The People's Portal: Ontology Management on Community Portals

Anna V. Zhdanova

DERI – Digital Enterprise Research Institute,
University of Innsbruck, Austria, and
National University of Ireland at Galway, Ireland
www.deri.org

anna.zhdanova@deri.org

Abstract. The Semantic Web will never be a reality unless the quantity of Semantic Web pages significantly increases. Our thesis is that enlargement of the Semantic Web content is possible and effective by acquisition of ontological structures from human users of community Semantic Web portals. An important aspect of the proposed approach is that ontology structure acquisition from humans takes place merely by the portal exploitation without significant effort from the user side. We describe the People's portal environment that motivates and allows the People's portal users to develop, populate and share ontologies, define the ways to manage the content on the community Semantic Web portals, and to reach dynamic consensus on the basis of heterogeneous ontologies. This work contributes to overcoming the limitations of the existing web portals by specifying and prototyping the People's portal environment, i.e., a framework for user-driven community Semantic Web portals.

1 Introduction

Nowadays, numerous community web portals related to business or leisure have been created, and many community web portals have proved to be highly popular and successful by acquiring millions of members [O'Murchu04]. However, the existing community web portals are rather inflexible in specification of member profiles, the portal content, the ways this content is internally structured and delivered to the portal members. The existing community web portals declaratively specify the data the user may provide to the portal and find there. The specification comes from the web portal creators and their view of the domain, which is usually comprehensive, but is definitely limited, and thus, makes the portal out of interest for the users when they want to go beyond this view.

However, a larger degree of the portal's flexibility and adaptation to the portal's members' real demands can be achieved by bringing the Semantic Web technologies and tools [Berners-Lee01] to the existing community web portals. To support this thesis, an experimental environment that comprises an

ontology management environment and a community Semantic Web portal, namely the People's portal, is built. The functions of the People's portal are to allow the Web developers and members creating and reusing the web content and the Semantic Web content by constructing, populating and using the People's portal. Also, the People's portal is aimed to provide the means to manage the undesired consequences of access to ontology editing and population by non-professional ontology engineers.

This paper is structured as follows. An example of a motivation scenario is provided in Section 2 in order to illustrate the key idea of our approach. The related works are outlined and compared with the proposed environment in Section 3. The ontologies employed at the People's portal environment, their data, interoperation and conceptual levels are described in Section 4. The user views of the environment are summarized in Section 5. The main architecture modules of the People's portal environment are presented in Section 6. Section 7 concludes the paper.

2 Motivation Scenario

Although the scenario described in this section is a very specific, it should be viewed as an instance of a general pattern of any information exchange that can take place at any Web portal.

The scenario is as follows: John and Mary are registered at a community web portal in a social networking domain, where they have their profile information (such as name, surname, e-mail addresses, hobbies, etc). Assume that the web portal does not have an ontology attribute (slot in the profile form) to allow the users specify their phone numbers, but John and Mary want to exchange the phone numbers between each other.

On the current portals the scenario usually runs as: John and Mary will have to use e-mail or an instant messenger or some other available mediums, but no longer the portal and its functionality. Further, if all the communities of John and Mary exchange all non-semantically annotated data by e-mail or any other chosen medium, the community members will obviously suffer from the numerous e-mails they get and/or overload with interfering requests for information and huge workflows of irrelevant data.

On the People's portal the scenario includes the following steps: Using the ontology management functionalities integrated into the Semantic Web portal environment, John goes to the community portal and creates a new concept "phone number" in a community ontology. Then he fills his phone numbers in his profile (he can do this now, since the concept is introduced by him). Mary reuses the concept introduced by John and fills in her phone numbers in her user profile. Thus, John and Mary can exchange information (e.g., easily find each other's telephone numbers on request or have them delivered by default) using the Semantic Web portal functionalities, without the need to involve any external mediums.

