SWAD-Europe deliverable 12.1.8: SWAD-E Demonstrators - Lessons Learnt

Project name:
Semantic Web Advanced Development for Europe (SWAD-Europe)
Project Number:
IST-2001-34732
Workpackage name:
12.1 Open Demonstrators
Workpackage description:
http://www.w3.org/2001/sw/Europe/plan/workpackages/live/esw-wp-12.1.html
Deliverable title:
SWAD-E Demonstrators - Lessons Learnt (demo_lessons_report)
URI:
http://www.w3.org/2001/SW/Europe/reports/demo-lessons-report/
Author:
Dave Reynolds, HP Laboratories, Bristol, UK
Steve Cayzer, HP Laboratories, Bristol, UK
Paul Shabajee, Graduate School of Education and ILRT, Bristol, UK
Damian Steer, (Contractor) HP Laboratories, Bristol, UK

Abstract:
This report provides an overall summary and update on the SWAD-E open demonstrators work package 12.1. The report serves three purposes. First, it provides a short summary of the two demonstrators produced within the work package. Second, it provides an update on changes in the first demonstrator, semantic blogging, since the time the demonstrator report was written. Third, it draws some overall conclusions from the experiences of developing the demonstrators.

Status:
First release

Comments on this document are welcome and should be sent to Dave Reynolds or to the public-esw@w3.org list. An archive of this list is available at http://lists.w3.org/Archives/Public/public-esw/

Contents

Executive Summary

The Semantic Web seeks to provide a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. The SWAD-Europe project [SWADE] provides targeted research, demonstrations and outreach to help semantic web technologies move into the mainstream of networked computing.

This report provides an introduction to the results of the *open demonstrators* work package. Within this work package we selected two applications designed to illustrate the nature of the semantic web and developed working and accessible demonstrations of both. For each demonstrator there are separate reports describing both the requirements and lessons learnt from the work. This report serves as a high level summary for that work and highlights general lessons that can be taken from the overall work package.

The first demonstrator, semantic blogging, takes the notion of weblogging or blogging as a lightweight publishing medium and enhance it by enabling authors to publish embedded structured information and metadata. As the test domain for the demonstrator we chose the problem of publishing bibliographic information and sharing it amongst small work-groups.

The second demonstrator, semantic portals, shows how to use the semantic web standards to create a rich community information portal. It shows how individual data providers can remain in control of their own data (thus helping with sustainability and reuse of the data) while still enabling other members of the community to enhance the data with additional classifications and relations. As a test domain for this demonstrator we constructed a directory of environmental and biodiversity organizations.

Whilst there were several specific lessons learnt from each individual demonstrator the overall lessons that we highlight in this report are:

- Decentralization and integration of distributed information are indeed key benefits of the semantic web which can be compelling in the right circumstances. However, benefits of such decentralization (sustainability, scaling, timely information updates) does carry risks to be aware of (consistency, trust and user interface issues).
- The reuse of ontologies and thesauri in such applications is not straightforward for a mix of technical and user

reasons. In particular, in applications with a substantial user-facing aspect then the appropriateness and navigability of the ontologies to the users context can out weigh the benefits of direct integration and reuse.
- Existing tool solutions are perfectly adequate to the types of demonstrator described here though higher level tools for producing data creation and data visualization interfaces would be welcome. Some of the software developed for the demonstrators, targeted at these issues, is designed to be a reusable resource for the community in its own right.
- Specific, grounded, demonstrators are indeed a valuable outreach aid which help communicate not just the nature of the semantic web but its practical utility.

In conclusion, we feel this has been beneficial exercise all round and that the demonstrators themselves, the collaborations which have emerged and the spin-off software components all have a future beyond the completion of the SWAD-E project.

1 Introduction

This report is part of ⊣ SWAD-Europe ⊣ Work package 12.1: Open demonstrators. This workpackage covers the selection and development of two demonstration applications designed to both illustrate the nature of the semantic web and to explore issues involved in developing substantial semantic web applications.

This report serves three purposes.

