Web Semantics: Science, Services and Agents on the World Wide Web 6 (2008) 283–290

Contents lists available at ScienceDirect

Web Semantics: Science, Services and Agents on the World Wide Web

journal homepage: www.elsevier.com/locate/websem

Recommendations based on semantically enriched museum collections

Yiwen Wang[a,*], Natalia Stash[a], Lora Aroyo[a,b], Peter Gorgels[c], Lloyd Rutledge[d], Guus Schreiber[b]

[a] *Eindhoven University of Technology, Computer Science, P.O. Box 513, 5600 MB Eindhoven, Netherlands*
[b] *Free University Amsterdam, Computer Science, Netherlands*
[c] *Rijksmuseum Amsterdam, Netherlands*
[d] *Telematica Institute, Netherlands*

ARTICLE INFO

Article history:
Received 1 May 2008
Received in revised form 31 August 2008
Accepted 15 September 2008
Available online 30 October 2008

Keywords:
CHIP
Semantics-driven recommendations
Content-based recommendations
Enriched collections
Cultural heritage vocabularies
Interactive user modeling dialog
Museum tours
Mobile museum guide

ABSTRACT

This article presents the CHIP demonstrator[1] for providing personalized access to digital museum collections. It consists of three main components: Art Recommender, Tour Wizard, and Mobile Tour Guide. Based on the semantically enriched Rijksmuseum Amsterdam[2] collection, we show how Semantic Web technologies can be deployed to (partially) solve three important challenges for recommender systems applied in an open Web context: (1) to deal with the complexity of various types of relationships for recommendation inferencing, where we take a content-based approach to recommend both artworks and art-history topics; (2) to cope with the typical user modeling problems, such as cold-start for first-time users, sparsity in terms of user ratings, and the efficiency of user feedback collection; and (3) to support the presentation of recommendations by combining different views like a historical timeline, museum map and faceted browser. Following a user-centered design cycle, we have performed two evaluations with users to test the effectiveness of the recommendation strategy and to compare the different ways for building an optimal user profile for efficient recommendations. The CHIP demonstrator received the Semantic Web Challenge Award (third prize) in 2007, Busan, Korea.

1. Introduction

Museum collections contain large amounts of data and semantically rich, mutually interrelated metadata in heterogeneous distributed databases [1]. Semantic Web technologies act as instrumental [2] in integrating these rich collections of metadata by defining ontologies which accommodate different representation schemata and inconsistent naming conventions over the various vocabularies. Facing the large amount of metadata with complex semantic structures, it is becoming more and more important to support users with a proper selection of information or giving serendipitous reference to related information. For that reason, as observed in [3,4], recommender systems are becoming increasingly popular for suggesting information to individual users and moreover, for helping users to retrieve items of interest that they ordinarily would not find by using query-based search techniques. From a museum perspective [5], personalized recommendations do not only help visitors in coping with the threatening "information overload" by presenting information attuned to their interests and background, but is also considered to increase user's interest and thus stimulate them to visit the physical museum as well.

The Web 2.0 phenomena enables an increasing access to various online collections, including also digital museum collections. The users range from first-time visitors to art-lovers, from students to elderly. Museum visitors have different goals, interests and background knowledge. With the help of Web 2.0 technologies they can actively participate on the Web by adding their comments, preferences and even their own art content. Meanwhile, Web languages, standards, and ontologies make it possible to make heterogeneous museum collections mutually interoperable [1] on a large scale. All this transforms the personalization landscape and makes the task of achieving personalized recommender systems even more challenging.

In this article, we present work done in the CHIP project. The rest of the article is structured as follows. In Section 2, we discuss the research challenges, in particular, for recommendations in the open Web context. Then, in Section 3 we explain how the museum collection is enriched by using common vocabularies and in Section 4 we elaborate on the content-based recommendations for artworks and topics. Further, in Section 5, we describe the user model specification and explain the technical architecture (Section 6) with an illustrative use case (Section 7). Results of two user evaluations are given in Section 8. Finally, we discuss our approach and outline directions for future work.

* Corresponding author. Tel.: +31 206747367.
E-mail addresses: y.wang@tue.nl (Y. Wang), n.v.stash@tue.nl (N. Stash), l.m.aroyo@cs.vu.nl (L. Aroyo), schreiber@cs.vu.nl (P. Gorgels), p.gorgels@rijksmuseum.nl (L. Rutledge), Lloyd.Rutledge@cwi.nl (G. Schreiber).

[1] http://www.chip-project.org/demo/.
[2] http://www.rijksmuseum.nl.

1570-8268/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.websem.2008.09.002

2. Research challenges

While the open world brings heterogeneous data collections and distributed user data together, it also poses problems for recommender systems. For example, how to deal with the semantic complexity; how to enable first-time users to immediately profit from recommendations; and how to provide efficient navigation and search in semantically enriched collections. To address the issues, we identify three main research challenges for recommender systems on the Semantic Web:

(i) Enhancing recommendation strategies
In [1,6], we see examples of how ontology engineering and ontology mapping enable content interoperability through rich semantic links between different vocabularies in heterogenous museum collections. This, however, raises new problems for recommender systems applied in such a context, for example, how to deal with the semantic complexity of different types of relationships for recommendation inferencing and how to increase the accuracy and define the relevance of recommendations based on the semantically enriched collection. Currently, there are many recommendation strategies [7,8,4] to address these issues: *collaborative filtering* compares users in terms of their item ratings (e.g. Amazon.com[3] and last.fm[4]); *content-based recommendation* selects items based on the correlation between the content of the items (e.g. Pandora[5] and MovieLens[6]). Ruotsalo and Hyvönen proposed an *event-based* [9] recommendation strategy that utilizes topics from multiple domain ontologies to enhance the relevance precision. In CHIP we have deployed a *content-based* [10] strategy, which uses users' ratings on both artworks and art topics in a semantically enriched museum collection.

