

Mashups for Semantic User Profiles

Riddhiman Ghosh, Mohamed Dekhil

Hewlett-Packard Laboratories

1501 Page Mill Rd., Palo Alto, CA

{riddhiman.ghosh, mohamed.dekhil}@hp.com

ABSTRACT

In this paper, we discuss challenges and provide solutions for capturing and maintaining accurate models of user profiles using semantic web technologies, by aggregating and sharing distributed fragments of user profile information spread over multiple services. Our framework for profile management allows for evolvable, extensible and expressive user profiles. We have implemented a prototype, targeting the retail domain, on the HP Labs Retail Store Assistant.

Categories and Subject Descriptors

H.3.4, H.3.5 [Information Storage and Retrieval]: Systems and Software – *User profiles and alert services*; Online Information Services – *Data Sharing*.

General Terms

Design, Management.

Keywords

User profiles, information management, semantic web, personalization

1. INTRODUCTION

Whether online on the Internet, or offline in brick-and-mortar environments, personalization of services, content and user interactions is seen as key to a superior customer experience. This requires service providers to build and maintain accurate models of a customer's preferences, interests, background, etc., i.e. a user profile. However to obtain better insight into customers and build more holistic user profiles it is not sufficient for the service provider to only look at *its* history and interactions with the customer, but look at “out-of-band” profile information—fragments of the user's interactions or preferences with other services, distributed perhaps across multiple web sites or service touch-points. For example, consider a retailer who would like to personalize offers/coupons for its customers. A large portion of the factors affecting a customer's purchase decisions occur outside the scope of the brick-and-mortar store or online shopping website. The “out-of-band” information in this case, which has a significant impact on what, when and how customers buy, include customers' social/life events, data from online calendars, online shopping lists/wish lists, customer personal information management systems, social network recommendations etc. And yet retailers largely only rely on demographic identifiers and past shopping history / point-of-sale data because building this holistic view of a customer is extremely hard given that the sources of this out-of-band information are diverse, typically cross-organizationally distributed and with differing schemas and semantics hindering aggregation and reuse.

Just as current Web 2.0 mashups allow for the creation of hybrid applications, we argue that user profiles can be significantly enriched

by building them through the mashup of profile fragments distributed across services. Our goal in this paper thus is to focus on user profiles not from a statistical prediction perspective, but from a data management one. Our desiderata for a profile management framework are as follows: to support expressive and extensible profile representations and query; enable aggregation of profile data from multiple heterogeneous data sources; and support distributed profile storage, specifically on user mobile devices. We have applied our ideas to the HP Labs Retail Store Assistant (RSA) [1], which is a kiosk and online service-based platform designed to enhance and create a personalized shopping experience. In order to *know* the customer and realize the different use cases we have envisaged, our profile management framework needs to capture and maintain the types of information as depicted in Figure 1 (which shows the different categories of user information as mentioned in [2] applied to the retail domain).

2. PROFILE DATA MANAGEMENT

In our *Semantic User Profile management framework (SUPER)* we address the problems mentioned above—of data integration from disparate sources, of semantic interoperability, of sharing profile data between organizations—by using semantic web technologies (RDF, OWL) to help us model profile data, make its meaning explicit through shared ontologies and integrate with other data sources.

2.1 Extensible Profile Representation

All information about a user in SUPER is expressed and stored in the form of RDF triples in a triple store. User profile create/read/update/delete operations are all based on the manipulation or SPARQL-based query of these RDF triples. There are several advantages to this approach: First, a user's information model becomes easily *extensible*, as the structure of RDF subject-predicate-object graphs makes it amenable to be readily combined with other graphs, perhaps RDF assertions made by external sources. The profile model is *post-hoc* able to store types of information we had not originally anticipated. Also, user profile data, such as in Figure 1, can be semi-structured, sparse or often too heterogeneous to fit neatly into relational database tables. RDF is designed to handle such types of information well.