First, and foremost, it includes short summaries of the two demonstrators developed in this work package along with some background on why they were chosen. This should be sufficient information to allow the reader to decide whether the demonstrators are of interest to them and, if so, whether to read the more detailed *requirements* and *lessons* reports for each demonstrator.

Second, the two demonstrators were developed in sequence so there has been some gap between the completion of the first demonstrator (semantic blogging) and the writing of this final report. We take the opportunity to note changes and updates that have occurred related to this demonstrator since the time of its initial report.

Third, we draw some overall lessons from the development of the two demonstrators. For each of the two demonstrators there is already a "lessons learnt" report giving the detailed lessons. So in this summary report we restrict ourselves to some higher level observations that span the demonstrators.

2 Background and objectives

2.1 SWAD-E Objectives

The Semantic Web seeks to provide a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. Much of the base work for this endeavour has already been laid. We have languages supporting exchange of data (RDF [⊣ RDF-LANGUAGE] backed by a model theory [⊣ RDF-MODEL]) and exchange of conceptual models or ontologies [⊣ OWL]. We have a number of mature software frameworks [⊣ JENA, ⊣ REDLAND] and some commercial activity [⊣ NETWORK-INFERENCE, ⊣ SEMAVIEW]. Yet there is still much misunderstanding and resistance, antagonism even, from the developer and wider community [⊣ SHIRKY-SYLLOGISTIC]. Therefore, while much foundational work remains to be done, there is an equally important need for education and outreach. It is with this need in mind that the Semantic Web Advanced Development Europe project, SWAD-E, [⊣ SWADE] was established to provide targeted research, demonstrations and outreach to help Semantic Web technologies move into the mainstream of networked computing.

2.2 Semantic web

In our application survey [⊣ SWADE-ANALYSIS] we characterised the semantic web as an attempt to turn the web from a large hyperlinked book into a large interlinked database. Viewed like this, we identified the characteristics which seemed to capture the value of a semantic web approach:

- **Data representation**. The foundation of the semantic web is a common format, RDF [⊣ RDF-LANGUAGE], to represent data. This format is designed for representing semi-structured data and metadata. Data is broken down into conjunctions of individual assertions in the form of subject/predicate/object triples. Each of the components (other than simple literals) is a web URI and thus has a defined place in the global namespace. This allows many sorts of data (property values of objects, relationships between objects, value annotations) to be represented uniformly and allows data from multiple locations to be combined without accidental clashing of property names or structure mismatches.
- **Semantics.** The aspiration of the semantic web is to be able to express meaning. The schema [⊣ RDF-SCHEMA] and ontology [⊣ OWL] layers of the semantic web begin to do this; for example, enable one to say whether two terms are distinct, equivalent or whether one term is a subset of another. This capability allows a data source to expose its conceptual model explicitly in machine processable form thus providing the basis for more sophisticated processing and better integration of data from different sources. The semantic web does not require some standardized global upper ontology to function, remaining decentralized so that data sources are free to mix and match terms from different ontologies.
- **Webness.** The critical innovation of the semantic web is to put both these values into a web framework. This is manifested in deceptively simple ways such as the use of URIs to provide a global namespace for both entities and concepts (properties, types). However, the impact is substantial. An agent accessing a data source now has a means for discovering the ontology associated with that data source. Ontologies can be developed in a decentralized way to suit particular needs, but the terms defined in different ontologies can be related and combined to enable transformation of data from one domain to another.

The aim of the open demonstrators (work package 12.1) was to construct two demonstration applications that would help to communicate the nature of the semantic web and the practical issues involved in applying it. They needed to illustrate the purpose of the semantic web (the exchange and reuse of data across boundaries) and to bring out these three core characteristics (semi-structured data representation, semantics, webness). The applications needed to be of a small enough scale to be accomplished within the limits of the work package yet be large enough to be realistic tests and demonstrations of the feasibility of developing semantic web applications using current tools.

After a survey of many existing and proposed semantic web applications [⊣ SWADE-SURVEY] we chose two demonstration applications. The first demonstrator, semantic blogging, sought to build upon the existing use of RDF in part of the blogging infrastructure (RSS 1.0 [⊣ RSS10]) and enable the creation and exchange of structured,

semantically rich, information items. The second demonstrator, semantic portals, sought to provide an integrated view of a set of information resources described by a distributed community.