(ii) Coping with cold-start and sparsity problems
The heterogeneous population of museum visitors increasingly grows. However, most users are still "first-time" or called "one-time" users to both virtual and physical museums [5]. Thus, coping with the *cold-start* problem becomes even more crucial for recommender systems applied in the museum domain. In other words, how do we allow first-time users to immediately profit from the recommender system, without requiring much user input beforehand? In addition, in the process of enriching the museum collections, there is an increase in the number of and the size of semantic structures used. This exceeds far beyond what the user can rate and thus creates the problem of rather *sparse* distribution of user ratings over the collection items. It becomes difficult to recommend effectively when there are not sufficiently many ratings in a large collection. To solve these two closely related problems, a *hybrid user modeling* approach is widely used [11,4], combining both user and content centered attributes for generating recommendations. In CHIP, we follow a twofold approach. On the one hand, we build a non-obtrusive and interactive rating dialog [12] to allow for a quick instantiation of the user model, and, on the other hand, we realize this dialog over the most representative samples for the collection of artworks in order to enable a fast population of ratings on artworks and topics [10].

(iii) Supporting recommendation presentation and explanation
Due to the heterogeneous character of the data, it is becoming more and more important to facilitate navigation and search in multi-dimensional collections [13]. How to let users explore a large amount of heterogeneous information and still allow for a comprehendable overview? Among the different techniques for visualization clustering [13], faceted browsers provide a convenient and user-friendly way for hierarchical navigation, as exemplified in MUSEUMFINLAND[7] and E-culture projects[8]. In CHIP, we focus on using and exploring the effectiveness of existing techniques like Spectacle[9] and Simile[10] to cluster multiple recommendations based on properties and present them with different views (e.g. timeline and museum map). Additionally, there is also the problem of explanation, i.e. how to provide users a logic insight in recommendations based on the semantic structure of the collection. Traditional ways to cope with this is using histograms of other users' ratings or likeness to previously rated items [4]. In CHIP, explanations are given based on semantic relationships of artworks and topics, which has shown to improve the transparency for recommendations [14].

Table 1
Mappings between ARIA data and other vocabularies.

Source data	Vocabulary	Mapped topics	Total topics
Metadata techniques, materials and artists styles	AAT	283	2825
Metadata artists names	ULAN	263	485
Metadata creation sites	TGN	69	507
Metadata subject themes	Iconclass	178	503

3. Metadata vocabularies

The Rijksmuseum digital collection is stored in two databases: ARIA[11](educational Website-oriented database) and ADLIB[12](professional curator database). The current CHIP demonstrator works with the ARIA database, which consists of 729 of the museum's most popular artworks, 486 themes, 690 encyclopedia keywords and 43 catalogue terms. The ARIA database has two main problems: (i) *inconsistent descriptions*: artworks are annotated with different descriptions without using any standard vocabularies; and (ii) *flat structure*: no semantic relationships are described except for general hierarchical relationships between topics (e.g. top, broader and narrower topics) and themes, which brings a severe obstacle for content-based recommendation inference. To address this problem we have focussed on enriching the ARIA database with shared vocabularies. For this, the E-culture project provided the RDF/OWL representation using three Getty vocabularies[13](ULAN, AAT, TGN) [15] and the CATCH STITCH project produced mappings to Iconclass thesaurus[14][2]. We also use SKOS Core[15], created for the purpose of linking thesauri to each other. It specifies the *skos:narrower, skos:broader* and *skos:related* relationships between ARIA topics. Mapping to common vocabularies introduces a semantic structure to the ARIA collection. Table 1 gives an overview of all mappings.

[3] http://www.amazon.com/.
[4] http://www.last.fm/.
[5] http://www.pandora.com/.
[6] http://www.movielens.org/login.
[7] http://www.seco.tkk.fi/applications/museumfinland/.
[8] http://e-culture.multimedian.nl/.
[9] http://www.aduna-software.com/products/spectacle/.
[10] http://simile.mit.edu/.
[11] http://www.rijksmuseum.nl/collectie/ontdekdecollectie.
[12] http://www.rijksmuseum.nl/wetenschap/zoeken.
[13] http://www.getty.edu/research/conductingresearch/vocabularies/.
[14] http://www.Iconclass.nl/libertas/ic?style=index.xsl.
[15] http://www.w3.org/2004/02/skos/.

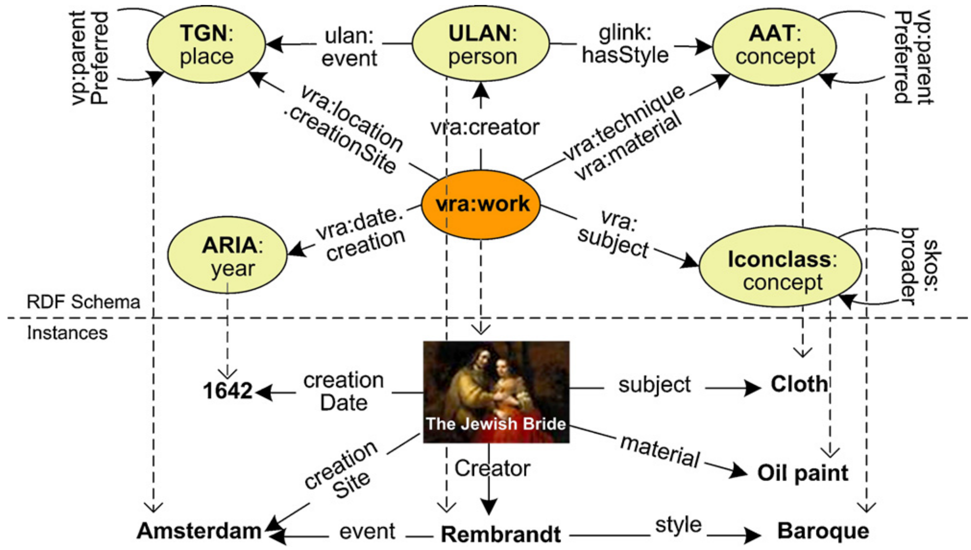

Fig. 1. Metadata vocabularies in RDF Schema.

