

# People's Community Semantic Web Portal and Metaportal

Semantic Web Portal Project

Anna V. Zhdanova

April 2004

**Deliverable:**  
D14

**Version:**  
0.1

**Date:**  
15-Apr -04

## **DERI Ireland**

University Road

Galway

IRELAND

[www.deri.ie](http://www.deri.ie)

## **DERI Austria**

Technikerstrasse 13

A-6020 Innsbruck

## Table of Contents

<b>0 SUMMARY .....</b>	<b>1</b>
<b>1 INTRODUCTION .....</b>	<b>2</b>
<b>2 DEFINITIONS OF WEB P ORTAL TYPES.....</b>	<b>4</b>
<b>3 PEOPLE'S PORTAL DESCRIPTION IN A NUTSHEL L.....</b>	<b>5</b>
3.1 COMMUNITY SEMANTIC WEB PORTALAND COMMUNITY SEMANTIC WEB METAPORTAL.....	5
3.2 DOMAINS .....	5
3.3 WHAT THE PEOPLE'S PORTAL IS NOT (BUT IS RELATED TO).....	5
<b>4 PEOPLE'S PORTAL FUNCTIONS .....</b>	<b>7</b>
4.1 AUTHENTICATION.....	7
4.2 PERSONAL PROFILE/ANNOTATION EDITING .....	7
4.3 PEOPLE'S PORTAL ONTOLOGY EDITING.....	8
4.4 SEARCH .....	9
4.5 VERSIONING.....	9
<b>5 ONTOLOGIES CAN BE EDITED BY EVERYBODY, AND THIS EDITING MAKES SENSE: MYTH OR REALITY?.....</b>	<b>11</b>
5.1 ONTOLOGIES CAN BE E DITED BY EVERYBODY: HOW TO HARNESS DEMOCRACY?.....	11
5.2 ONTOLOGIES CAN BE E DITED BY EVERYBODY: HOW TO HELP DEMOCRACY? .....	12
5.2.1 <i>One Social Role – One Ontology Class, or Personality Split</i> .....	13
5.2.2 <i>Possession of a Social Role of a Subclass is Sufficient for Possession of a Social Role of the Class</i> .....	14
<b>6 TECHNICAL DETAILS OF THE PEOPLE'S PORTAL IMPLEMENTATION .....</b>	<b>17</b>
6.1 INTEROPERABILITY BETWEEN THE USER, PORTAL AND METAPORTAL.....	17
6.2 DATA STORAGE AND ACCESS.....	18
6.3 NAMESPACES .....	19
6.4 DISCOVERY ASPECT .....	19
<b>7 OVERCOMING THE SW CHALLENGES .....</b>	<b>20</b>
<b>8 CONCLUSIONS .....</b>	<b>21</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>22</b>
<b>REFERENCES .....</b>	<b>22</b>



## 0 Summary

This document describes the People's portal that comprises a community Semantic Web portal in a social networking domain and a community Semantic Web metaportal. The community Semantic Web metaportal allows the People's portal end users to define the content structure (i.e., develop ontologies), populate ontologies and define the ways the content is managed on the community Semantic Web portal of the People's portal. The users are motivated to develop and populate the own portal ontology in order to use the People's portal functionalities that employ the ontology and its instances.

The People's portal is being prototyped as one of the DERI SWPP case studies.



## 1 Introduction

Nowadays, a multitude of related to business or leisure community web portals has been created [O'Murchu04], and many community web portals have proved to be highly popular and successful by acquiring millions of members. However, the existing community web portals are rather inflexible when it comes to specification of user profiles, the content of the portals, the ways this content is organized, search options. Nowadays, the existing community web portals simply specify what and how the users can contribute there and search about. The specification comes from the web portal creators and their view of the domain, which is normally comprehensive, but is definitely limited, and thus, makes the portal out of interest for the users after they deplete this view. However, a far larger degree of the portal's flexibility and adaptation to the portal's member's real demands can be achieved by bringing the Semantic Web technologies to the existing community web portals. Basically, the intelligent application of Semantic Web technologies would allow the portal's members to specify what their (informational) demands are and how to fulfill them.

