Personalizing Human Interaction through Hybrid Ontological Profiling: Cultural Heritage Case Study

Nima Dokoohaki[1], Mihhail Matskin[1 2]

[1] Department of Electronics, Computer and Software Systems (ECS),
School of Information and Communications Technology (ICT),
Royal Institute of Technology (KTH), 164 40 Kista, Sweden
[2] Department of Information and Computer Science,
Norwegian University of Science and Technology (NTNU),
Trondheim, Norway
{nimad,misha}@kth.se

Abstract. In this paper we present a novel user profile formalization, which allows describing the user attributes as well as history of user access for personalized, adaptive and interactive experience while we believe that our approach is applicable to different semantic applications we illustrate our solution in the context of online and onsite museums and exhibits visit. We argue that a generic structure will allow incorporation of multiple dimensions of user attributes and characteristics as well as allowing different abstraction levels for profile formalization and presentations. In order to construct such profile structures we extend and enrich existing metadata vocabularies for cultural heritage to contain keywords pertaining to usage attributes and user related keywords. By extending metadata vocabularies we allow improved matchmaking between extended user profile contents and cultural heritage contents. This extension creates the possibility of further personalization of access to cultural heritage available through online and onsite digital libraries.[1]

Keywords: personalization, user profiling, hybrid user modeling, semantic user profiles, cultural heritage metadata.

1 Introduction

Lessons learned from the adoption of Web technologies helped researchers realize the importance of human factors. From user perspective, Semantic Web faces many challenges in successful deployment [1]. Interactivity and usability of technologies are dependable on infrastructure. Semantic Web enabling infrastructure continues being developed and utilized. Heterogeneity of resources and users across Internets and intranets hinders the mass adoption of these tools. Considering increasing proliferation of semantic Web driven technologies and tools, personalization

[1] This work has been done within the FP7-216923 EU IST funded SMARTMUSEUM project. The overall objective of the project is to develop a platform for innovative services enhancing on-site personalized access to digital cultural heritage through adaptive and privacy preserving user profiling. Using on-site knowledge databases, 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.

techniques could build a bridge between the users and Semantic Web [2]. In short, personalization is customizing the information content or the adapting the visualized experience of the system to the user's preferences and interests. Personalization and personalized systems came into life, as a result of research and study into such drastic increasing problem across Internet and intranets world-wide. By constructing personalized information systems which wrap existing legacy systems, their databases and their contents, we can facilitate access towards existing information content decentralized across different databases through personalized information retrieval. In order to personalize the process of information retrieval, on behalf of the user personalized system creates a model of usage behavior and further develop such model by creating a profile of the user which documents the history of usage and attributes of the usage. Through this observation, system understands the preferences of the users. By understanding the preferences of the users, system can tailor the information to users' needs. In order to construct a correct model of the user, we need to understand from which dimensions and perspectives users will observe and experience the system. As a matter of fact, a profile that documents the experience and behavior of the user will incorporate certain attributes which pertain to these dimensions and perspectives. In order to point out features of such profile structure and format, we take a generic approach which allows incorporation of all sorts of user attributes as well as allowing the documentation of history of experience of user as well. We incorporate ontological description of user data in order to ease the interoperability of content, when shared across multiple systems, as well as adding meaning and semantics into the concepts of usage domain. Along with the content, we also describe how this format and structure allows the contextual information to be documented and presented.

While we consider our approach to personalization as general enough to be used in different semantic web applications our case study mostly concerns systems providing access to cultural heritage. In this context we extend cultural heritage metadata keyword sets to include attributes pertaining to user behavior. By including usage attributes in the vocabularies along with existing metadata we provide the possibility of query extension to include a subset of these keywords which are already documented in ontological user profile. When querying digital cultural heritage content, these keywords are matched against museum digital metadata content and allow personalized matchmaking of items and user profiles.

The rest of the paper is organized as follows; in the second section, background is presented, in the third section the profile structure is introduced, in the forth section we introduce the extension of a sample metadata vocabulary with user attributes and we also describe the process of matchmaking of profiles and extended metadata. We finally conclude in the fifth section.