The metadata of artworks in CHIP is defined by VRA Core[16] interpreted here to be a specialization of Dublin Core[17] for describing works of art and images of works of art. Fig. 1 gives a top-level overview of the RDF Schema used in CHIP, where concepts for places (creation places, birth and death places) in ARIA refer to the geographic location concepts in TGN; artist names in ARIA refer to artist names in ULAN; art styles in AAT are linked to artists in ULAN, and via the link to artists in ARIA the concept of 'style' is introduced in the Rijksmuseum collection; and, finally, subject themes in ARIA refer to concepts in Iconclass. For example, in Fig. 1, the artwork "The Jewish Bride" is created by "Rembrandt" (ULAN concept) in "1642" (ARIA concept) in "Amsterdam" (TGN concept). It uses material "Oil paint" (AAT concept) and has a subject "Cloth" (Iconclass concept). Artist "Rembrandt" is born in "Amsterdam" (TGN concept) and has a style of "Baroque" (AAT concept).

To enlarge the scope of the recommendations and to address the scalability aspects of our approach, we plan to include also the ADLIB database (70,000 objects) in the current demonstrator. The enrichment of this collection has already been provided by the E-culture project.

4. Content-based recommendations for artworks and topics

In CHIP, a user can start the exploration of the Rijksmuseum collection by first building a user profile, which is driven by an interactive rating dialog [16] over the museum collection. In this rating dialog, we distinguish three steps:

Step 1. The user gives ratings to both artworks and associated topics on a 5-degree continuous scale of preference.

Step 2. Based on the semantic relationships, the Art Recommender calculates a *Belief value* to predict the user's interest in other artworks and topics.

In this calculation of belief values for directly linked topics, a smoothing method, (called *Laplace smoothing*), is used: $\theta_j = (N_j + \lambda)/(N_{\text{presented}} + N_{\text{states}} \times \lambda)$ where θ_j is the probability that the user likes a topic with *j* stars, N_j is the number of times the topic appears in a set of rated artworks (e.g., artworks the user rated as "I like it"), $N_{\text{presented}}$ is the number of times the topic is presented among rated artworks, λ is the smoothing parameter (often set to 1), and N_{states} is the number of rating states (5 in our case).

Using this formula, we then calculate the belief value for topics and artworks:

$$\text{Belief}_{\text{topic}} = \sum_{j=1}^{5} \theta_j \times W_j \quad \text{Belief}_{\text{artwork}} = \frac{\sum_{t=1}^{T} \text{Belief}_{\text{topic}}}{N_{\text{topics}}}$$

where W_j is the rating of the artwork and N_{topics} is the number of topics.

In other words, the rating of an artwork propagates a belief value to all topics that are directly linked to this artwork and likely to some semantically related topics. The belief value of each topic is used, in turn, to determine the belief value for artworks.

Step 3. The user may give a rating to either recommended artworks or topics and this is collected as user feedback on the recommendations in the same scale to refine the recommendations presented.

The use of common vocabularies makes it possible to infer additional artworks and topics via semantic properties such as *vra:creator, vra:creationSite* and *vra:materialMedium* [17]. Following the content-based recommendation strategy, we allow for the enlargement of the recommendation scope through meaningful links. Also, it is partially helpful for solving the cold-start and sparsity problems. Even with a limited amount of ratings, the demonstrator still may produce recommendations through the semantic relationships and order them based on the belief value. For example, if the user rates the artwork "The Nightwatch" with 5 stars, the artwork "The Sampling Officials" and the topics "Rembrandt van Rijn" and "Lastman, Pieter" will be recommended. The underlying inference is that "The Nightwatch" has a creator "Rembrandt van Rijn", who also painted "The Sampling Officials", and he has the *student-of* relationship with "Lastman, Pieter". The rich semantic relationships offer explanations for users to understand

[16] http://www.vraweb.org/resources/datastandards/vracore3/categories.html.
[17] http://dublincore.org/.

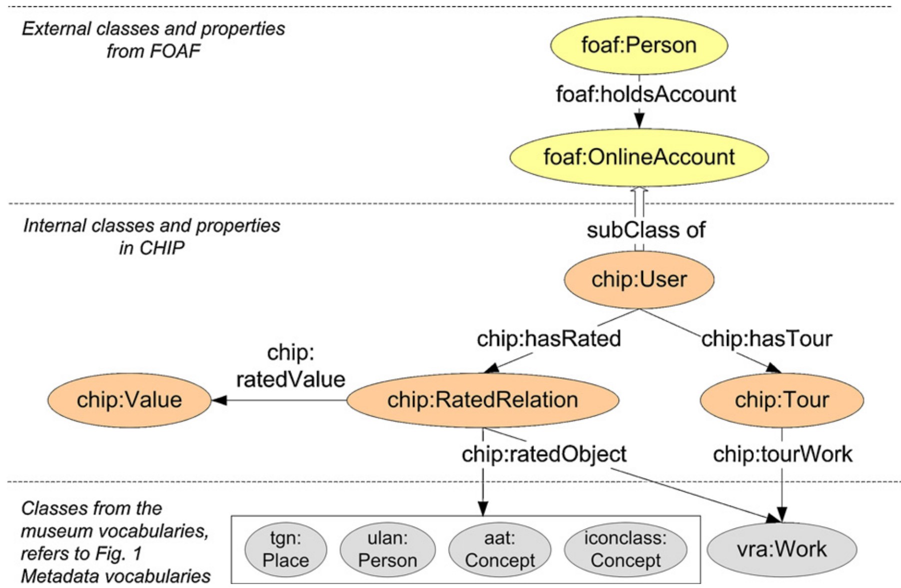

Fig. 2. Main classes and properties in the CHIP User Model.