2.3 Why did we choose semantic blogging?

In our applications survey [⊢ SWADE-ANALYSIS], we noted that semantic web technologies are well suited to tasks where a user community is incrementally publishing structured and semantically rich (categorized and cross-linked) information. We also noted that blogging is a very successful paradigm for lightweight publishing, providing a very low barrier to entry, useful syndication and aggregation behaviour, a simple to understand structure and decentralized construction of a rich information network.

The notion of semantic blogging builds upon this success and clear network value of blogging by adding additional semantic structure to items shared over the blog channels. In this way we add significant value allowing view, navigation and query along semantic rather than simply chronological or serendipitous connections. Our semantic blogging demonstrator design emphasises three key behaviours:

- **Semantic View**: Semantically enriched blog metadata enables context sensitive, metadata driven views of the blog content (over and above fixed templates).
- **Semantic Navigation**: Semantically enriched blog metadata enables new blog navigation modalities (over and above unlabelled links).
- **Semantic Query**: Semantically enriched blog metadata enables richer query and discovery mechanisms (over and above free text search).

We should note here that the use of the term 'semantic' emphasises the use of semantic web technology to enable these behaviours. It is true that in the current instantiation of the demonstrator, the new capabilities are enabled primarily by using rich metadata, and require little in the way of actual semantic machinery. However, in each case it is easy to see how the behaviour can be further extended by adding inferencing over a semantic model. A simple example would be subcategory inferencing for semantic query. We prepare our blog for such possibilities by encoding its metadata in RDF.

In summary then, the rich structure and query properties enabled by the semantic web greatly extends the range of blogging behaviours, and allows the power of the metaphor to be applied in hitherto unexplored domains. One such domain (and the one explored in the demonstrator) is small-group management of bibliography data.

2.4 Why did we choose semantic portals?

The second demonstrator was chosen partly to complement the first. Whereas the semantic blogging is primarily aimed at publication of information the semantic portal application is primarily aimed at aggregation and viewing of information published using semantic web formats. Though, in the event the semantic portal demonstrator also involved substantial work to support data entry and publication as well.

The core notion of the semantic portal application is to illustrate how the decentralized information sharing enabled by the semantic web could benefit the construction of community information portals. We identified several ways in which such portals could benefit from semantic web technologies.

Ontologies

The use of an explicit, shared domain ontology enables both data sharing and richer site structure and navigation including multidimensional classification and browsing schemes. Use of the Semantic Web standards for encoding these ontologies also enables the ontologies themselves to be shared and reused across portals. Several projects have already derived benefits from ontology-driven portal designs [⊢ SEAL][⊢ WEB-PORTALS].

Evolution

Requirements change over time leading to a need to extend the information model. The semantic web helps in two ways. Firstly, the user interface and submission tools can be generated from the declarative ontology. Secondly, the semi-structured data representation of RDF permits data extensions to be added in a new format, without invalidating existing data, in such a way that both original and extended formats can be used side by side. This suggests an alternative approach to information portal design. Instead of a long top-down design cycle, we start from a seed ontology and information structure that we extend incrementally.

Community extensions

Whilst many portals support constrained community annotations, such as comments and ratings, the semantic web approach allows more extensive community customization. For example, during previous work on a portal for wildlife multimedia it became clear that many user communities would like specialized navigation of the data (based on formal species taxonomy or behavior depicted), which was unfeasible for the centralized portal provider. Using the decentralized approach it is possible for communities to develop these specialist navigation structures as a set of external RDF annotations on the portal data. The central site can then aggregate the community-provided enrichments and make them accessible to all users.

Aggregation and decentralization

One problem with traditional information portals is that they are often dependent on the responsiveness of the central maintainers, so that if funding disappears, so may the data. In the semantic web approach the supplying groups host their own data and the portal becomes an aggregating service. Central organization is still needed (for example, to provide the initial impetus and ensure that appropriate ontologies and controlled vocabularies are adopted). However, once the system reaches a critical mass it can more easily be self-sustaining - anyone can run an aggregator service and ensure continued access to the data or a new supplier can add data to the pool without a central organization being a bottleneck.

