Personalized Context-aware recommendations in SMARTMUSEUM: Combining Semantics with Statistics

Innar Liiv
Tallinn Uni. of Technology and ELIKO
Tallinn, Estonia
innar.liiv@ttu.ee

Tanel Tammet
Tallinn Uni. of Technology and ELIKO
Tallinn, Estonia
tammet@staff.ttu.ee

Tuukka Ruotsalo
SeCo, Helsinki University of Technology
Helsinki, Finland,
tuukka.ruotsalo@tkk.fi

Alar Kuusik
ELIKO Technology Competence Centre
Tallinn, Estonia
alar.kuusik@eliko.ee

***Abstract* - Our goal is to enhance on-site personalized access and recommendations for cultural heritage. We have designed and implemented the SMARTMUSEUM platform using adaptive and privacy preserving user profiling. The described recommendation system relies on combining a semantics/ontologies based approach with a data mining/statistics based approach. The paper presents the architecture and main methods of the system.**

Keywords: recommendations, user profiles, semantics, ontologies, data mining.

I. INTRODUCTION

One of the main trends people are daily confronted with is an ever growing load and diversity of information and the associated complexity to find specific objects and content based on their personal interest. The SMARTMUSEUM FP7 [23] project aims to improve personalized on-site cultural heritage access with a strong focus on user preferences and context based recommendations. The expected applicability of the recommendation solution for several museums relies on the use of widely accepted cultural domain ontologies validated by museums in Italy, Malta and Finland. Additional statistics based methods should enable personalized access to unannotated data.

Recommendation systems targeted to travelers and museum visitors have become perhaps most popular recommendation systems after customer relationships management and news filtering solutions. They have a history over ten years [2][3][4].

Some authors classify recommendation systems into three categories [6]: *content-based recommendations* - the user will be recommended items similar to the ones preferred in the past; *collaborative recommendations* - the user will be recommended items that people with similar tastes liked; *hybrid approaches* combining both methods. The hybrid approach has significant advantages over single method recommendation systems: content based solutions require implicit user preference and content metainformation is usually difficult to acquire, collaborative approaches are facing “cold start” problems.

Alternative classification of recommendation systems describes rule-based, collaborative filtering and content-based personalization systems [9][10]. The particular SMARTMUSEUM recommendation solution can be more accurately described through this scheme.

The collaborative recommendation systems require implicit feedback on whether the user likes the object/content or not. Browser activity logging is a widely used method for web content relevance evaluation. A description of a recent implementation taking into account web page access duration can be found in [11]. It is even more difficult to monitor user impressions regarding physical artifacts. For such purposes even eye movement monitoring has been used [12]. A list of currently more practical monitoring sensors is presented in [19]. 2D barcode and especially short range (13.56MHz) RFID tags accessible with portable readers are frequently used to monitor user attention as a handy and cost efficient solution, e.g. [13], the Stockholm Post Museum exhibition.

Well-standardized cultural domain and generic WordNet ontologies allow simple creation of rather flexible recommendation solutions. Therefore several practical museum recommendation systems, e.g., Amsterdam Rijksmuseum [15], MuseumFinland [7] and iJADE FreeWalker [8] are ontology-based. The additional advantage of this approach is that it removes the need for technically complex user feedback monitoring for simple and static content driven solutions. However, large amount of cultural heritage sites and related content remains unannotated, thus forcing the need for hybrid solutions. Since content based recommendation systems with learning capabilities require user feedback, the hybrid approach appears to combine the best of two worlds.

The paper presents the user scenarios, recommendation system architecture, user profile handling, ontology usage and data mining approaches.

II. SCENARIOS

We use two principal scenarios as a motivation for our architecture: the inside scenario - user visits a museum - and the outside scenario: - user walks around the city, looking for interesting places/buildings to visit/look at.

In both scenarios the user has a PDA as a main device for both locating user and presenting information. The outside scenario relies on GPS for recognizing objects and the inside scenario relies on RFID tags attached near the objects. Both scenarios have to:

- Locate the user. We use GPS outside and RFID reading or manual numeric input inside.
- Send the user profile from PDA to the recommendations server and calculate the probable interests of the user.
- Present the suggested places to visit.
- Present detailed information at interesting places on users request.
- Store user interests/feedback in the profile.
- Allow the user to modify the profile by hand.
- Offer administrative tools both for museum and city places-of-interest administrators.

Let us consider the typical activities of a hypothetical John visiting a museum.