2 Background

Nora Koch [3] describes a user profile as a simple user model. A user profile is a collection of personal information. The information is stored without adding further description or interpreting this information. It is comparable to a getting-setting mechanism of classes in object-oriented programming, where different parameters are set or retrieved. User profiles represent cognitive skills, intellectual abilities, and

intentions, learning styles, preferences and interactions with the system. These properties are stored after assigning values to them. These values may be final or may change over time. Depending on the content and the amount of information about the user, which is stored in the user profile, a user can be modeled. Thus, the user profile is used to retrieve the needed information to build up a model of the user. Koch also describes a user model as the representation of the system's beliefs about the user. The "real world" user is perceived by the system through the human computer interface. According to Wahlster and Kobsa [4], information about the user is usually collected in a so-called user model and administrated by a user modeling system. They define (in the context of a dialog system) the following two fundamental concepts: "A user model is a knowledge source in a system which contains explicit assumptions on all aspects of the user that may be relevant to the behavior of the system. These assumptions must be separable by the system from the rest of the system's knowledge." User profiling is either knowledge-based or behavior-based according to Middleton et al [5]. Knowledge-based approaches engineer static models of users and dynamically match users to the closest model. Questionnaires and interviews are often employed to obtain this user knowledge. Behavior-based approaches use the user's behavior as a model, commonly using machine-learning techniques to discover useful patterns in the behavior. Behavioral logging is employed to obtain the data necessary from which to extract patterns, according to Wahlster et al [4]. Fröschl [6], states that difference between user profiling and user modeling relies in the different levels of sophistication. He states that, in general, the profiles contain "raw material" gathered and acquired from a user while, when such data is processed it will be used to build up a model of user, creating a sophisticated perception of user. Hybrid modeling and profiling have been discussed extensively [7] [8]. Hybrid user modeling can be defined as combining user attributes and content attributes for improving personalization effect. Profiles created and composed based on matching extended user profiles and items, are referred to as hybrid profiles. Hybrid approaches to user modeling and profiling, are either focused on combining strategies for profiling [9] or on combining user models [10]. We have introduced a hybrid model which is expressive enough to allow different sort of information about the user (such as attributes and preferences, observational data, user context) to be documented and presented. We have mostly considered combination of usage attributes and content attributes for personalization of access.

3 Profile Structure

In this section, we define and introduce a structure that could be used for saving and retrieving different types of information that document both behavior and knowledge aspects of the user. We define a user profile as a structured collection of personal information about the user that has certain perspectives or dimensions which covers different aspects of the personal attributes of the user. Profile content, documents the personal information about the user as well as history and evidence of the experience of the user who is being profiled. Profile has depth (hierarchy) as well as length (flat structure), allowing us to create a level measure for details incorporated into the profile.

3.1 Profile Segments

In addition to user details we need a structure which could record the history user experience as well as the weight of the information presented. Since privacy is a crucial and very important aspect of user profiling, we would like to also incorporate security information describing the privacy of the profiled information as well as trusted arguments pertaining to profiles. In order to incorporate such information, along with the weight of the information and security credentials protecting and enriching such information we will divide the structure into different sections. We refer to each section of the records a segment. The first segment of profiled materials is context segment. We assume that all surrounding facts could be considered context and information contained within them could be considered contextual information. In order to make such format more generic, we can avoid inserting values directly to context segment and instead we can give a reference to existing contextual information. For instance we can define a context ontology that documents contextual concepts and it becomes populated when contextual information is available. Instead of using solely attribute and value pairs, we use RDF (resource description format) described triplets of predicate, subject and objects to describe information contained in the profile. This segment will contain the actual material which is profiled. For

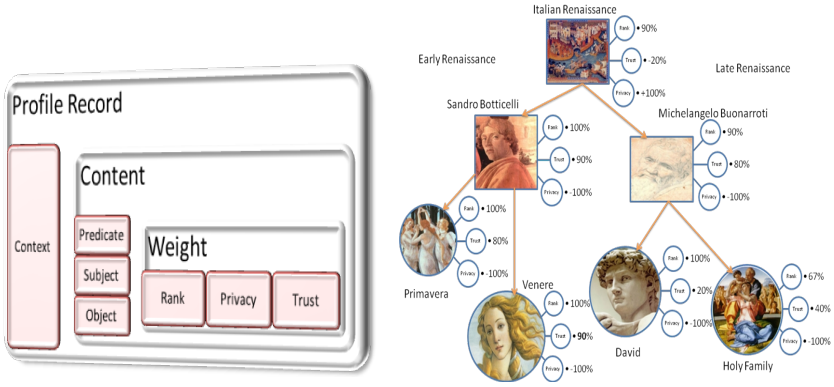

Fig. 1. Structure of User Profiles: Each profile could be composed of multiple records. Structure of a single record is depicted on left side, while a hierarchical (high-level) presentation of a user's profile is depicted on right side.