The People's portal presented in this paper is about letting the Web developers and users to create the Semantic Web content through developing, populating and using the People's portal, and thus becoming the Semantic Web developers and users. Among the key factors of the People's portal prototype success in actual production and use of the Semantic Web content is its involvement of social networking portal successful features that are good at motivating many individuals to bring lots of data at one place. Among the key ideas behind the People's portal is that the Semantic Web is more likely to come true if everybody is provided with means and motivation to weave the Semantic Web, in a similar way as some time ago everybody was provided with means and motivation to weave the Web. The mission of the People's portal is to provide the means and motivation to weave the Semantic Web for a large number of Web users.

### **Note on the lexicon used in this paper**

In this paper, general, depersonalized terms are used when talking about the portal's features. Namely, the general terms "schema" and "ontology" are equally applied in the same usage contexts, while it is implied that these terms denote also the specific ontology embodiment present at the People's portal, e.g., an RDFS instance describing



the structure of the People portal's data. Another case for this is when the term "web-forms" is used and specific web forms of the People's portal are implied, e.g. , the ones, which at the moment are built dynamically by Java servlets and can be viewed by Internet Explorer and many other browsers. The general, depersonalized terminology is employed in order to separate the theoretical solutions from the implementation features. For instance, referring to the portal's ontology as to RDFS instance would be correct from the implementation point of view for a certain moment. But as the portal is being used and extended, an additional expressivity may be required, and for this an extension from RDF/S encoded data to OWL/S encoded data may be performed. However, after such an extension or even an ultimate change of ontology language, most of the ontology features and construction principles would remain intact. Thus, the specifications will be still relevant as the terms do not become obsolete, what precisely is attained by using a general, depersonalized terminology as described above.

### **Structure of the paper**

The paper defines the People's portal (Sections 2, 3), describes the portal functionality traits (Section 4), architecture and implementation issues (Sections 6).

Reading about supported principles of the creation of the People's portal ontology (Section 5) improves understanding of the People's portal outlook.

Sections 6 and 7 on challenges and conclusions are only sketched and are to be filled in later.



## 2 Definitions of Web Portal Types

**Definition 1:** A web portal is web site that collects information for a group of users that have common interests.

**Definition 2:** A semantic web portal is web portal that is based on semantic web technologies.

**Definition 3:** A community semantic web portal is a semantic web portal that is maintained by a community of users.

**Definition 4:** A community semantic web metaportal is a community semantic web portal that is maintained by a community of users who have an interest to define and manage a web portal.



### **3 People's Portal Description in a Nutshell**

#### **3.1 Community Semantic Web Portal and Community Semantic Web Metaportal**

Two interconnected semantic web portals are being implemented:

- a) a community semantic web metaportal allows to describe and manage a community semantic web portal,
- b) the community semantic web portal that is described and managed by the community semantic web metaportal (a).

#### **3.2 Domains**

The community semantic web metaportal is domain neutral.

The community semantic web portal is in the domain of social networking. Specifically, the portal collects information about people. This domain was chosen due to its superiority to any other existing domain in the number of web portal members it can draw and in its general understandability.

The defined above community semantic web metaportal and community semantic web portal collecting information about people are also referred as "People's portal".

#### **3.3 What the People's Portal is not (but is Related to)**

It is NOT another social networking community portal, because it comprises a community Semantic Web metaportal that allows the Semantic Web portal members to specify knowledge representation issues of their Semantic Web portal, and thus,



provides more functions. But it also has functions that are typical for social networking community portals.

It is NOT another attempt to simply FOAF<sup>1</sup> people around, because it is application-centered, thus annotations *have to be* created to get involved in the portal. But it also provides means (similar to foaf-a-matic) to create semantic annotations on people's personal details and thus make machine readable homepages.

It is NOT another Chimaera<sup>2</sup>, because it is application centred, thus makes the users truly interested to create, extend and reuse ontologies (of their own portal). But it also provides means for collaborative development of ontologies.