First, John indicates which museum he is visiting, either by reading the museum RFID tag with a museum URI at the entrance or alternatively typing in the museum URI. The . PDA sends his user profile to the profile server along with museum URL.

The profile server uses the methods developed in our project to calculate the recommendations and sends John back the information about the suggested rooms / areas to visit.

After obtaining the initial recommendations, John walks around the museum. He can get information – when he explicitly wants – for rooms and separate objects:

- Each object is represented by a URI.
- When entering an interesting room or looking at an interesting object he either scans the RFID tag at the door/object or types in the number near the tag.

The PDA fetches the required information: it reads both the basic data directly from the RFID along with the the object URI for accessing additional information from the server.

The PDA will give John initial information about the object, taking into account John's user profile. For example, the profile determines whether John prefers audio or text, whether John is typically interested about the author or historical context of the object. Further information about the object is available by browsing.

The PDA regularly updates John's personal profile. First, John can – if he so wishes – mark the object in the PDA as “I like it” or “I do not like it”. Second, the software measures the approximate time of looking at the object and browsing the available information about the object. Finally, the system allows John to type in short comments for the object, possibly including links to his blog where more remarks are given. All such information is stored in John' s profile.

When back at home, John can have a look at his profile at his PC using a browser. He can have a more detailed look at interesting items using ordinary web browser, not the PDA with a tiny screen used in the museum. John can – if he so wishes – also edit his profile in the web application.

III. THE RECOMMENDATION SOLUTION

Several museum trial systems have been deployed to improve user experiences through the webpage bookmarking solutions - e.g., Tate Modern (Multimedia Tour), Getty (Guide). These systems enable post-visit digital content study through the collected URLs. However, it appears that such delayed content access services are not attractive for the normal users who would prefer to acquire information immediately at the exhibition site [5]. This is one of the reasons why the SMARTMUSEUM solution mainly aims to improve *on-site* cultural heritage experiences by shortening and simplifying selection processes through the personalized recommendations in two domains:

- object (artifact) recommendation,
- content recommendation for particular artifact.

SMARTMUSEUM combines rule-based, collaborative and content (semantics) based recommendation solutions for both domains to improve user satisfaction. Comparing with the above described implementations, the current solution uses more physical context information for all the recommendation components. Additionally to GPS positioning and RFID based user interest monitoring we also collect web page access duration information for ranking purposes similarly to [11], which is especially important for structuring user generated content links.

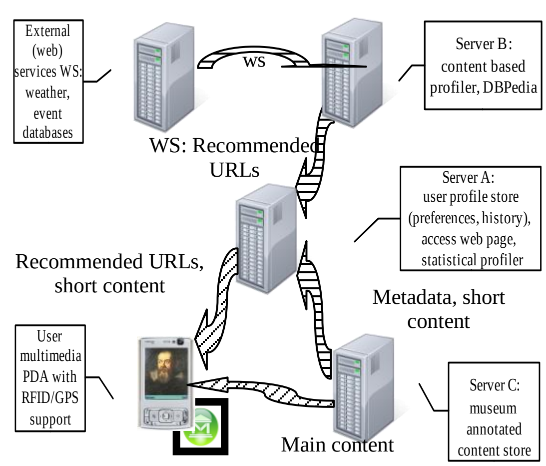

Figure 1. SMARTMUSEUM main data flow diagram

IV. USER PROFILE

The profile consists of three Resource Description Framework (RDF) [25] coded segments: user ability information, preference/interest information and visit history.

First two components are presented using widely accepted ontologies, namely GUMO and Getty - but can be extended (see [1]). Visit history as a data source for further recommendations contains URI-s of items *and* content pages user liked or disliked. The visit history component is a list of records

VisitHistory: = <*Record 1, Record 2, Record 3* … >

where *Record* contains *objectID* (global artifact identifier, usually created by a museum), visit context and group information described below, visited content URL-s and ranking for object and content:

<*Rank* rdf:datatype="&xsd;float">0.9</Rank>

After the analysis of user scenarios a combined scale of manual preference input and implicit monitoring of preference and behavior was developed:

- 1: strong like, manually input.
- 0.7: user fetches a lot of information about the object, e.g., follows several provided url-s or browses content more than 3 minutes.
- 0.3: user fetches basic information about the object.
- -0.01: user visits room with the object, shows no interest.
- -1: strong dislike, manual input.