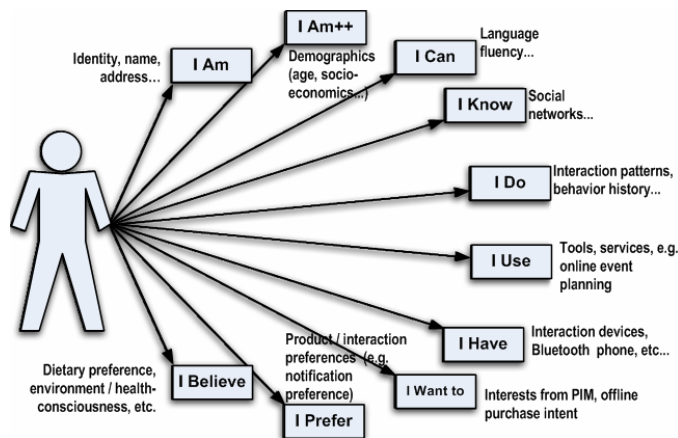


Figure 1: Types of profile information captured

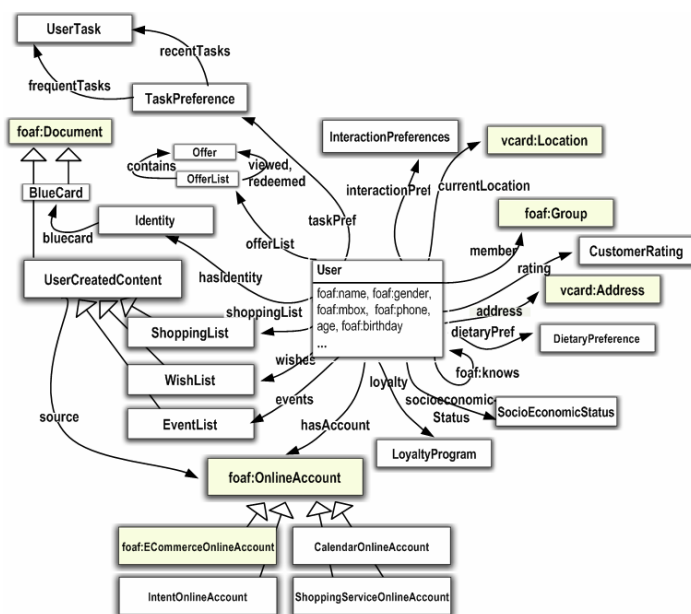


Figure 2: Partial snapshot of Retail User Profile Ontology

2.2 Retail User Profile Ontology

We have designed an OWL ontology to explicitly define the semantics of a user profile targeted towards the retail domain—the Retail User Profile Ontology (RUPO). It uses shared concepts from several existing ontologies such as FOAF and vCard, and uses concepts related to the retail domain from the National Retail Federation’s Association for Retail Technology Standards’ IXRetail specification; it is a ‘living’ entity and is being revised based on our ongoing experiences with retail solutions. A simplified snapshot of the RUPO in UML-like notation is represented in Figure 2. The linkages in the ontology of a user’s profile to user-created content such as their wish lists and shopping lists are strengthened by the increasing emergence of Web 2.0 services that export their data through APIs (e.g. using the RDF-based RSS, or Atom formats).

2.3 Profile in your Pocket

The SUPER framework supports distributed and mobile profiles, based on our previous work [3]—users can carry their profiles in their pockets via their cell phones or PDAs. Generally, user profiles can be stored on the service-side (users don’t have enough control), or entirely user-side (disadvantage being the information is not specific or rich enough to be useful to different services). We believe in a combination of these two approaches, with users using their mobile device to assert general preferences and information about themselves, which can then be combined with service-specific user profiles that are maintained at the service end. We have been using RUPO, and similar to [4] the FOAF (friend-of-a-friend) ontology, to express these user-authored profiles.

3. PROOF OF CONCEPT

We have built a prototype of the SUPER profile management framework to validate our approach, as shown in Figure 3. Our implementation focused on one of the primary use cases in retail scenarios—a user walking up to the RSA kiosk in a store, identifying himself and expecting a list of offers/coupons customized for him based on his aggregate profile. Multiple data sources (representing sources internal and external to a retailer) were used. We assume the user has *a priori* informed SUPER which sources it is allowed to use. Calendar/event information can be a valuable source of insight into what a customer wants to buy or

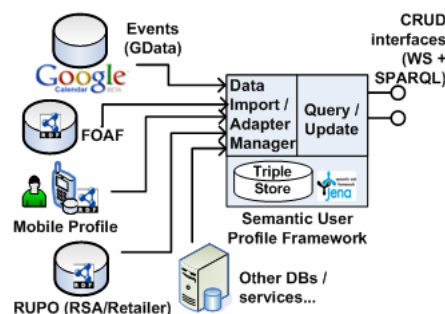


Figure 3: Overview of SUPER framework

is interested in, e.g. throwing a party, or attending a birthday, going on vacation etc. We used Google Calendar, a popular online personal information/event management system that provides REST-style APIs (called GData) that export calendar information.

An RDF calendar vocabulary adapter was used. Similarly FOAF is a widely popular RDF format for users to describe themselves, their interests and also their social networks. We use FOAF as a source since using this data can inform user–service interactions with not only general profile information about a user, but also holds the potential for us to infer what a user’s interests or preferences are based on those of the people in his social network. We also used profile information from a user’s cell phone. (This allows for ‘anonymous personalization’—personalization can be offered by services with which the user has no established interaction history, enabled by RUPO which makes meaning of profile data explicit. Moreover the user can adopt different *personas* while interacting with different services by changing the profile on his phone). We also used simple “shopping list” information available from a user’s RUPO document. The source of this is the RSA ‘intent capture’ system we have built, but could also be another online service, such as Amazon wish-lists. The Jena semantic web toolkit from HP Labs was used in this prototype.

4. DISCUSSION

We described a framework to enable the creation and management of extensible and expressive user profiles, and use traditionally “out-of-band” data systems to enrich user information models. This differs from other profile services such as Microsoft Live ID [5] (cannot easily combine with external data sources) or the Liberty Alliance personal profile service [6] (has different goals and does not address profile semantics, capture or distribution). Other efforts such as Universal Profiling Interoperability [2] (recommendations for data interchange between domains) or the Connected Services Framework [7] are not targeted towards the retail domain, do not contribute to a domain ontology or tie to personal information management systems to enrich user profiles. Our next steps include making profiles adaptive based on the evolution of profile-linked user-created content over time, and further investigating use of FOAF/RUPO social network information for personalization use cases.

5. REFERENCES

- [1] Hewlett-Packard Press Release, HP shows off system that affords every customer a personal shopper, May 2007
- [2] Houben, J., et al., State of the art: semantic interoperability for distributed user profiles, Telematica Institut Report, 2005.
- [3] Ghosh, R., Dekhil, M., “I, me and my phone: identity and personalization using mobile devices”, HP Labs Technical Report HPL-2007-184, November 2007.
- [4] Ankolekar, A., Vrandecic, D., Personalizing web surfing with semantically enriched personal profiles, in *Proceedings of the Semantic Web Personalization Workshop*, Budva, 2006.
- [5] Windows Live ID, accounts.services.passport.net/
- [6] Liberty Alliance, projectliberty.org/
- [7] Microsoft Connected Services Framework, microsoft.com/serviceproviders/solutions/connectedservicesframework.k.msp