instance, "interest of user in science of art" can be described as (*hasInterest, user," Science_of_art*") tuple. Here subject is used to describe the attribute and object is used to describe the value and the predicate describes the semantics of the relations of concepts. Using RDF-triplets eases the later extraction and mapping of RDF data to lower level formats such as XML or higher level formats such as OWL. Each profile record contains a weight segment which can specify the weight of information being profiled. For instance if the profiled information is about user's cognitive patterns such as interests, then the weight would describe the intensity of interest of user in the object atom specified in RDF triplets, meaning for instance, how much user is interested in a certain artifact or artist. In addition to the profiled information and its designated weight we incorporate a segment for security credentials. These credentials can have different semantics in different cases and scenarios. But in

general, trust describes the trust, belief and confidence of the user towards the piece of information profiled while privacy describes the privacy of the piece of information recorded. For instance, privacy could have values between range [-1, 1] where positive values could describe the positive consent of the user towards sharing, while negative value could describe the negativity of the user towards disclosure of the content recorded to outside world. Using such weighted information user can specify which atomic piece of information he/she would like to disclose or not to disclose to outside world. Trust can be interpreted differently depending on the case scenario. For instance, trust in an artwork could document the originality of work being experienced by the user, while trust in profiled information documents the trust of user in the information piece documented and profiled or just trust of the user in system. Such atomic representation of trust could also be used for describing the trust in individuals while describing relationships between users.

4 Extending Metadata with Human User Metadata

In order to be able to extend existing keywords used for describing cultural heritage to include user attributes, first we need to understand and identify attributes that document user behavior.

4.1 Constructing a Smart User Model Ontology

We have identified four major categories attributes: attributes that document the user's demographics such as age, languages and gender, attributes that document preferences of the users such as system, device and personal preferences, attributes that document abilities and disabilities of user such as hearing abilities or walking

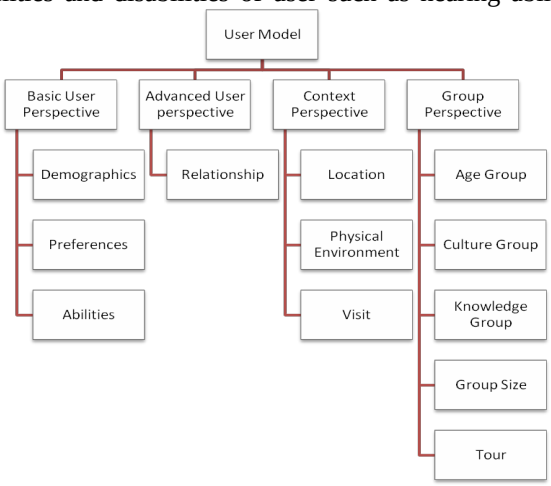

Fig. 2 . Attribute Categories: identified user attributes are categorized into super and sub-categories that allow development and design of ontological classification.

disabilities and attributes that document social aspects of the users such as role and relations. In addition to user attributes we have identified and modeled attributes that document the context of use [11], as context plays an important role in personalizing

the user's experience. Attributes considered for documenting and presenting context are for instance, location, environmental attributes such as humidity, and visit, such as goal of visit or companion of the visit. Attributes mentioned so far pertain to individual usage, rather than group usage. As a matter of fact, we have also identified and documented attributes such as grouping users based on age, knowledge or culture. We have considered utilizing these attributes for target grouping the users. Tours are group activities and we have also considered introducing them in our attribute sets, such as tour name, fee and size. In order to categorize these attributes, we have created super categories for these attributes and we have referred to each super category as perspective. For creating this structure, we have partially used ontological structure and organization of GUMO [12] (general user model ontology), and UserML [12] to describe the attributes and perspectives which pertains to users.