V. THE BASIC RULE SYSTEM

The rule based recommendation component of the SMARTMUSEUM solution is used for both filtering out the objects and content that is based on

- individual abilities and preferences implicitly described in the personal profile
- visit context information specified at the beginning of the tour.

For example, objects accessible via the staircase are only filtered out for visitors requiring wheelchair access, content presentation in foreign languages may be disabled and users may specify if they prefer instant playback of the related multimedia content.

For ability-based recommendations the SMART-MUSEUM solution employs a limited subset of the well known GUMO user modeling ontology [16]. Supported user properties include physical abilities, demographics and a limited set of social roles.

The POI (point of interest) descriptions of responsible content providers shall include GUMO compatible metainformation describing

- physical access restrictions of POI,
- content features that may be related to abilities.

The SMARTMUSEUM-compatible html content pages should include invisible control tags with specific uri-s indicating e.g., the location of related audio-video information. For example, the suitably tagged html fragment may be used by the PDA software to automatically launch video playback for the majority of users. The on-device text-to-speech synthesis is used for visitors with visual impairments.

Figure 2. Initial profile setup user interface

Additionally, embedded metatags are used for specifying (main) *target group* e.g., busy, greedy; *education level* and (expected) use *motivation* e.g., adventure, art_and_culture, of particular piece of content. For example, content prepared by the Institute and Museum of the History of Science in Florence for evaluating the developed recommendation solution is the following:

<sm-schema:**ageGroup** rdf:resource="http://e-culture.multimedian.nl/ns/getty/aat#300154397"/>
<dcterms:isPartOf rdf:resource="urn:imss:multimedia:500035"/>
<dc:subject xml:lang="it">Collezionismo scientifico</dc:subject>
<dc:title xml:lang="it">Strumento topografico</dc:title>
<dcterms:**educationLevel** rdf:resource="http://e-culture.multimedian.nl/ns/getty/aat#300254154"/>
<rdf:type>http://smartmuseum.eu/types/Instrument</rdf:type>
<sm-schema:**purposeOfVisit**>ArtAndCulture</sm-schema:purposeOfVisit>

More complex filtering rules combining statistical and semantic recommending engines are used for limiting the amount of recommended objects and content URL-s based on the *context* and *group* information given by the user at the beginning of the museum tour:

- expected visit duration;
- purpose of the visit (selection from the list): adventure, advanced study, leisure, education, etc.
- group context: alone, with family, with friends, with a tourist group, with a school group.

The motivation of the context information and related rules stems from the natural multidimensionality of user preferences, rarely addressed in recommendation systems [14]. Obviously, the selection of the recommended museum objects shall depend on the role of the visitor - is he together with a child, a father or a friend at the particular moment and is the purpose of the visit collecting detailed information or just leisure. When recommendation rules based on the user ability information are most applicable for object recommendations, the context and companion information is mainly targeting content optimization. Although the rules for limiting object recommendations and content triggers are simple, selecting the appropriate context dependent content is a complex task where the usage of the statistical information from the real visits is crucial.

VI. Ontologies, annotations and context

As mentioned before, content based recommendation solutions are in the leading position in the cultural heritage access domain. The method does not have serious cold start problems and works fine with a small amount of users - especially since the cultural preferences are rather stable. If the user tends to visit objects that represent the style *renaissance* and are made by *Italian artists*, it is very likely that the user would prefer such objects also in the future.

The ontology-based recommendation component is used to provide recommendations based on the user interest profile and the current semantically described context. The interest profile is approximated based on the metadata of the items the user has tagged in different contexts. Approximation of user preferences based on metadata requires annotations that describe the objects in terms of ontologies and metadata schemas. In this way the ontologies can be used to overcome the semantic gap between the annotation and the user profile. For example, if the user has a tendency to visit objects that represent C*arolingian* style, but exactly those are not available in the museum that the user is visiting, the system is able to recommend the user other objects representing *medieval style* (Carolingian is a special case of medieval styles).

Annotations and metadata corresponding to metadata schemas are stored in the SMARTMUSEUM server and can be further used as a reference to the actual content that is stored in a decentralized way as HTML-pages. Annotations are described in the RDF language and use ontologies that provide a controlled set of concepts and instances. For this study we used the Getty vocabularies [18] that contain authoritative data on the artist names (ULAN), current and historical place names important for cultural heritage domain (TGN) and the general indexing terms for art and architecture (AAT).