Together these features serve to illustrate each of the semantic web characteristics we identified earlier.

We grounded the demonstrator by choosing a specific portal application and sought to be provide genuine value to that community, to keep ourselves honest by working with real users. The application we chose to develop was a directory of environmental, wildlife and biodiversity applications.

2.5 Demonstrator Objectives

To summarize the demonstrators had three main objectives.

First, they had to illustrate the nature of the semantic web. We wanted it to be possible to see, by looking at the demonstrators, that the semantic web is concerned with sharing of information across boundaries and to understand the

mix tools that it offers - semi-structured data representation, representation of conceptual models and ontologies, anchored in the technology and design principles of the web.

Secondly, they had to be practical and convincing demonstrations. They were intended to be valuable enough within their chosen domains that other members of the community would be motivated to build upon them or upon the ideas. It was important that they should be seen as practical and give evidence of what can be done with the technology in its current state. These were not research activities.

Thirdly, whilst the demonstrators were aimed at specific domains we wanted it to be possible to generalize from them. To learn lessons about semantic web applications that could apply to related domains and, where possible, have the software developed be itself reusable and a useful spin off from the work.

3 Semantic blogging

3.1 Semantic blogging - summary

The core notion of semantic blogging is to take the blogging approach to light weight publishing and enhance it by enabling authors to publish embedded structured information and metadata. As the test domain for the demonstrator we chose the problem of publishing bibliographic information and sharing it amongst small work groups. A typical usage would be that an individual, writing some comments on an article they have just read, would publish the comments as a blog entry along with a full bibliographic reference embedded using RDF. The bibliographic reference could be classified using appropriate domain-specific vocabularies. In that way a community could build up a decentralized database of bibliographic information with associated comments and classifications.

The vision of semantic blogging is that semantic web technologies enable new blogging modalities that would be difficult or impossible otherwise. For the purpose of the demonstrator, we chose three axes:

- **Semantic View** using semantic metadata to enhance the view of rich content. For example, viewing bibliographic items as record cards, arranged in tables or grouped in clusters. Crucially, the semantic view is chosen by the consumer, not the provider, on the basis of the metadata. So it would be possible for example to arrange someone else's blog entries according to your own categorisation scheme.
- **Semantic Navigation** using semantic metadata to more efficiently find blog items of interest. For example, using dynamically constructed tree type interfaces to browse related items. Or following labeled links such as 'agrees with' or 'part of'. Another possibility is defining a semantic similarity measure so that one can ask the question "find me similar entries".
- **Semantic Query** using semantic metadata to build rich queries over a community's knowledge. For example, "Find all blog entries about a paper written by this author", or "find me all blog items about my friends".
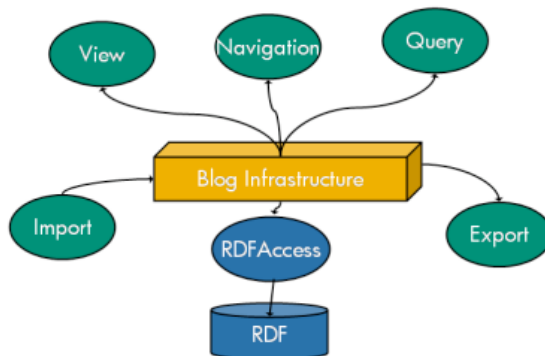

Figure 1- High level architecture of semantic blogging demonstrator

Figure 1 shows the basic structure of the semantic blogging demonstrator, an online version of which is available at [⊣ SEMBLOG_DEMO]. We can see that an RDF metadata store is built underneath a blog infrastructure. Input and output mechanisms complete the metadata pipeline, over which the semantic capabilities described above can be built. We chose to built over a Java(tm) blogging platform called blojsom and used the HP Labs semantic web toolkit Jena [⊣ JENA]. A full description of the demonstrator's design can be found in [⊣ SEMBLOG-REPORT].