It is NOT another Wiki, because it provides means to edit and populate the portal's ontologies. But it is also an environment where the users are totally responsible to populate and develop their own portal themselves.

It is NOT another Open Directory Project<sup>3</sup>, because it gives the users an access to and the right to manage a higher, ontology level. But it also provides means for bringing metadata on the web by the whole community of web users.

---

<sup>1</sup> The FOAF project: <http://www.foaf-project.org>

<sup>2</sup> Chimaera: <http://www.kslstanford.edu/software/chimaera>

<sup>3</sup> The Open Directory Project: <http://dmoz.org>





## 4 People's Portal Functions

### 4.1 Authentication

Two user authentication and data storage ways are being created for the People's portal. Namely:

1) ID + password login,

data storage on the portal

2) homepage URL login,

data storage on the portal and on the member's resource, changes/updates are valid only when these changes/updates are actually detected by the portal's crawler on the members resource.

Advantages of the way(2) are that:

- 1) creation of the user ID + password combination is not performed (so the user can not lose or forget his authentication details)
- 2) one can supplement/modify instance data without visiting the portal(s)
- 3) one can easily reuse of his/her instance data in other environments

Clearly, whether the first or the second authentication way should be used depends primarily on the way a portal member stores his/her instance data. If the portal member's instance data is stored at the web-portal only, then the ID-password authentication is possible only.

### 4.2 Personal Profile/Annotation Editing

The ontology instances (i.e., personal details) are expected to be added by all the portal's users who create their profiles and thus get annotated with the help of the People's portal. Naturally, all ontology instance additions proceed according to the specified earlier schemata.



Physically, annotation editing can be performed in two ways:

- editing on server,
- editing on client.

Editing on client is reasonable when schema editing/extension is not necessary. The user just modifies his/her personal details or adds some new and stores the modified annotation on his location. Then, the portal's crawlers reach the modified annotation and update the search index of the portal accordingly. However, editing on client can not be applied to the cases when the user chooses (or has no other opportunity than) to store his or her profile on the server or when the user wants to modify a commonly shared schema. In these cases, the user should refer to the portal's annotation and schema editing services.

### **4.3 People's Portal Ontology Editing**

The People's portal members are allowed to extend/edit the "Person" ontology that is used to search for people and produce personal annotations.

The considered ontology extensions:

- class extensions
- attribute extensions

The considered inputs for class extension:

- subclass of which class
- name
- synonyms for the name and the name and its synonyms in other languages

The considered inputs for attribute extension:

- attribute of which class
- name
- synonyms for the name and the name and its synonyms in other languages
- cardinality
- data type



The implementation of the People's portal starts with comprising the most primitive extensions of the listed above.

#### 4.4 Search

Search is very important, since it is one the core functionalities the People's portal provides. Employing ontology structure search and/combined with keyword multilingual search are applicable.

Why would people search for people:

- to really find an individual(s) with some specific properties
- not to find somebody in particular, but out of an internal reason, e.g., a demand for self-assessment or curiosity
- not to find somebody in particular, but out of an external reason, e.g., to get statistics or for a market study

Search will be discussed in more details in the next versions of this deliverable.

#### 4.5 Versioning

One known way to support versioning is to ask the validity period when the instance is defined. For example, the user who adds a document describing a call for papers for some conference, can specify that the document is valid until the data of the paper submission deadline. However, in the "Person" domain, the validity period is often unknown (e.g., a person normally does not know well in advance when precisely he/she changes an apartment or dies), or the user simply does not care to take an effort to communicate the validity period.



Stronger motivations than just “keeping the People’s portal up-to-date” should be evoked in the People’s portal members to keep their personal annotations up-to-date. It is important to make the convenience of the users’ life as much as possible dependent on the personal semantic annotations, e.g., by bringing them to use also different systems/portals/applications that are important for the users and need the up-to-dated personal annotations to function properly.