For structuring the annotations we used a metadata schema that extends the Dublin Core [24] to represent different aspects of the objects, such as a type of the object, place of manufacture, creator and subject matter.

The initial SMARTMUSEUM user interest profile can be set up through the web interface offering keyword selection of the whole Getty AAT vocabulary. Later the user profile is updated based on a tagging behavior of a user. When the user sees an interesting item inside a museum, the user is able to tag the item with the ”I Like”-tag through the PDA interface. The user profile is then updated accordingly. Each triple in the user profile is attached to a context in which the tagging of the object occurred. For example, if a user tags Italian paintings only in Helsinki and outdoor locations in Italy, the system can prefer outdoor locations in southern Europe and museums exhibiting Italian paintings in northern Europe.

A user profile is a set of profile items that represent individual RDF(S) triples that originate from the annotations of objects. A profile item *pi* is a triple

$$pi = \langle t, ct, w \rangle$$

where *t* is a triple, *ct* is a context of the triple and *w* is a weight for the triple *t* in a context *ct*. The weight of the profile item can be observed from its maximum likelihood.
The maximum likelihood of a triple is its count normalized by the number of all triples in the profile. Referring to our former example, if a user has tagged Italian paintings in Helsinki, say 10 times, and there are 20 entries in the profile, the $P(\langle Painting, manufacturedIn, Italy\rangle\ Helsinki) = 0.5$. However, the problem with direct estimation like this is that although the user has indicated liking of Italian paintings, the contexts in which these observations are done can be very sparse. Therefore, we use Laplace (i.e. add one) smoothing to shave a share of the probability mass to contexts for which no observations are available. For example, determining a probability $P\ (\langle Painting, manufacturedIn, Italy\rangle\ Italy)$ with Laplace smoothing would be 1/21 = 0.047. In this way we are still able to determine probabilities other than zero, even in case the observed context triples do not match the users current context.

Context profile represents the user current context: for example, the location of the user at the moment when the recommendations are requested. At this point the context profile consists of positioning information that gives spatial restrictions for items to be recommended. It is based on the location of the user. Spatial restrictions can be a triple in the context profile indicating that the user is inside a certain museum (inside scenario) that limits the objects to be retrieved to those that are exhibited in the particular museum that the user is in. Spatial restriction can also be a bounding box of WGS84 coordinates that limit the objects retrieved to those that are close to the user's current location (outside scenario). The third use case, where user has read the RFID aside an physical object, the URI of the physical object is inserted to the user profile. In this case, the recommendations are limited to the ones that are directly attached to this specific item.

Recommendation retrieval is based on the triples in the user profile and in the context profile. Based on the earlier phases we have a set of profile triples that each have a weight that is calculated by determining their probability based on user current context. Each triple may be expanded using query expansion to multiple triples. Here we expand the query to all triples having Wu-Palmer measure [20] smaller than 0.5. This ensures that each triple matches all of its subsumers (e.g., Italy is included if Europe is in the profile) and sufficient amount of its super cases (i.e. Italy is included if Florence is in the profile). The Wu-Palmer measure ensures that concepts that are deep in the hierarchy can be expanded up in the tree (Florence can be expanded to

Italy), but more general concepts are not expanded (Europe is not expanded if Italy is in the profile). This query expansion reduces the sparsity problem that occurs often in the recommendation scenarios, because ontologically similar objects can be recommended to users even if they do not exactly match the triples in the users profile.

The weighted triples determined from the user and context profiles are then used to perform a query to the knowledge-base. We index each triple as all combinations that can be determined using subsumption reasoning. We use a matrix of documents times triples. For example, for the triple

<Item1,manufacturedIn,Helsinki>

we store all combinations reachable by reasoning, i.e.

<Item1,relatedTo,Helsinki>,
<Item1,rdf:Property,Helsinki>,
<Item1,manufacturedIn,Finland> , ...

These combinations formulate a vector space for the documents that is normalized on the basis of the number of triples in the annotation and based on each triples information values using tf-idf [10]. The triples in the context and user profiles formulate a vector space for the profile. To determine the recommendations we calculate a cosine similarity of the profile and document vectors. This determines the ranking for the objects to be recommended for the user. The vector space indexing and retrieval is implemented using Apache Lucene.