4.2 Extending Schemas with Human Usage Keywords

Fig. 3 is an example of how cultural heritage metadata can be extended using the attributes identified previously. In order to describe metadata keywords describing the attributes of the visual artworks, such as place of creation, artist name and material of artwork, we have used Getty vocabularies [13][14][15]. ULAN [13] has been used for documenting the artist name and relating artist, AAT [15] has been used for documenting the material of the artwork and TGN [14] has been used for documenting the place the artwork is created. We have utilized VRA Core [16] concepts and properties to describe work of art dimensions as well as relationships in

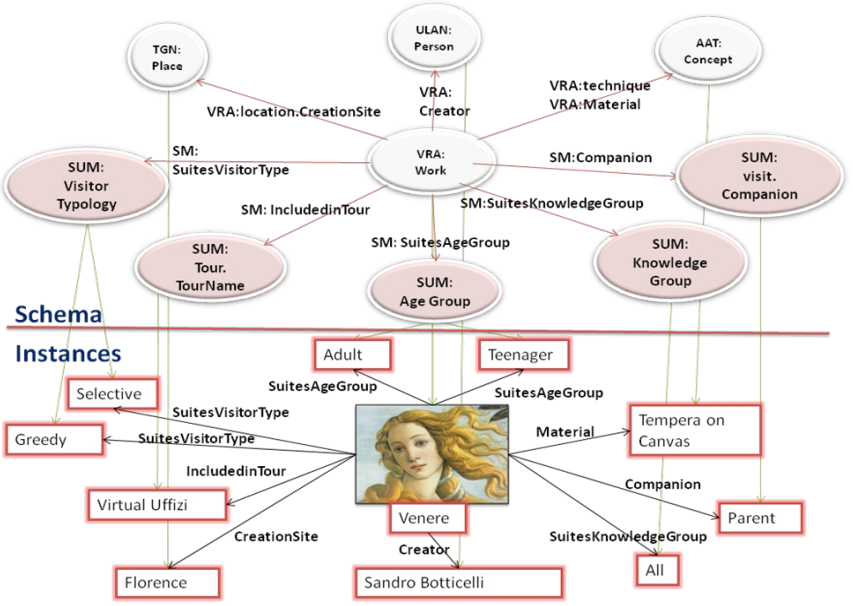

Fig. 3 . Extended CH Metadata with Human Keywords: Schema has been extended to include new concepts describing human usage domain. When these concepts become instantiated, they extend the artworks with individual, group and contextual attributes.

between the concepts, such as relation between artwork creator and artwork itself. The concepts of our user model ontology has been defined on the namespace SUM (smart user model) while the edges are defined on the common namespace of SM (smart museum). We have extended the cultural heritage ontological schema with group attributes distinguishing user groups' based on age and knowledge, allowing us to describe the recommended target group that this artwork could be useful to. Extension of Contextual attributes used here has allowed us to state the recommended companion for the visit. We have also used visitor typology that allows us to state the type of visitors [17] (greedy, busy and selective). In order to recommend the user with an existing tour that includes the artwork we have extended concept set with a tour attribute, *tourname*. When schema becomes instantiated and populated with the values, then in addition to instances that document the artwork's name and properties we have instances that document human user side that interacts with this artwork.

For instance, in our example we have Venere as a visual artwork painted by Sandro Botticelli in Florence using Tempera on Canvas. In addition to this legacy metadata, we have added that this artwork suites selective and greedy visitors, who look for lots of artistic and cultural details. This work can be subject to any audience with different knowledge backgrounds. This work suites best adult and teenage target groups, while this fact is also presented by recommending visit with parent. This work is included in Virtual Uffizi tour which can be recommended to user if interested.