The obvious shortcoming of the proposed approach to the users' unlimited access to the content ontology editing options is that the ontologies of the Semantic Web community portal run into risk of becoming badly structured, far too large and redundant to support the activities of the community and an individual efficiently. Handling this problem by making the ontological structure adapted to the human communication formalisms by organizing an efficient way to operate and render ontological data is a challenge for the practical use of our approach. In practice, this challenge will be addressed in our future work.

3 Related Work

The related works lie in the following areas: Semantic Web portals, collaborative ontology management and knowledge acquisition, ontology alignment, personalization and community support.

Ontology development and editing policies are quite simple on most of the current Seamntic Web portals [Stollberg04]: ordinary portal users do not participate in construction of ontologies, though they often can introduce their ontology instances (e.g., as in KnowledgeWeb and Esperonto Semantic Web portals based on ODESeW [Corcho03]). Exceptionally, the users can propose changes to ontology structure, but these changes need to be approved by the main ontology editor [Pinto04]. Obviously, this approach to ontology development and editing is not dynamic, does not consider heterogeneity, personalization and community aspects, is not scalable, and thus can not serve as a basis for organization of an effective communication process. Though the People's portal environment supports functions that are typical for Semantic Web portals in general, it is different, because of allowing the portal members to specify knowledge representation issues of their Semantic Web portal, and thus, develop their own portal themselves.

In analogy with FOAF project, the People's portal environment provides means (similar to foaf-a-matic) to create semantic annotations on people's personal details or other portal content the portal members might want to bring in. The specifics of the People's portal environment is that its users actually produce machine readable pages to make use of the portal, whereas FOAF project approach focuses on the promotion and improvement of a specific ontology, but not on the FOAF ontology application, usage and dynamic user-driven evolution. Meanwhile, recent research has shown effectiveness of knowledge acquisition from web users, and the same research also brought understanding that in order to be a success knowledge acquisition applications need to move out from the game and toy area and be tightly integrated with applications that are of actual use to the community [Chklovski03].

In comparison to Wiki and Open Directory Project approaches, where "netizens" are encouraged to bring structured knowledge on the web, the People's portal environment aims at reaching more semantic granularity in specifying the portal content. The People's portal environment provides the means for collaborative development of ontologies. However, it is different from environments for explicit web-based collaborative ontology development [Domingue98] [Farquhar97], which resulted to be of limited practical usage. The People's portal environment makes the users involved in creation, extension and reuse of ontologies implicitly in order to increase the value of the portal.

Personalization is another field that aims at making applications more useful. Personalization is traditionally defined as the ability to customize each individual user's experience of electronic content [McCarthy01]. The known areas of personalization application are:

- handling different sources of content
- arrangement of content on a screen
- delivery mechanisms ("push" vs. "pull")
- delivery vehicles (web browser, mobile phone, etc.).

The objective of personalization for the purpose of delivery of personalized information is fairly straightforward. It is to deliver information that is relevant to an individual or a group of individuals in

the format and layout specified and in time intervals specified [Won02]. While personalization was applied extensively on the ordinary Web portals for the individual users (especially in eCommerce area) [Aggarwal02; Instone04; Schiaffino04], the studies for community and consensus aspects of personalization in the Semantic Web context are still lacking.

The community issues are currently usually studied with respect to computing communities by means of clustering based techniques and identifying the communities to which pages belong [Greco04]. For the issues of individual personalization issues, there are no solution frameworks that support these issues extensively. The state of the art is mainly in establishing the theoretical basics for the further work on the application level, e.g., developing languages, such as a view language that picks up the unique situation of data in the Semantic Web and allows easy selection, customization and integration of Semantic Web content [Volz03].

4 Ontologies and Ontology Layering

We distinguish three main levels and six ontology types in the specification and prototype of a community Semantic Web portal ontology submodule in the People's portal environment. The proposed classification simplifies the environment development and allows introducing similar editing and storage policies for the ontologies and data that are assigned to the same level.