VII. Data mining based recommendations

The data mining based approach predicts new objects using the activity pattern similarity without any knowledge of the objects themselves. Thus, the data mining approach requires more historical information than content driven solution. However, the power of collaborative recommendation systems in the cultural heritage domain should not be underestimated: to date a huge amount of legacy content is not sufficiently annotated for the purposes of the content based recommendation method. Additionally, average users cannot or are not interested in expressing their interests using field specific terminology for proper interest profile setup. Therefore the cold start problem still exists and the "interest evaluation" interviews or demo object presentations are practical even for mostly semantics oriented profilers e.g., CHIP [17].

From the SMARTMUSEUM perspective the choice between various data mining methods boiled down to a choice between two classical approaches. Whether to find binary recommendations: based upon previous information, this user will like a following list of cultural heritage objects. Or to try to predict the probable value or score the user would give to a specific object. Based upon such information and additional information from the profile, the goal of the system is to find users with similar preferences and make recommendations regarding objects and content URL-s upon that. Mining for *association rules* e.g. [21]**,** *frequent closed itemsets* (basically a two-mode clustering approach), e.g. [22] and *collaborative filtering* approaches account for more than 80% of the solutions typically used for this kind of problems. According to our dataset size estimations of a million objects/content URL-s, one or even a few servers would not support the calculation of full object times object distance matrix. Same holds for several different versions of collaborative filtering algorithms, which do not provide exact numeric predictions and are essentially inferring the users' taste similarly to classical "market basket analysis" methods like association rule and frequent closed itemset mining.

Frequent closed itemset mining was chosen for item-based recommendations and is implemented in the SMARTMUSEUM system for the following reasons:

- Fast answers to queries.
- The construction of the full object*object distance matrix is not required.
- Symmetric recommendation is quicker and more reasonable in practice.
- Frequent closed itemsets are essentially binary biclusters from two-mode clustering, which has the advantage to identify patterns which are not valid for the whole database, but only a small set of users and objects.

For technical reasons we transmit the user rating information just once to the profile repository. Manual ratings override the background monitoring results. That way the data to be processed by the clustering algorithm is also minimized.

A walkthrough from the use case and data storage point of a view would be as follows. A browsing and walking session has to be linked with a specific *userID*. Therefore, if we visit a specific object (with *objectID*) and either rank it manually or tag with other scores as described earlier in this section, we have to store at least the following data to meet the requirements of our algorithms:

<userID,objectID,ranking>

Knowledge about the preferences of users is built up on such information periodically, employing the following scheme:

- Preprocessing of the database to meet the data mining algorithms input.
- Postprocessing of the found patterns to meet the operational level recommendations to be uploaded in the handheld interface or browsed via a web browser.

As a result, a database of patterns of objects which people tend to like together is built up offline. It is possible to query beforehand (at the entrance point) all the object recommendations or - for example - the recommended route for the visit. One needs to query all the patterns containing at least some of the objects in that museum. With nested queries and database joins it is possible - even without additional data analysis - to link the *objectID* with the museum and other preferences, therefore filtering the result in several ways. For example:

- Exclude everything that the person has explicitly stated not liking or rank higher (ranking is based on the patterns support by frequency) objects with the properties/attributes that the person has explicitly stated to like.
- Filter out cross-patterns between places of interests: the *objectID* has to be linked with the museum and a query could show all the patterns, where there are objects that this person has liked, and filter out only objects which are located in particular museum.

Proper combination of content and statistics based recommendation results, especially selecting proper weight factors is a task of the upcoming evaluation phase of the SMARTMUSEUM solution.

CONCLUSION AND FUTURE WORK

We have presented the design and methods used in the SMARTMUSEUM platform, employing adaptive and privacy preserving user profiling. The described recommendation system relies on combining a semantics/ontologies based approach with a data mining/statistics based approach. We will continue by deploying the system for actual use and collecting feedback, with the goal to improve the described methods in realistic setting.

ACKNOWLEDGMENTS

This work is supported by the European FP7 project SMARTMUSEUM [23].

REFERENCES

[1] N. Dokoohaki and M. Matskin, Personalizing Human Interaction through Hybrid Ontological Profiling: Cultural Heritage Case Study", 1st Workshop on Semantic Web Applications and Human Aspects, (SWAHA08), M. Ronchetti, Ed., In conjunction with Asian Semantic Web Conference 2008, AIT e-press, 2008, pp.133-140.

[2] G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper and M. Pinkerton, (1997). Cyberguide: A mobile context-aware tour guide, Wireless Networks, 3, pp. 421–433.