why a recommendation is produced. By allowing users to rate recommended artworks and topics, it enables a fast rate-recommend loop for refining the user's preferences and increasing the accuracy of recommendations.

Besides the semantic-driven recommendation based on content, we have explored various approaches to address the cold-start and sparsity problems. By consulting museum domain experts, we present users a subset of artworks containing representative topics to rate first in the rating dialog. In such a way, the user profile collects user ratings with well-balanced distributed topics in a short time and make it possible to quickly generate recommendations through the entire collection.

As an example of distributed user data integration, we have mapped a small set of iCITY[18] user tags to CHIP art topics. The result of this experiment [18] suggests that the user tags may be used to populate the user model in CHIP and enable instant generation of recommendations. However, as we discussed in [19], this approach depends heavily on the correctness of the mappings. Another constraint is that the user tags are mostly seen as a stream of concepts that can be interpreted in various of ways, where the museum vocabularies are static.

5. A user model specification

Our goal of building a user model in CHIP is to provide a shared and common understanding of user information and behaviors for enhancing the personalized access to museum collections. Ideally, the user profile needs to store: (i) user's personal information; (ii) objects that the user has interacted with; (iii) user's activities over the objects (e.g. the user *rates* an object with a value); and (iv) the corresponding contextual information such as time, place and device. All these data allow us to get information of the user in context.

Currently, we have built a minimal user model as a specialization of FOAF[19]. Main classes and properties from FOAF used in CHIP are *foaf:Person* and *foaf:holdsAccount*.

- Class: *foaf:Person* is used to represent the information about a person who holds an account *chip:User* on a Web site. Account specific information is described by *chip:User*, a subclass of *foaf:OnlineAccount*.
- Property: *foaf:holdsAccount* is used to link a *foaf:Person* to a *chip:User*.

The core class in the user model is the *RatedRelation*. It uses the definition of semantic N-ary relations[20] to represent additional attributes describing a relation. For example, Saskia rates artwork "Nightwatch" with a value of 5. This rate relation contains information in the original three arguments: who has rated (Saskia), what is rated (Nightwatch), and what value the rating gives. Each of the three arguments in the original N-ary relation gives rise to a true binary relationship. In this case, there are three properties: *hasRated, ratedObject* and *ratedValue*, as shown in Fig. 2. The additional labels on the links indicate the OWL restrictions on the properties. We define both *ratedObject* and *ratedValue* as functional properties, thus requiring that each instance of *RatedRelation* has exactly one value for Object and one value for Value.

There are in total 5 classes in the range of ratedObject property: *vra:Work, ulan:Person, tgn:Place, aat:Concept* and *ic:Concept*. These objects are well-defined with properties in Fig. 1 Metadata vocabularies in CHIP RDF Schema. In the definition of the *User* class (of which the individual Saskia is an instance), we specify a property *hasRated* with the range restriction going to the *RatedRelation* class (of which *RatedRelation_1* is an instance). In addition, we have defined the *Tour* class and two related properties: *hasTour* and *tourWork*. The range of *tourWork* is the class *vra:Work*.

Further extension of this specification would require more indepth treatment of contextual information (e.g. device, time, location) and how this is linked to user activities, such as rating an artwork or creating a tour. In addition, also observational data, e.g. artworks visited, time spent with artworks, could be useful to collect, and may possibly be used to increase recommendation efficiency, effectiveness and relevance. For example, does recording the time spent with an artwork, allow us to infer an actual preference for that artwork, even it is not included in the tour or not rated?

[18] http://icity.di.unito.it/.
[19] http://www.foaf-project.org/.
[20] http://www.w3.org/TR/swbp-n-aryRelations/.

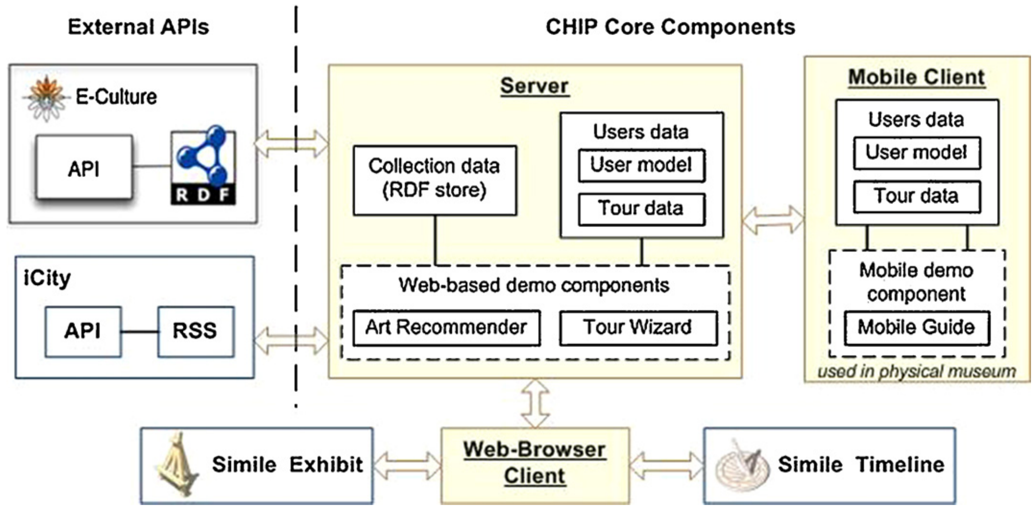

Fig. 3. CHIP overall architecture.