Levels of the Semantic Web portal environment:

1. User level – user profile and personalization data specified according to ontologies of the community level.
2. Community level – ontologies and rules associated with a community, used and evolved by the community.
3. Portal level – ontologies and rules for cross-community information exchange, that also support inter-portal integration and communication.

Ontology types:

1. User profile ontologies – ontologies that specify the content of the portral. For example, if the main content of the community Semantic Web portal is data about people, the user profile ontology data are person's first name, last name, phone numbers, hobbies, etc.
2. User personalization ontologies-- ontologies that specify how user profile ontologies and user profile data are delivered to the individual user. These personalization ontologies can be Semantic Web portal specific. The personalization ontology data can also specify which user profile ontology concepts are instantiated by the user and which content and content links the user wants to share and which not.
3. Community profile ontologies – ontologies that specify community data, such as lists of the members of this community, their general anonymous interests and preferences.
4. Community personalization ontologies – ontologies that specify how and which Web portal content is delivered to a community. These personalization ontologies can be Semantic Web portal specific. The personalization ontology data can also specify which content and content links the community wants to share and which not.

5. Portal profile ontologies – ontologies that specify mappings and data transfer protocols across community and user ontologies. These ontologies define both mapping within ontologies (helping to reach consensus at the data level: example of a problem taken from Instone [Instone04]: "if users can specify they are interested in "PlayStation 2" but the information about the product is tagged "PS2" there will be gaps in the personalization") and also specific inter-community ontology mappings (helping to reach consensus at the metadata level: problem of the type "she uses FOAF, he uses VCard").
6. Portal personalization ontologies – ontologies that specify inter-portal mappings (helping to reach consensus at the physical level: problem of the type "she is on Friendster, he is on Orkut").

The six ontology types and assignment of the ontology types and instance data to the three levels of the Semantic Web portal are shown at Fig. 1.