## **5 Ontologies can be Edited by Everybody, and this Editing Makes Sense: Myth or Reality?**

Our thesis is that Semantic Web users can acquire skills of collaborative ontology development, in a similar way the Web users have acquired skills to contribute to web portals, act collaboratively on Wikis and Blogs, organize their own files in folders in their file systems, on the intranets and Internet. However, the process of acquiring of these skills will take some time, and meanwhile uncontrolled unskilfulness may cause inconveniences in using the portal. In this section, ways to minimize the harmful side effects of unprofessional ontology development are discussed.

### **5.1 Ontologies can be Edited by Everybody: How to Harness Democracy?**

Here, three cheap ways to harness the democracy on the People's portal are described. The first implementation of the People's portal follows the first way.

#### **(1) Grant schema editing permissions to all portal's members.**

A permission to extend the base Person ontology is given to all the portal members. Totally democratic.

#### **(2) Spread schema editing permissions to verified portal's members.**

Initially, the permission to extend the base Person ontology is held by the portal's modifier only. Later on, the portal's modifier can grant a permission to modify an ontology to anybody whom s/he considers competent of ontology modifying. After this, anybody who has a permission to modify the ontology can grant the permission to anybody s/he considers competent of ontology modifying. Such an approach will guarantee that only people who are acknowledged of having enough ontology engineering expertise will be permitted to extend the ontology. Restricting the number of ontology modifiers in this way is inexpensive for the portal's maintenance, and favorable for the portal's ontology quality.



A similar approach is being successfully practiced in the Open Directory Project<sup>4</sup> where the submitted web-resources are being finally classified by human-editors to the appropriate web-directory within a hierarchy of directories. Being an editor of a directory means to review web-resources submitted for addition to this directory and add these resources to the directory in case they are indeed related to this directory. There, to become an editor of some web-directory, one has to verify his or her understanding of the principle of this directory filling by providing a set of links to web-resources that s/he considers being relevant for this directory to be evaluated by other editors who initially created the directory. Clearly, this approach of collaborative building of “intelligent” web get substantially extended in the People’s portal: now besides adding of instances, the task of ontology extension and semantic enrichment is also delegated to the web’s community.

**(3) Grant schema editing permissions to all portal’s members and delete unpopular concepts with the portal’s functions.**

The portal can maintain a function that checks if the ontology concepts created by users are reused and extended by the other users. If this is not the case, unpopularity of some concepts can result from the bad design/accessibility of these concepts. The function may propose to reconsider unpopular parts of the ontology or/and delete them.

**5.2 Ontologies can be Edited by Everybody: How to Help Democracy?**

Learning is easier by following examples [a link to supporting literature is to be found]. If a sample, starting ontology is already provided when the People’s portal is launched, the portal members can see how this ontology is constructed and they are likely to follow the construction pattern when further developing the ontology. However, to be consistent in the ontology example, the portal’s developers need to make the ontology design principles explicit. In this subsection the design principles of the People’s portal ontology “Person” are discussed.

---

<sup>4</sup> The Open Directory Project: <http://dmoz.org>



### 5.2.1 One Social Role– One Ontology Class, or Personality Split

The People's portal implementation employs Description Logic (DL) based ontology languages. Coming to ontology design, DL naturally brings an object-centered flavor to ontological data representation. On the other hand, the fact that one person usually has multiple social roles and leads multiple activities has been discovered a long time ago [a link to supporting literature is to be found]. For instance, a person can be at the same time a researcher at the university, a parent at home, a customer in the supermarket, etc. Thus, the object-centered flavor of DL and the opportunity to represent a person as a set of his/her social roles and activities provide a good grounding for creation of an ontology where the basic (root) class is "Person", and all its subclasses are social roles of "Person".