If we know where a user has been, when visiting a city, does this allow us to infer a consistent interest in particular topics?

6. Architecture and implementation

Fig. 3 shows the core CHIP components, third-party open APIs, which deliver semantic search results in CHIP (E-Culture API) or additional user data (iCity API) and tools that CHIP uses for data visualization.

The server-side CHIP core components are described below:

- *Collection data* refers to the enriched artwork collection, currently the Rijksmuseum ARIA database, maintained in a Sesame Open RDF memory store and queried with SeRQL.
- *User data* contains user models stored in OWL and tour data stored in XML. To be used by the Mobile Tour Guide, the user models currently have to be transformed to XML.
- *Web-based components* are an Art Recommender and a Museum Tour Wizard realized as Java Servlets and JSP pages with CSS and JavaScript.

Another CHIP client, implemented on a PDA (MS Windows Mobile OS) contains a standalone application Mobile Guide. It is an RFID-reader-enabled device and could also work offline inside the museum and subsequently be synchronized with the server-side on demand. The user profile and the tour data (both in XML) can be downloaded from the CHIP server to the mobile device to be used during the tour in the museum. When the museum tour is finished, the user data can be synchronized with the user profile on the server.

Fig. 4 presents the details with respect to the usage of the E-Culture API for semantic search in CHIP. Each user query in CHIP is sent to the E-Culture server, which sends a JSON file back with a list of artworks related to the search query. For every artwork we get a score (relevance of the search result) and a path (search path in the graph). We then further process the JSON file and add more CHIP-specific information to each artwork, like concepts that are associated with this artwork (from the collection data) and the artwork rating (from the users data). The resulting CHIP JSON file is sent to Simile Exhibit tool to be presented in a faceted view.

In order to experiment with user tag interoperability between the CHIP demonstrator and third party applications, we have adopted an open API to request and link user data from iCity using RSS feed. Once the user's personal (login) information is authenticated in a dialog between iCity and CHIP, we map the iCity user tags to the CHIP vocabulary set (ARIA shared with Getty and Iconclass) [18], by using the SKOS Core Mapping Vocabulary specification.

7. Usage scenario

In this section we describe a typical usage scenario of the CHIP demonstrator in order to illustrate the main user–system interactions.

Saskia is planning her first-time visit to the Rijksmuseum Amsterdam. She does not know a lot about the collection and she would not be able to spend much time there either. Here is how the CHIP demonstrator could help her:

- finding out what she likes in the Rijksmuseum collection;

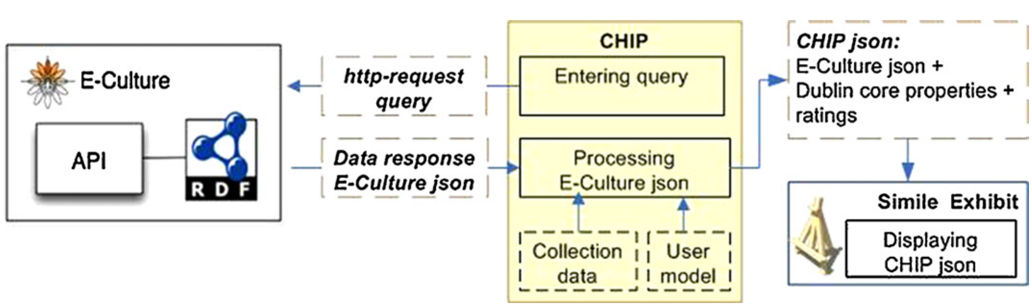

Fig. 4. Application of E-Culture API in CHIP.