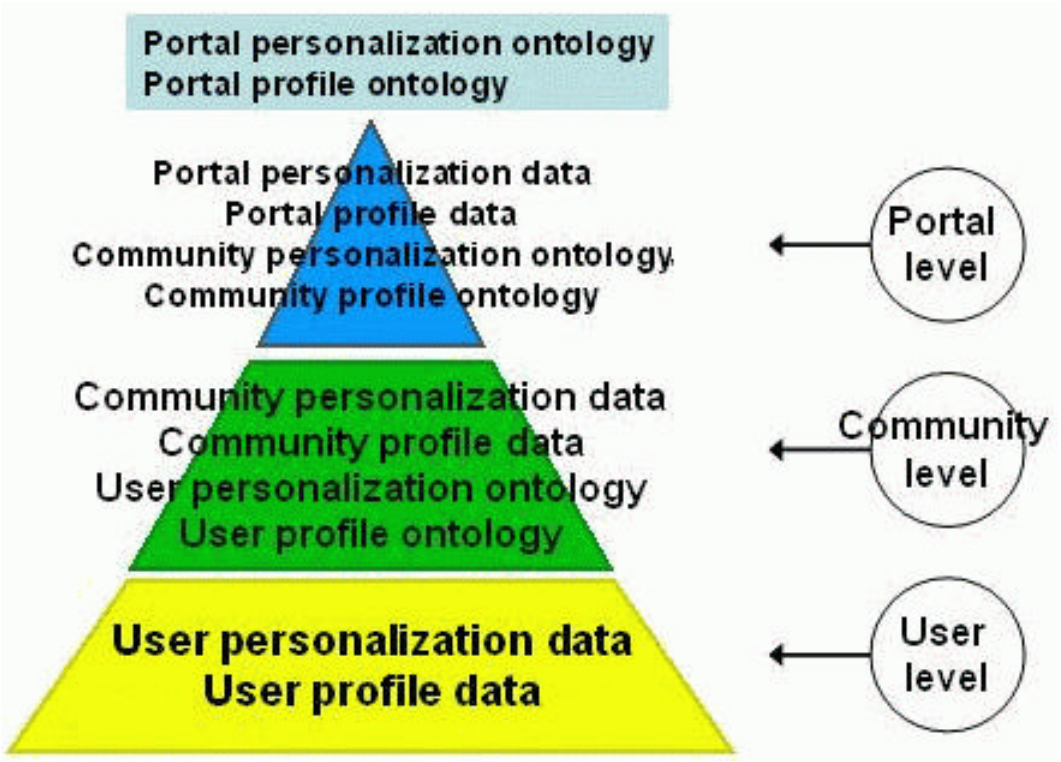

Figure 1: Layering ontologies and instance data

5 Views

The implementation of the People's portal supports four portal views in the user interface. The user interface is written in JSP and HTML and can be accessed with an ordinary web browser. The four views are as follows:

1. Content view displays content brought by the portal to the human user by request or default. This view is typical for many existing portals and may display such information as the latest economy news, weather forecast for tomorrow and scores from yesterday's football games.

2. Access to the user profile. This view gives an opportunity to modify user profile ontology and user profile data.

3. Access to the user personalization data. This view gives an opportunity to modify user personalization data. At the early stage of the prototype testing, an access to this view will be given to web-portal administrator only, to avoid involvement of the portal user in too many new features at once.

4. Access to portal's profile. This view gives an opportunity to modify portal profile data, specifically, introduce inter-community mappings. External modules for semi-automatic mapping can be integrated in the portal to help the user to find and map ontology items.

Each of the four views is rendered on separate web pages that include a menu which allows switching between the views.

6 Modules

The People's portal environment incorporates the following modules:

1. Ontology management module
   This module allows the user to develop and instantiate ontologies providing an access to the user profile, user personalization data and the portal's profile. In the People's portal prototype, this module is built employing Java and JSP as programming languages, Jena [Jena] as a Semantic Web toolkit and languages of XML/RDF/OWL family for knowledge representation. The ontology management module runs on the Tomcat server [Tomcat] and currently allows the users to extend lightweight ontologies and populate these ontologies with instances. The module can be accessed through an ordinary web browser such as Internet Explorer.

2. Publishing/delivery/personalization module
   This module presents the information collected by the portal to the user, i.e., this module generates the content view of the portal. Ideally, having a Semantic Web approach in mind, the work of this module should be based on employing personalization ontologies as described in Section 3. Employing a customized content management system suitable for web-site construction is also possible for generation of content views from Semantic Web. However, the existent content management systems that do not process ontology data pose certain limitations on the use of Semantic Web content. An example of such a limitation is the need to eventually convert Semantic Web content in a non-semantic knowledge representation format (e.g., XML file) that is supported by a typical content management system.

3. Ontology alignment support module
   An external software library for ontology alignment will be integrated in the People's portal framework. The criteria for choosing the ontology alignment library are:
   - the library is written in Java
   - the library supports alignment of ontologies represented in XML/RDF/OWL languages

- the library is preferably based on the Jena ontology model
- the library is lightweight
- the library's API is understandable and preferably well documented
- the library is open source.

4. Interportal and intercommunity integration support module
   This module provides an access to the portals' profile view to the portal members and interoperates with ontology alignment support module.

7 Conclusion

The resulting principles and specifications are purposed to form the basis of user-oriented community Semantic Web portals of the future. The prototyped software will be further elaborated to comprise the next generation of ubiquitous web applications aimed at effective publishing, access and efficient management of personal and community information taking into account personalization issues and community support.

Acknowledgements

The work is funded by the European Commission under the projects DIP, Knowledge Web, Ontoweb, SEKT, SWWS, Esperonto, COG and h-TechSight; by Science Foundation Ireland under the DERI-Lion project; and by the Vienna city government under the CoOperate programme. The author would like to thank Andreas Harth, Stefan Decker and all the members of the Semantic Web portal working group for their advises and inputs to this work.