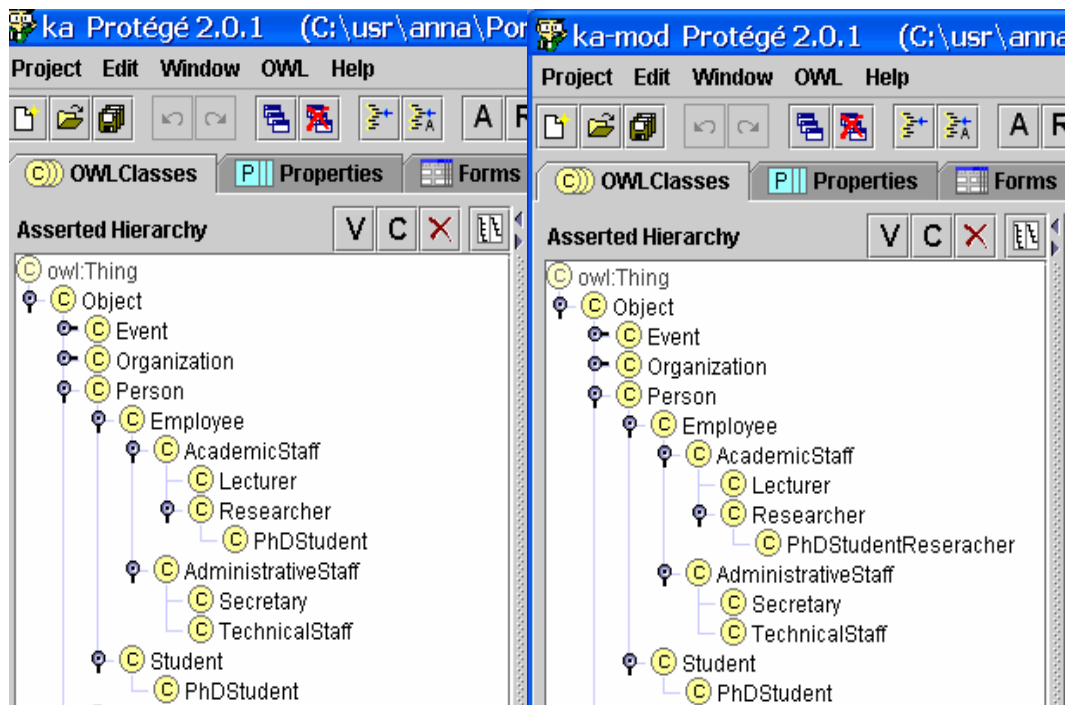
Examples of ontologies where "Person" is modeled with its social roles are shown at Figure 1. The left part of the Figure 1 shows how a concept "Person" is represented in [ka.owl](#) ontology that defines concepts from academic research (contributed by Ian Horrocks to Protégé Ontologies Library<sup>5</sup>). The ontology design principles of the People's portal ontology will generally propagate the principles that are illustrated by the example on the right side of Figure 1. The example on the right side of Figure 1 is a modified [ka.owl](#) ontology. In contrast to the [ka.owl](#) ontology construction principles, the People's portal ontology construction principles are focused on precise definitions of people's social roles. Thus, for instance, an example from the left side of Figure 1 does not follow the principle of People's portal "Person" ontology entirely, because the same class "PhD student" describes two different social roles a person may have. Being a "PhDStudent" as a subclass of "Researcher" means working as a researcher who does not have a PhD degree, but undertakes actions in order to get a PhD degree, also by being a student at the place of work or elsewhere. Being a "PhDStudent" as a subclass of "Student" means being a student who undertakes actions in order to get a PhD degree. Thus, the social role of the "PhDStudent" as a subclass of "Researcher" belongs to the employment domain, and the social role of the "PhDStudent" as a subclass of "Student" belongs to the education domain. Thus, to separate these roles and comply

---

<sup>5</sup> Protégé Ontologies Library: <http://protege.stanford.edu/ontologies/ontologies.html>



with the principles of ontology construction of the People's portal, a modification such as renaming "PhDStudent" that is a subclass of "Researcher" into "PhDStudentResearcher" can be introduced as shown on the right side of Figure 1.



