on unique hardware ID of the tag. Using this ID for referencing an object would require changes in the SMARTMUSEUM database when tags are replaced. For this reason we made tag data area usable for content. Two solutions have been studied and evaluated: 1) store html contents directly on tags, 2) store an object URI on tag and to retrieve information through wireless network. We made several measurements with 4K tags. Our experiments showed that retrieval of a small html page (3071 octets, without image and CSS style) requires 5 sec. The reading of a simple URI takes only 150 msec. Here the problem is that the user must keep the device near the tag until the end of reading. With the first approach, there is a risk of retrieving partial information. In the second solution, tag reading is almost immediate so that user behaviour is not constrained. Information are automatically loaded and displayed after tag reading. We made some performance measurements with the server of IMSS. It takes about 1.5 sec to read tag and to load a simple html page (without image of object). Finally the second solution has been fully integrated in the current release of our software.

4.1.2 Monitoring user interest feedback

Receiving pertinent user feedback is a crucial factor to process further statistical recommendations for both objects and content pages. Browser activity logging is a widely used method for web content relevance evaluation, especially from server side. In many studies page access duration is taken into account. However, time counting is quite unsuitable for mobile users, because the content access is highly fragmented. So it is difficult to detect idle periods when user, for example, is moving from one object to another one. The second issue is that since content is fully distributed, the activity monitoring cannot be performed in a centralized manner. But in SMARTMUSEUM context, most of the available information is split into separate html pages. In this way it becomes possible to infer user's interest for each museum object and atomic content piece. Mobile software stores all visited URL, and/or visited object URI. And a score is associated to each URIs and URLs.
After the analysis of user scenarios, a combined scale of manual preference input and implicit monitoring of preference and behaviour have been proposed and implemented (see Figure 1):

- score = -0.01: user receives recommended object. If the user shows no interest, and leaves associated recommended link unused, this initial score remains unchanged.
- score = score + 0.3: user fetches basic information about the recommended object or user reads object tag.
- score = score + 0.4: user requests more information about an object to receive a list of additional content URLs.
- score = 1: strong like, manual input on user interface.
- score = -1: strong dislike, manual input on user interface.

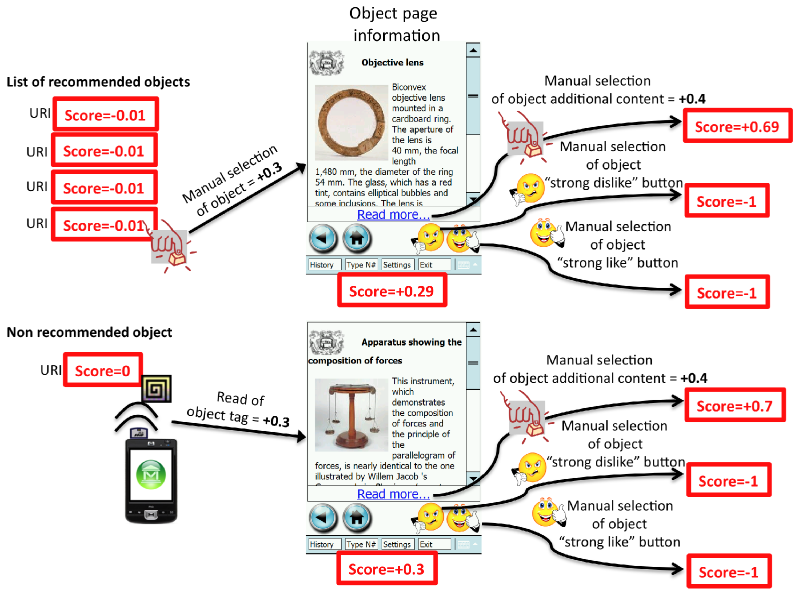

Figure 1. Rating of a SmartMuseum object

Currently the user device client software is implemented for Windows Mobile (WiMo) and Symbian operating system platforms. Two screenshots of WiMo user interface optimized for larger touchscreens are presented in Figure 2 (recommendation list and object information page). Our objective is to minimize user interventions. Upper part of screen is an embedded Web browser window, on bottom part rating buttons can be used. On this screen area multimedia and text-to-speech controls appear automatically when content html page includes respective hidden tags.

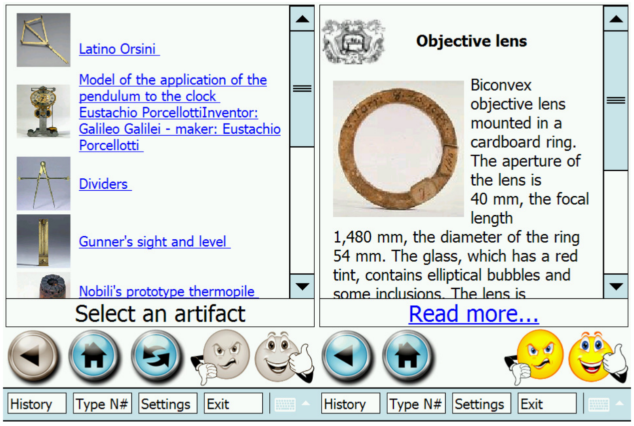

Figure 2: Mobile device main UI (WiMo)

5. Conclusions

SMARTMUSEUM is a versatile knowledge exchange platform for hosting and publishing museum collections and POIs for user's of mobile phones and PDAs. In this paper we described different components used for realizing the SMARTMUSEUM.

6. Acknowledgements

The research described in this chapter was done in the EU project SMARTMUSEUM supported within the IST priority of the Seventh Framework Programme for Research and Technological Development. Partners of the SMARTMUSEUM include Apprise (Estonia), National Institute for Research in Computer Science and Control INRIA (France), Helsinki University of Technology (Finland), Royal Institute of Technology in Stockholm (Sweden), Webgate JSC (Bulgaria), Heritage Malta (Malta), Institute and Museum of the History of Science (Italy) and Competence Centre of Electronics-, Info and Communication Technologies ELIKO (Estonia).

7. References

[1] Mulrenin, A. (Ed.). (2002). *The digicult report. technological landscapes for tomorrow's cultural economy. unlocking the value of cultural heritage. executive summary*. Luxembourg: European Commission.

[2] Tuukka Ruotsalo, Eetu Mäkelä, Tomi Kauppinen, Eero Hyvönen, Krister Haav, Ville Rantala, Matias Frosterus, Nima Dokoohaki and Mihhail Matskin: Smartmuseum: Personalized Context-aware Access to Digital Cultural Heritage. Proceedings of the International Conferences on Digital Libraries and the Semantic Web 2009 (ICSD2009), September, 2009. Trento, Italy.

[3] http://www.getty.edu/research/conducting_research/vocabularies/

[4] Christopher D. Manning and Hinrich Schuetze. Foundations of Statistical Natural Language

Processing. The MIT Press, 1 edition, June 1999.
[5] G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. Commun. ACM, 18(11):613–620, 1975.
[6] Gerard Salton and Christopher Buckley. Term-weighting approaches in automatic text retrieval. Inf. Process. Manage., 24(5):513–523, 1988.