References

[Aggarwal02] Aggarwal, C., Philip S. Yu, P. S., 2002. An Automated System for Web Portal Personalization. VLDB 2002, Hong Kong, China, pp. 1031-1040.

[Berners-Lee01] Berners-Lee, T., Hendler, J., Lassila, O., 2001. The Semantic Web. Scientific American 284(5), pp. 34-43.

[Corcho03] Corcho, O., Gomez-Perez, A., Lopez-Cima, A., Lopez-Garcia, V., Suarez-Figueroa, M., 2003. ODESeW. Automatic Generation of Knowledge Portals for Intranets and Extranets. In: Fensel, D. et al. (Eds.), Proceedings of the Second International Semantic Web Conference; Springer, LNCS 2870, pp. 802-817.

[Chklovski03] Chklovski, T., 2003. LEARNER: A System for Acquiring Commonsense Knowledge by Analogy. In Proceedings of Second International Conference on Knowledge Capture (K-CAP 2003). October 2003.

[Domingue98] Domingue, J., 1998. Tadzebao and WebOnto: Discussing, Browsing, and Editing Ontologies on the Web. 11th Knowledge Acquisition for Knowledge-Based Systems Workshop, April 18th-23rd. Banff, Canada.

[Farquhar97] Farquhar, A., Fikes, R., Rice, J., 1997. The Ontolingua Server: Tool for Collaborative Ontology

Construction. Int. J. Human-Computer Studies, 46(6), pp. 707-728.

[Greco04] Greco, G., Greco, S., Zumpano, E., 2004. Web Communities: Models and Algorithms. World Wide Web: Internet and Web Information Systems, 7, pp. 59-82.

[Instone04] Instone, K., 2004. An Information Architecture Perspective on Personalization. In: Karat, C.-M. et al. (Eds.), Designing Personalized User Experiences in eCommerce, the Netherlands, Kluwer.

[Jena] Jena: http://jena.sourceforge.net.

[McCarthy01] McCarthy, J. F., 2001. The Virtual World Gets Physical: Perspectives on Personalization. IEEE Internet Computing 5(6), pp. 48-53.

[O'Murchu04] O'Murchu, I., Breslin, J.G., Decker, S., 2004. Online Social and Business Networking Communities. To appear in Proceedings of Workshop on Application of Semantic Web Technologies to Web Communities, ECAI 2004.

[Pinto04] Pinto, S., Staab, S., Sure, Y., Tempich, C., 2004. OntoEdit Empowering SWAP: a Case Study in Supporting DIstributed, Loosely-Controlled and evolvIng Engineering of oNTologies (DILIGENT). 1st European Semantic Web Symposium (ESWS 2004), May 2004, Heraklion, Crete, Greece. Springer, LNCS 3053, pp. 16-30.

[Schiaffino04] Schiaffino, S., Amandi, A., 2004. User-interface agent interaction: personalization issues. Int. J. Human-Computer Studies 60, pp. 129-148.

[Stollberg04] Stollberg, M., Lausen, H., Lara, R., Ding, Y., Sung-Kook, H., Fensel. D., 2004. Towards Semantic Web Portals. In Proceedings of WWW2004 Workshop on Application Design, Development and Implementation Issues in the Semantic Web, 18 May 2004, New York, USA.

[Tomcat] Apache Tomcat: http://jakarta.apache.org/tomcat/index.html.

[Volz03] Volz, R., Oberle, D., Studer, R., 2003. Implementing Views for Light-Weight Web Ontologies. 7th International Database Engineering and Applications Symposium (IDEAS 2003), 16-18 July 2003, Hong Kong, China. IEEE Computer Society 2003, pp. 160-169.

[Won02] Won, K., 2002. Personalization: Definition, Status, and Challenges Ahead. J. of Object Technology (1) 1, pp. 29-40.