**Figure 1. An example of ontology built regarding the personal role centered principle**

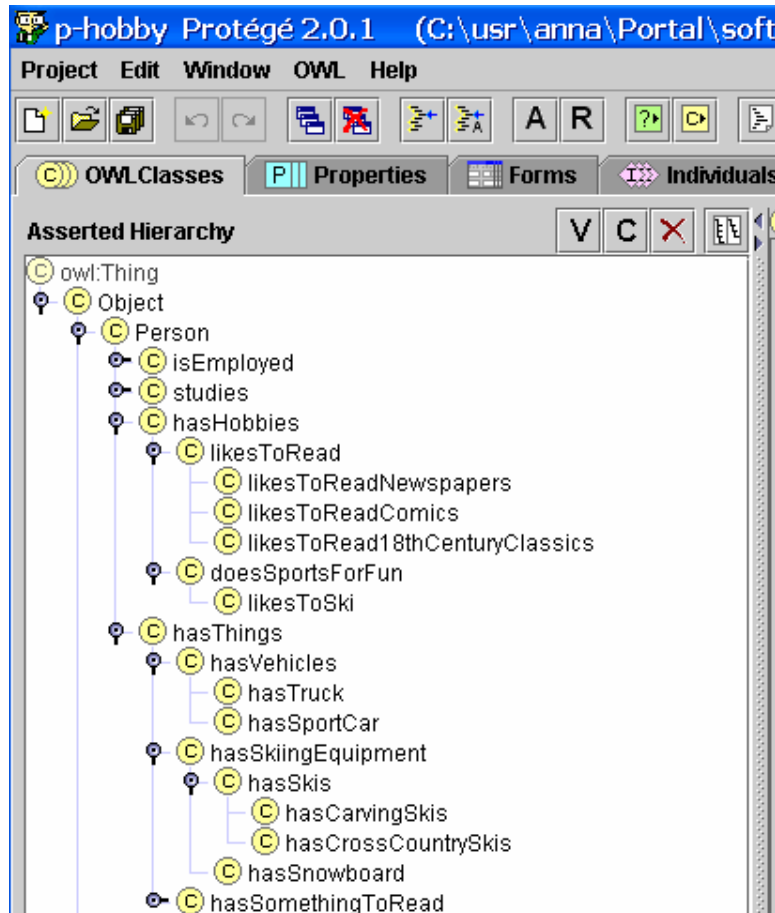
### **5.2.2 Possession of a Social Role of a Subclass is Sufficient for Possession of a Social Role of the Class**

In order to decide whether a class should be modeled as a subclass of another class, a guideline is to check whether possession of a social role described by a subclass is sufficient for possession of the social role described by the class. Figure 2 illustrates this principle. For instance, it is sufficient to be a person who "likesToSki" in order to have





“doesSportsForFun” and, continuing, “hasHobbies” as one’s social roles. However, the social role “hasCarvingSkis” is not sufficient to state that the owner of the skis actually skis or likes to ski. Thus, the classes describing “hasSkis” social roles are substantially separated from the class describing “likesToSki” social role, although they are both related to skis. If the goal was to build an ontology on ski equipment and its usage, probably, the concepts describing the facts that “someone has skis” and “someone likes to ski” would be located close to each other, but the goal is to build an ontology describing a person.



**Figure 2.** An example of the People’s portal “Person” ontology possible design

A major advantage of the proposed ontology design principles is simplicity of understanding of these principles. High understandability of how the ontology is built is



required, because the People's portal provides an opportunity to modify the basic "Person" ontology to all its members, inclusive the members with no ontology engineering skills. To build an ontology in the proposed way is just a little more difficult than to build a folder-structure which nowadays became a commonplace for all computer users.



## 6 Technical Details of the People's Portal Implementation

The technologies and tools employed in the implementation are mainly consistent with the ones described in deliverable 6 of DERI SWPP [Zhdanova04].

### 6.1 Interoperability between the User, Portal and Metaportal

The diagram on Figure 3 displays components that fulfill conversions of knowledge representation formats within the People's portal and interaction between these components.