4.3 Matching Profiles against Extended Personalized Metadata

In order to further facilitate personalized access to digital content, a filtering methodology should be implemented. We have considered item-user filtering for access personalization. By taking advantage of extending user profiles with user model keywords and attributes, we can allow item-user matchmaking to be implemented. This process involves expanding the query with additional human user keywords, which describe the profile of the user. Since a similar subset of these keywords is used to extend the schema for describing the artworks, then the query is personalized according to user's profile. As a matter of fact, slices of the user profiled are used to expand the semantic query to digital content. The results of the query represent the matching of user's partial interests in digital cultural heritage and provide the user with more personalized access to digital cultural content. In this process instances of cultural metadata schema are matched against the user profiled record instances.

5 Conclusion

We have presented a novel ontological user profile structure and formalization that allows documentation and presentation of user information. This structure can record and present the weight dimensions and context dimensions for user profiles on an item-wise basis. We presented this profile in ontological format. A user model hierarchy of keywords or attributes were developed to allow expansion of legacy cultural heritage metadata for personalized access to cultural contents. As a future work, we intend to create algorithms for self-adaptive profile management based on our profile, as well as developing and implementing services allowing onsite and

extra-site systems to utilize our profiles for providing users with recommendations and personalized information.

References

1. Heath. , T. Domingue., J. Shabajee., P .User Interaction and Uptake Challenges to Successfully Deploying Semantic Web Technologies. Proceedings of the 3rd International Semantic Web User Interaction. 2006.
2. Dolog., P. and Nejdl, W. Challenges and Benefits of the Semantic Web for User Modelling - Proceedings of Adaptive Hypermedia, 2003.
3. Koch., Nora. Software Engineering for Adaptive Hypermedia Systems. Munich, Germany: PhD thesis, Ludwig-Maximilians University, May 8th, 2005.
4. Wahlster. , W. and Kobsa. , A. User models in dialog systems. [Book auth.] A. Kobsa and W. Wahlster. User Models in Dialog Systems (Symbolic Computation Series). Berlin: Springer, 1989, pp. 4–34.
5. Middleton. , Stuart E. Shadbolt. , Nigel R. and De Roure. , David C. Ontological User Profiling in Recommender Systems. ACM Transactions on Information Systems. 2004.
6. Fröschl, Christoph. User Modeling and User Profiling in Adaptive E-learning Systems. Graz, Austria: Institute for Information Systems and Computer Media (IICM), Graz University of Technology, November 2005. PhD Thesis.
7. Brusilovsky, P., Kobsa, A. and Nejdl, Wolfgang. s.l.: The Adaptive Web, Methods and Strategies of Web Personalization. Springer, 2007.
8. Zakaria, M., et al. 2002. The hybrid model for adaptive educational hypermedia. Ah'02: Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web Based Systems.
9. Poo, Danny, Chng, Brian and Goh, Jie-Mein. Hawaii. A Hybrid Approach for User Profiling. Proceedings of the 36th Annual Hawaii International Conference on System Sciences 2003, p. 103. 0-7695-1874-5.
10. Billsus, Daniel and J.Pazzani, Michael. Banff. A Hybrid User Model for News Story Classification. Canada: Springer-Verlag New York, Inc, 1999. Proceedings of the seventh international conference on User modeling. pp. 99 - 108. 3-211-83151-7.
11. Niederée, Claudia, Stewart, Avaré and Mehta, Bhaskar. A Multi-Dimensional, Unified User Model for Cross-System Personalization. Gallipoli, Italy: Advanced Visual Interfaces (AVI'04), Workshop on Environments for Personalized Information Access, May 25, 2004..
12. Heckmann, D, et al. GUMO - the General User Model Ontology. Edinburgh: Birkhäuser, 2005. User Modeling UM 2005: 10th International Conference, 2005.
13. The Getty Foundation. ULAN: Union List of Artist Names. http://www.getty.edu/research/tools/vocabulary/ulan/.
14. The Getty Foundation. TGN: Getty Thesaurus of Geographic Names. http://www.getty.edu/research/tools/vocabulary/tgn/.
15. The Getty Foundation. AAT: Art & Architecture Thesaurus. http://www.getty.edu/research/tools/vocabulary/aat/.
16. VRA Core Categories. Visual Resource Association. http://www.vraWeb.org/resources/datastandards/vracore3/categories.htm
17. Sparacino, Flavia. Museum Intelligence: Using Interactive Technologies for Effective Communication and Storytelling. Berlin: ICHIM, 2004.