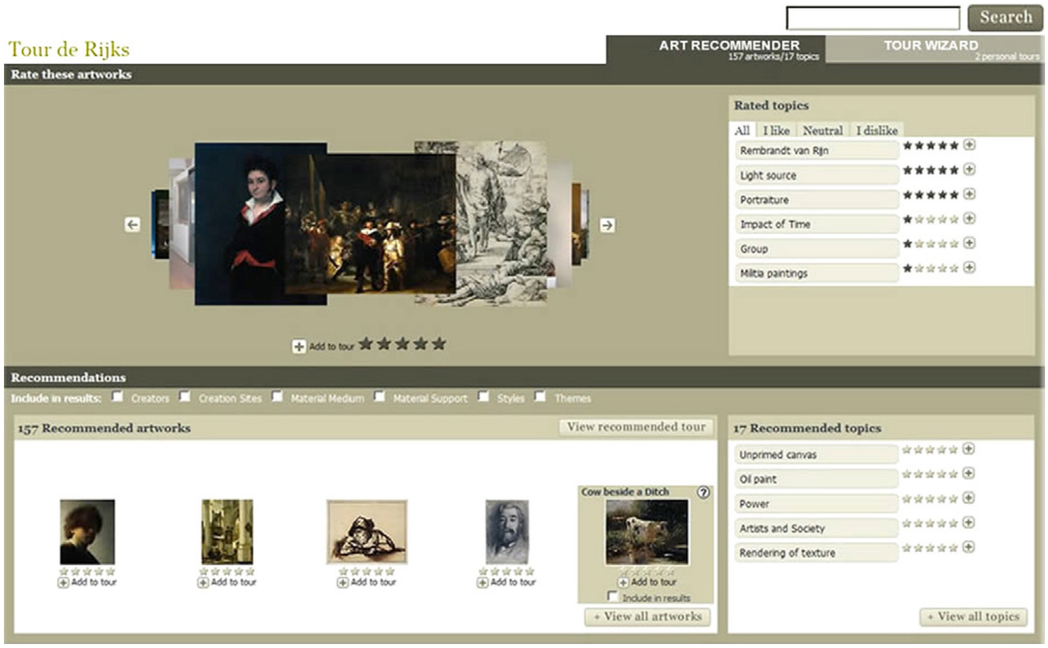

Fig. 5. Screenshot of Art Recommender.

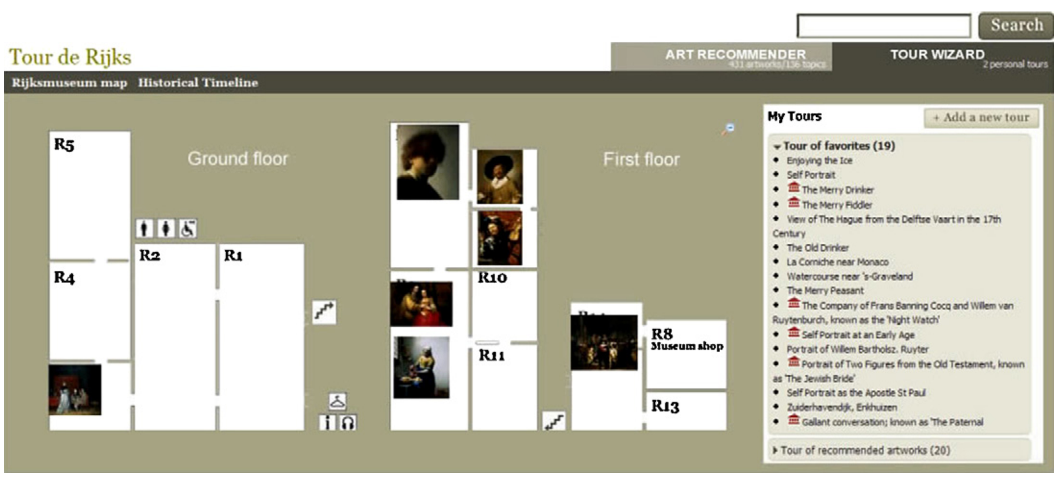

Fig. 6. Screenshot of Museum Tour Wizard.

- preparing a personalized museum tour (in terms of time to spend and number of artworks to see);
- storing the data of her visit so that she can later on use it.

To login on the CHIP online demonstrator Saskia needs to create a user account. Once logged in, she can choose either the *Art Recommender* tab, to quickly get acquainted with the Rijksmuseum collection and find out her art interests, or she can choose the *Tour Wizard* tab to create different personalized tours and see their layout on the Rijksmuseum map or on a historical timeline. A general *Semantic Search* option supported with an autocompletion function is available, if she wants to search for artworks or topics.

Everywhere in the CHIP demonstrator Saskia can give a rating (in a *5-degree rating scale*) from 1 star (I hate it) to 5 stars (I like it very much) on an artwork or a topic presented on the screen. Each rating of an artwork results in: (i) directly including the artwork with the rating in her user profile, (ii) using the updated user profile to generate a *list of recommended artworks* and a *list of recommended topics*. For each recommended artwork or topic, Saskia can click on the *"why"* (see Fig. 5[21]) for an explanation. For recommended topics, "why" explains which artworks with this topic have been rated positively, and for recommended artworks, it explains which topics from these artworks have been rated positively. Also, Saskia can rate recommended artworks or topics and update her user profile for a further refinement of recommendations.

Based on the collected ratings from Saskia, the *Museum Tour Wizard* generates automatically two tours: "Tour of favorites" containing all her positively rated artworks and "Tour of recommended artworks" containing the top 20 recommended artworks. Saskia can explore the tours by viewing the artworks on a museum map (see Fig. 6) or on a historical timeline. She can also create new tours by using the search option for finding topics or artworks to add to the tour.

[21] The screenshots are based on the design by Fabrique (http://www.fabrique.nl/).

When Saskia is in the museum she can upload her tours on a PDA and use it for guidance. Artworks currently unavailable in the exhibition are filtered out, but are still to be seen on the PDA as background information [18]. For example, Saskia's tour of favorites consists of 15 artworks and is estimated to last for 75 min. But she wants to spend at the maximum one hour, so the Mobile Guide reduces her tour to 12 artworks. When she is ready to start, the Mobile Guide recommends her a sequence of artworks and a route to follow.

The usage scenario assumes that all artworks in the museum are tagged with RFID tags. During the tour, Saskia can request information about new artworks by using the RFID tag reader attached to the PDA, which plays an audio and provides an option to rate this artwork. After listening to the audio and rating the artwork, she follows the initial tour. When the tour is finished, Saskia may synchronize her updated user profile on the PDA with the user profile that was created earlier online. In this way, she has saved all her interactions in the museum and maintained an updated user profile online.

8. Evaluation

The overall rationale of the evaluation is to follow a user-centered design cycle in the construction of each part of the CHIP demonstrator. We have performed two initial evaluations at Rijksmuseum Amsterdam with real users to test particular aspects of the demonstrator and derive requirements for further development.