3.2 Semantic blogging - lessons learnt

Previous Lessons: an update -

The previous report [⊣ LESSONS_LEARNT_SEMBLOG] detailed a number of lessons learnt. A number of these are worth reiterating, and expanding upon, here.:

Bibliographic management is an excellent choice of domain

We performed an initial user study [⊣ USER_STUDY] which indicated that current bibliographic management tools are poorly suited to the task of small group bibliographic management. In addition, the idea of a blog for a carrier of information snippets (in this case bibliographic records) is one that has found resonance both in our previous work [⊣ EPERSON] and in our current knowledge management explorations [⊣ MEME]. The notion of blog entries as annotations about information items is one that has been built into the data model from the beginning, and is one that allows the metaphor to be extended beyond blogs to information channels like wikis, email, portals, databases and custom feeds. In addition, designing for a community needing to share information snippets is clearly relevant for the semantic portal demonstrator.

For flexibility, we need to generate metadata 'on the fly' and at an externally retrievable location.

The type of metadata generated can be, and should be, client controllable. For example, various queries might return different combinations of blog entries, or an RSS feed might limit the number of blog entries to a certain number. In

addition, certain aspects of the blog metadata might be more important in some contexts (for example, do we want to return all the bibliographic information in an RSS feed?).

Importing metadata can be semantically assisted, and we show one simple way to achieve this

In the demonstrator, we developed an RDF representation of the blog categories (expressed in TIF [-| TIF], the precursor to the SWAD-E developed SKOS [-| SKOS]). Using this representation we were able to suggest relevant blog categories to the creator of a blog entry. Our thoughts have progressed beyond this simple mechanism. For example, we can use machine learning techniques to provide more refined suggestions. More importantly, we can extend the idea to other domains. In our recent thoughts about a departmental snippet manager (see below), we have considered how to maintain a simple set of ontologies (for example, "Who", "What", "Where") which can be used to mark up snippets in a very lightweight manner. The snippet manager work is at an early stage, but our thinking is certainly influenced by the experience gained on the semantic blogging demonstrator.

Use of RDF in an export feed leads to consumer issues.

The issue here is that many RSS clients parse RSS1.0 feeds as XML rather than RDF. Thus they operate at the syntax rather than the model level, and the variety of RDF serializations may pose a problem. To take one specific example, some aggregators display channel metadata (such as the blog title) only if the <channel> element precedes the blog items in the RSS feed. Fortunately, Jena [-| JENA] allows one to have some control over the XML serialization of an RDF model, for example the RDFXML writer will tend to create RDF/XML with `prettyTypes` at the top level. To take another example, the current instantiation of the software assigns each blog entry to be of type `semblog:blogItem` as well as `rss:item`. Depending on the serialization order, the item can be presented in one of the following ways:

```
<rdf:RDF
xmlns="http://purl.org/rss/1.0/"
    ...
    xmlns:semblog="http://jena.hpl.hp.com/semblog#">
```

```
<item rdf:about="http://jena.hpl.hp.com:3030/semblog/blog/news/2004/08/01/sample_entry">
    <rdf:type rdf:resource="http://jena.hpl.hp.com/semblog#blogItem"/>
    ...
</item>
```

OR

```
<semblog:blogItem rdf:about="http://jena.hpl.hp.com:3030/semblog/blog/news/2004/08/01/sample_entry">
    <rdf:type rdf:resource="http://purl.org/rss/1.0/item"/>
    ...
</semblog:blogItem>
```

The latter option, though isomorphic in RDF model terms, may cause aggregators to ignore the blog entry.

Values

Semantic view is a good way to show how a blog can look very different with a little effort. Semantic navigation is potentially useful but requires richer metadata for a full exploration of the possibilities (eg facet browsing). Semantic query enables one to expose the rich structure of a semantic blog (eg "show me all blog entries about this paper"). The idea of a 'semantic view' is one that has been explored and extended in the semantic portals project. Semantic navigation (in particular facet browsing) has been utilized extensively in the environmental portal, and has proven quite popular with our early adopters. Semantic query is expected to be a key feature of our snippet manager (see below).