[3] K Cheverst, N. Davies, K. Mitchell, A. Friday and C. Efstratiou, Developing a context-aware electronic tourist guide: some issues and experiences. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp.17–24). The Hague, The Netherlands, 2000.

[4] J. Fink and A. Kobsa, User Modeling for Personalized City Tours. Artificial Intelligence Review, 18, pp.33–74, 2002..

[5] S. Filippini-Fantoni and J. Bowen, “Bookmarking In Museums: Extending The Museum Experience Beyond The Visit”, Museums and the Web 2007, the International conference for culture and heritage on-line.

[6] M. Balabanovi and Y. Shoham. Fab: Content-Based, Collaborative Recommendation. *Comm. ACM*, vol. 40, no. 3, pp. 66-72, 1997.

[7] E. Hyvönen, E. Mäkela, M. Salminen, A. Valo, K . Viljanen, S. Saarela, M. Junnila and S. Kettula, MuseumFinland—Finnish museums on the semantic web. Journal of Web Semantics, 3(2), pp. 224–241, 2005.

[8] T. H. W. Lam, and R. S. T. Lee, “iJADE FreeWalker – An Intelligent Ontology Agent-based Tourist Guiding System”, Studies in Computational Intelligence (SCI), 72, pp. 103–125. 2007.

[9] F. M. Shipman, C. C. Marshall, and T. P. Moran, “Finding and using implicit structure in humanorganized spatial layouts of information,”in *Proc. Of CHI*, pp. 346–353. 1995.

[10] C. Chen, “Generalised similarity analysis and pathfinder network scaling,” *Interacting with Computers*, vol. 10, pp. 107–128, Feb. 1998.

[11] N. Hoebel and R. V. Zicari, "Creating User Profiles of Web Visitors Using Zones, Weights and Actions," cecandeee, 2008 10th IEEE Conference on E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services, pp.190-197, 2008

[12] F. Sparacino, “Real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences”, In: Proceedings of Museums and the Web (MW2002), pp. 17-20 2002.

[13] W. Rudametkin, L. Touseau, M. Perisanidi, A. Gómez and D. Donsez , NFCMuseum: an Open-Source Middleware for Augmenting Museum Exhibits, International Conference on Pervasive Services (ICPS'08), Demonstration track, July 6-10, Sorrento, Italy, 2008.

[14] C. Niederée, S. Avaré and M. Bhaskar, “A Multi-Dimensional, Unified User Model for Cross-System Personalization”, workshop on Environments for Personalized Information Access. - Gallipoli, Italy: Advanced Visual Interfaces (AVI’04), May 25, 2004.

[15] Y. Wang, N. Stash, L. Aroyo, P. Gorgels, L. Rutledge, and G. Schreiber, Recommendations based on Semantically-enriched Museum Collection, Web Semantics: Science, Services and Agents on the World Wide Web, Volume 6, Issue 4, pp. 283-290, November 2008.

[16] D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz and M. von Wilamowitz-Moellendorff, GUMO – the General User Model Ontology. User Modeling 2005: 10th International Conference, UM 2005. - Edinburgh : Birkhäuser, 2005.

[17] L. Aroyo, R. Brussee, L. Rutledge, P. Gorgels, N. Stash and Y. Wang, Personalized museum experience: The Rijksmuseum use case. In J. Trant, & D. Bearman (Eds). Museums and the Web 2007: Proceedings. Toronto: Archives & Museum Informatics, 2007.

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

[19] M. Baldauf, S. Dustdar, and F. Rosenberg, A survey on context-aware systems. International Journal Ad Hoc and Ubiquitous Computing, 2(4), pp. 263–277, 2007.

[20] Z. Wu and M. Palmer. Verbs semantics and lexical selection. In Proceedings of the 32nd annual meeting on Association for Computational Linguistics, pages 133–138, Morristown, NJ, USA, 1994. Association for Computational Linguistics.

[21] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules”, in Proc. 20th Int. Conf. Very Large Data Bases, VLDB, 1994.

[22] M. J. Zaki and C.-J. Hsiao, "CHARM: An Efficient Algorithm for Closed Itemset Mining", SIAM International Conference on Data Mining, Chicago, 2002.

[23] Smartmuseum, http://www.smartmuseum.eu

[24] Dublin Core Metadata Element Set, Version 1.0: Reference Description,. http://dublincore.org/documents/1998/09/dces/

[25] RDF, http://www.w3.org/RDF/