8.1. Evaluation I: effectiveness of recommendations, novices vs. experts

The goal of the first evaluation [10] is to test the effectiveness of the content-based recommendations with the CHIP Art Recommender. 39 Rijksmuseum visitors participated in this study with an observer. They used the CHIP Artwork Recommender in an average of 20 min. The knowledge of the users of the Rijksmuseum collection was tested with questionnaires before and after the test session with the CHIP demonstrator. Our hypothesis was:

The Art Recommender helps novices to elicit or clarify their art preferences from their implicit or unclear knowledge about the museum collection.

To test the hypothesis, we have compared the *precision* of user's topics of interest before and after using the Art Recommender (rating and getting recommendations) [10]. Looking at the wide variety of users, we defined an *expert-value* as a weighted sum of user's personal factors (e.g. prior knowledge of the museum collection, frequency of visiting the museum, interest in art) collected from the questionnaire to distinguish between novice and expert users. As reported in [10], the results confirmed our hypothesis, a significant increase of precision was found for novices, while there is a slight increase for experts. However, the distinction between novices and experts is not clear-cut. Plotting the precision on a continuous range of the expert value, we observed, ignoring extreme values, a convergence as expert level increases.

In addition, we have derived four dominant factors about the museum visitors target group. Most of the users appear to be:

- Small group with 2–4 persons and a male took the leading role (67%).
- Mid-age people in 30–60 years old (62%).
- No prior knowledge about the Rijksmuseum collections (62%).
- Strong interest in art (92%).

Table 2
Evaluation II: results in six groups.

Group	1	2	3	4	5	6
Sequence of artworks	R	R	E	E	E + S	E+S
Target of ratings	Ra	Ra + Rt	Ra	Ra + Rt	Ra	Ra + Rt
Number of user ratings	96	151	170	224	157	203
Match of preferences	24%	30%	45%	48%	49%	44%

From this, we get a clear image what are the characteristics of the main target users. The main questions in this context are: (i) what kind of interaction and personalization topics do we need for providing personalized access to the museum collection?; (ii) How to structure, store and use the user characteristics to refine the current user model?

8.2. Evaluation II: Representative samples for rating, sparsity and cold-start

The second evaluation was performed online with 63 participants, most of them are first-time users of the CHIP demonstrator. Based on a functionally enhanced CHIP Art Recommender, which allows to search for artworks and topics, we explored different alternatives for getting recommendations through the entire collection, to solve the *sparsity* and partially the *cold-start* problem. The evaluation consists of two parts: Part 1 is to let users assess 45 well-distributed topics and Part 2 is to randomly split users into six different groups to rate artworks and topics in a short time (limited to 5 min). These six groups follow different alternatives to build their user profiles according to two independent variables: (i) sequence of artworks, which are presented in the Art Recommender for users to rate; and (ii) target of ratings. These two variables ranged over the following values: *Sequence of artworks* (random, expert-sorted, expert-sorted + self-selected); and *Target of ratings* (rate artworks, rate artworks and topics). Here "expert-sorted" means that domain experts selected the first 20 artworks, which overall cover a well-balanced distribution of topics through the entire collection. After that, artworks appear in the order of the number of topics each contains. The "expert-sorted + self-selected″ condition allows to search for artworks and topics based on "expert-sorted". Table 2 gives an overview of the results according to the six groups using different approaches, where R (Random), E (Export-sorted), S (Self-selected), Ra (Rate artworks) and Rt (Rate topics).

The results show that: first, the "expert-sorted" sequence of artworks works very well for first-time users to quickly build their user profiles with well-distributed topics through the entire collection; and second, "rating both artwork and topics synchronously" increases the total amount of the user's contributions (ratings) and it seems to improve the precision of recommendations; however, at some moment, it might lead to information overload.

All in all, the two evaluations gave us some critical insights in: (i) how to further specify the target group and adapt the user interaction and interfaces for the main groups of users; (ii) how the sequence of artworks affects the recommendation relevance and ranking. Further we learned about the context in which the users are visiting the museum, e.g. in small groups of 2–4 persons, and usability issues of the mobile device.

9. Discussion and future work

In this article, we demonstrated how Semantic Web technologies are deployed in a realistic use case to provide personalized recommendations in the semantically enriched museum collection. The semantic enrichment provides relational and hierarchical structure which we further exploit in a combined artwork and topic based recommendations. The evaluation suggests that this approach helps especially novices to elicit their art preferences about the collection.

However, it also brings up a new problem with respect to calculating the recommendation relevance. For example, if the user rates an artwork, we currently treat all its properties, such as "creator", "creationSite" and "material" equally in the recommendation strategy, where they could carry different importance for each user. In other words, the "creator" could be more interesting to the user than the "material". Moreover, material is likely to be a less discriminative factor for recommendations, as most of the artworks in this collection are of the same material. Thus, each artwork property should be assigned with a different weight in the recommendation strategy. Even more, the relevance of each property for a given user should be dynamically adjusted according to the user's ratings, or used with a default value when not enough user ratings are available. If a user mostly rates values of the property of "creationSite", these should have a priority in recommendations. To solve this problem, we are now looking for strategies to define a dynamic weight for properties when calculating the *Belief value* of an artwork and topics for recommendations.

Web 2.0 enjoys increasing popularity and offers a rich network with a large number of user communities and a staggering amount of user generated content. For recommender systems this suggests, as a main opportunity, the integration of distributed user data for recommendations. Such integration would amount to a unified user model that can be used across multiple applications, enriching the potential for recommendations by using the distributed user data. However, to realize such a user model, issues of storage, linking, representation and inference must be solved.