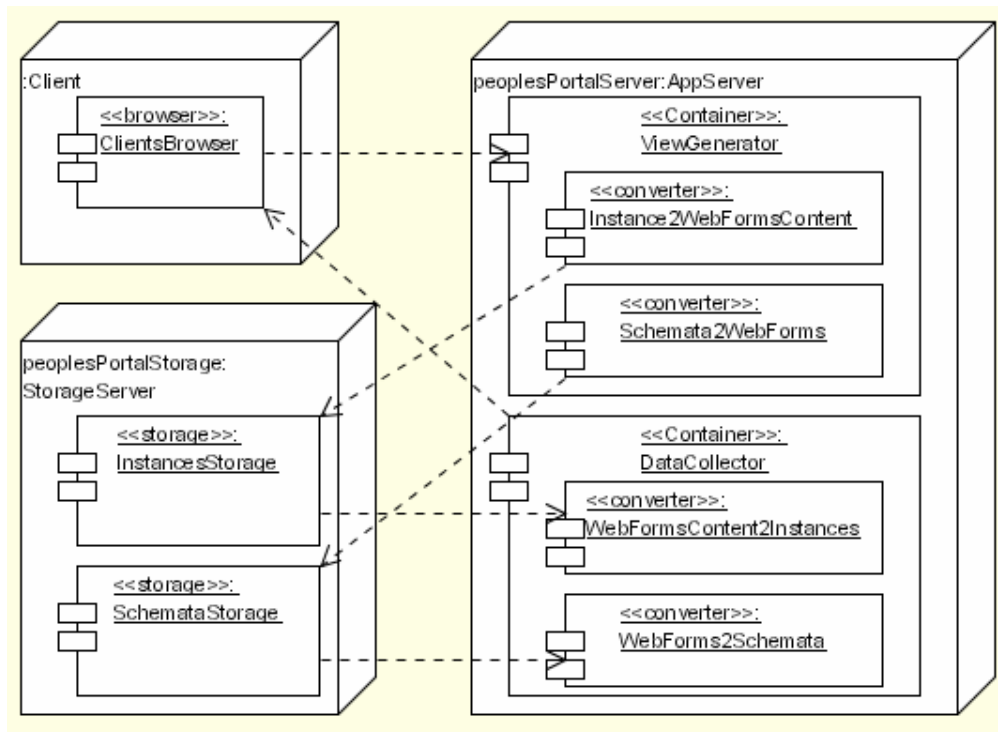


Figure 3 . Deployment Diagram for the People's Portal

## 6.2 Data Storage and Access

In the first prototype of the People's portal, the available data is stored in RDF and RDFS files, separating instances and schema. When performing web form generation, search, the data is extracted from these files. Because of scalability issues, an optimization of the data access in the search process will be required. Such future optimization is likely to involve introduction of databases (e.g., as done by Corcho et al. [Corcho03]) and domain specific search indices.

However, keeping the data in the form RDF and RDFS files has another advantage: simplified data sharing and integration with other systems, e.g., other Semantic Web portals. The advantages of such data representation are considered in the paper by Hyvönen et al. [Hyvönen04]. One of the practical problems considered in this paper was to access heterogeneous distributed databases of museums located throughout Finland. For the purpose of integration, the database tables were transformed to XML cards and XML cards were transformed to RDF cards. Beside the resolution of the integration issue, another advantage of such a transformation is semantic enrichment where new meaning is automatically added to the collection data (by discovering shared resources /URIs, relation inheritance and logical rules). On the other hand, reaching integration and interoperability in this way can not be done fully automatically due to unknown terms and complicated descriptions encountered in the databases and homonymous terms. An annotation system can identify such situations and point them out to a human editor that has to make the right decisions and corrections by hand, but involvement of handwork tends to make the solution costly. The rather typical situation leads to the conclusion that the process of data integration and system interoperation would be simplified, if all the involved parties used a common, semantically friendly way of data storage, but not heterogeneous semantics-deprived databases. Thus, storing ontologies in relational databases can be viewed as a step backwards in the technology evolution. The necessity to maintain a semantic version of the data anyway (e.g., such as RDF and RDFS data representation) along with a database storage means duplication of the effort spent at the data storage.



### **6.3 Namespaces**

Two types of namespaces are used, one to describe schema and the other to describe instances. The namespaces that name different schema concepts start with “www.deri.org/person” and end with respective extensions. The namespaces that name instance data of a particular user start with the user’s real homepage URL address or a fake address derived from the user’s ID on the portal and end with respective extensions identical to the endings of the namespaces of the first type.

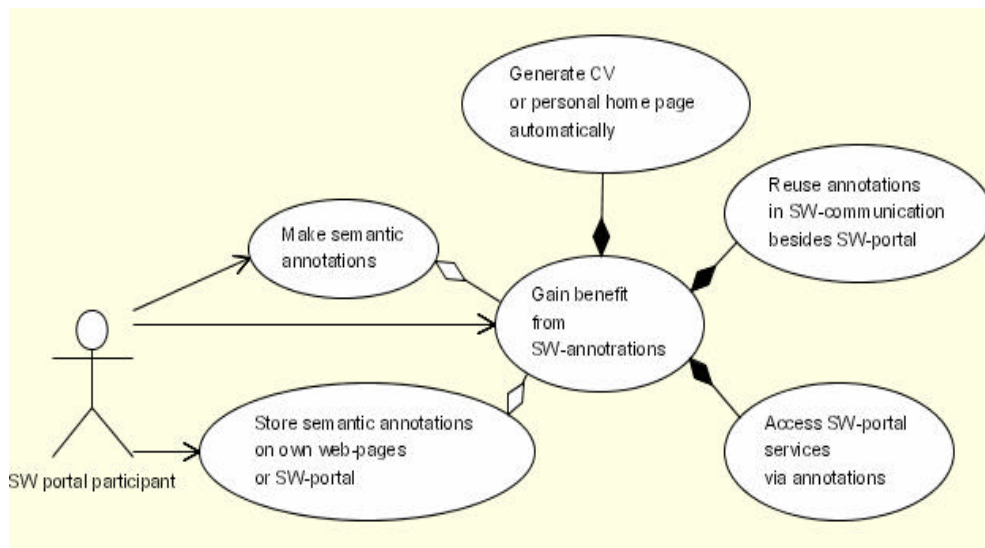
### **6.4 Discovery Aspect**

Discovery takes place when the users provide a link to the place where the personal annotations are stored or when the personal annotations are stored directly on the portal. Discovery of the externally created annotations is not an issue, because the People’s portal comprises the only relevant annotation service, and a report of all the changes in annotations and annotation schemas to the portal’s data storage is one of the side-effects of this service.



## 7 Overcoming the SW Challenges

- list the challenges
  - boost the number of SW pages
  - SW page reuse, add value to metadata, releases from data input overhead



**Figure 4 . Added Value of Metadata for the Human User**

- overcome limitations of existing portals
- bringing the whole world to enrich semantic representation, i.e. no more need for wordnet, since now it is created and extended by the people interested
- language problems – adding terms in synonyms section in any languages will work for the key-based multilingual search
- why earlier works failed to do it

## 8 Conclusions

popular domain.

presented user view, tech view

unusual metaportal approach with ontology extension, development by community of users for themselves, its application, sw is brought to reality thereof

overcame sw challenges

**future work is to get the thing running**



## Acknowledgements

The authors thank members of the DERI Semantic Web portal work group (<http://www.nextwebgeneration.org/projects/sw-portal/sw-portal.htm>) for creation of a stimulating research environment.

## References

[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.

[Hyvönen04] Hyvönen, E., Junnila, M., Kettula, S., Mäkelä, S., Saarela, S., Salminen, M., Syreeni, A., Valo, A., and Viljanen, K., 2004. MuseumFinland — Finnish Museums on the Semantic Web: User's Perspective. Proceedings of Museums and the Web 2004 (MW2004), Arlington, Virginia, USA, March 31 – April 3, 2004.

[O'Murchu04] O'Murchu, I., Breslin, J.G, Decker, S., 2004. Survey of Existing Community Portals, DERI SWPP, deliverable 9.

[Zhdanova04] Zhdanova, A.V., Zaremba, M., Bischof, L., Möller, K., 2004. Semantic Web portal – Grounding Technologies and Tools Evaluation, DERI SWPP, deliverable 6.