As a first step of defining such a user model specification, we proposed to extend the existing FOAF specification with possibilities to express user activities and interests in objects. Moreover, as observed in [20], Web 2.0 is a user-centered community, whereas the Semantic Web must be regarded as primarily a network connecting professional data through semantic relations. When we extrapolate this observation to our approach in CHIP, the major challenge is not to linking data from social networks and other Web 2.0 applications, but to bridge the gap between the semantic structure of museum collection data, which is professional semantics, and the variety of meanings found in open social networks, which rely on what is commonly called emergent semantics. The direction of bridging this semantic gap, as suggested by [21], is to add structure to user data, as a function of how this data links to repositories of information. One way of creating such a structure, as proposed for SIOC in [22], is to characterize social networks not as relations between people, but rather as object centered sociality. Objects could simultaneously be characterized by semantically linked meta data, obtained from professionals. Admittedly, this is still a long way from collective intelligence [21], but it is likely a significant step towards providing better recommendations, that take the users social context into account.

Acknowledgement

The CHIP (Cultural Heritage Information Personalization) project is funded by the Dutch Science Foundation funded program CATCH[22](Continuous Access to Cultural Heritage) in the Netherlands.

References

[1] E. Hyvönen, E. Mäkelä, M. Salminen, A. Valo, K. Viljanen, S. Saarela, M. Junnila, S. Kettula, MUSEUMFINLAND—Finnish Museums on the Semantic Web, Journal of Web Semantics 3 (2) (2005).
[2] M. van Gendt, A. Isaac, L. van der Meij, S. Schlobach, Semantic web techniques for multiple views on heterogeneous collections: a case study, in: Proceedings of the European Conference on Research and Advanced Technology for Digital Libraries (ECDL'06), 2006.
[3] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, Journal of IEEE Transactions on Knowledge and Data Engineering 17 (6) (2005).
[4] P. Brusilovsky, A. Kobsa, W. Nejdl, The Adaptive Web, Methods and Strategies of Web Personalization, 2007, pp. 291–324.
[5] J.P. Bowen, S. Filippini-Fantoni, Personalization and the web from a museum perspective, in: Proceedings of the Museums and the Web Conference (MW'04), 2004.
[6] G. Schreiber et al., MultimediaN E-Culture demonstrator, in: Proceedings of the International Semantic Web Conference (ISWC'06), 2006.
[7] K. Höök, J. Karlgren, A. Wærn, N. Dahlbäck, C.G. Jansson, K. Karlgren, B. Lemaire, A glass box approach to adaptive hypermedia, Journal of User Modeling and User-Adapted Interaction (UMUAI) 6 (2–3) (1996).
[8] S. Berkovsky, T. Kuflik, F. Ricci, Cross-domain mediation in collaborative filtering, in: Proceedings of the User Modeling Conference (UM'07), 2007.
[9] T. Ruotsalo, E. Hyvönen, An event-based approach for semantic metadata interoperability, in: Proceedings of the International Semantic Web Conference, 2007.
[10] Y. Wang, L. Aroyo, N. Stash, L. Rutledge, Interactive user modeling for personalized access to museum collections: the Rijksmuseum case study, in: Proceedings of the User Modeling Conference (UM'07), 2007.
[11] M.R. Zakaria, A. Moore, H. Ashman, C.D. Stewart, T.J. Brailsford, The hybrid model for adaptive educational hypermedia, in: Proceedings of the Adaptive Hypermedia Conference (AH'02), 2002.
[12] R. Denaux, V. Dimitrova, L. Aroyo, Integrating open user modeling and learning content management for the semantic web, in: Proceedings of the User Modeling Conference (UM'05), 2005.
[13] R. Albertoni, A. Bertone, M. De Martino, Semantic web and information visualization, in: Proceedings of Semantic Web Application and Perspective, 2003.
[14] H. Cramer, et al. The effects of transparency on trust and acceptance in interaction with a content-based art recommender. Journal of User Modelling and User-Adapted Interaction (UMUAI), 4560 (2008) 324–333.
[15] M. van Assem, M. Menken, G. Schreiber, J. Wielemaker, B. Wielinga, A method for converting thesauri to RDF/OWL, in: Proceedings of the International Semantic Web Conference (ISWC'04), 2004, pp. 17–31.
[16] L. Aroyo, N. Stash, Y. Wang, P. Gorgels, L. Rutledge, Chip demonstrator: semanticsdriven recommendations and museum tour generation, in: Proceedings of the International Semantic Web Conference (ISWC'07), 2007.
[17] Y. Wang, R. van Sambeek, Y. Schuurmans, L. Aroyo, N. Stash, L. Rutledge, P. Gorgels, Be your own curator with the CHIP tour wizard, in: Proceedings of the Museums and the Web Conference (MW'08), 2008.
[18] Y. Wang, F. Cena, F. Carmagnola, O. Cortassa, C. Gena, N. Stash, L. Aroyo, RSS-based interoperability for user adaptive systems, in: Proceedings of the Adaptive Hypermedia Conference (AH'08), 2008.
[19] F. Carmagnola, F. Cena, O. Cortassa, C. Gena, Y. Wang, N. Stash, L. Aroyo, Tag interoperability in cultural web-based applications, in: Proceedings of the Hypertext and Hypermedia Conference (Hypertext'08), 2008.
[20] M. Greaves, P. Mika, Semantic Web and Web 2.0, Journal of Web Semantics 6 (1) (2008) 1–3.
[21] T. Gruber, Collective knowledge systems: where the social web meets the semantic web, Journal of Web Semantics 6 (1) (2008).
[22] U. Bojars, J.G. Breslin, A. Finn, Using the semantic web for linking and reusing data across Web 2.0 communities, Journal of Web Semantics 6 (1) (2008).

[22] http://www.nwo.nl/catch.