3.3 Semantic blogging - update and new developments:

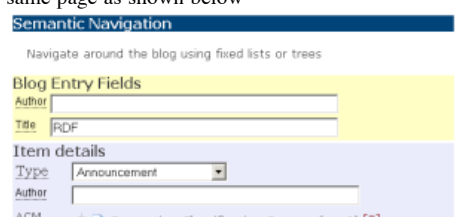### Semantic blogging taken up by KM group within HP

The use of semantic web for informal knowledge management has been greeted with interest and enthusiasm by the KM group in HP. These themes are explored in an upcoming HP Labs technical report [-| SEMBLOG_KM]. We have had a number of discussions with this group, and are likely to work on an internal pilot project, inspired by semantic web technologies in general and semantic blogging in particular.

Semantic Blogging demonstrated at XMLEurope 2004 and HP TechCon

Semantic blogging has been demonstrated successfully to the XML community at -| XML Europe 2004 (April 2004) and to the HP leadership at HP TechCon 2004 (June 2004).

Semantic navigation and query conflated

In the original demonstrator there was a distinction between semantic navigation and query. However, although conceptually these are distinct (one is a browsing activity, one a seeking behaviour), in practice navigation is often presented as a kind of constrained query. Therefore in the context of the semantic blog, the two are now presented on the same page as shown below
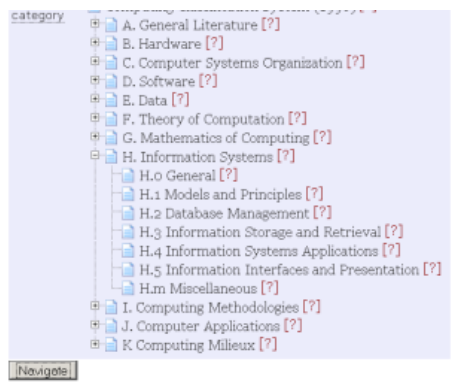
Departmental Snippet Manager

Inspired by the lessons of the semantic blog, we are actively planning a internal project to construct a departmental snippet manager. The idea here is to capture, in a lightweight manner, snippets of information which can be shared within a small group. The small group setting allows both informality and a reasonable chance of being able to share a conceptual model. In fact, many of these themes have been explored previously within our group [⊣ EPERSON], but we are now able to take on board lessons learnt from the semantic blogging demonstrator. For example, people are generally unwilling to invest time in adding snippets, so the process should be very simple (like creating blog entries, or sending emails to a distribution list). This extends to a reluctance to download & install specialised snippet manager applications, although bookmarklets appear to be acceptable. Secondly, people may have different preferred containers for capturing information - blogs, wikis, newsgroups, databases, mind mapping tools and so on. Thirdly, one should be able to integrate the snippet data with external sources (newsfeeds, enterprise directories, technical report archives...). Finally, the idea of a distributed portal, as developed in the semantic portals project, allows a community view over snippets without forcing everyone to submit their knowledge to a centralised store.

    The snippet manager is still in early planning stage. However, even now two things are apparent. Firstly, the data model (blog entries are annotations over information items) is a good one, which allows semantic blogs to integrate well with other data sources. The idea of blogs as carriers of information plays well in the snippet manager space. Secondly, the real challenge (and one that is not really explored by either of the two demonstrators) is to allow the evolution of shared ontologies, while keeping data creation and import simple and lightweight.

Software feedback

The semantic blogging demonstrator software was made available under open source license and we have had some take up. A small number of users have also given qualitative feedback on the process. In general, the semantic blogging demonstrator is not straightforward to set up and configure. This is partly because of the underlying platform, blojsom, which used to require some setup configuration. But the semantic blogging software itself had a large number of configurable parameters, which has hampered people's ability to get it up and running. With regard to the former, blojsom is now deployable pretty much 'out of the box', given a servlet container. The latter is a current focus of development; and serves as a reminder that to lower the barrier of entry, one should strive to provide an easy 'try it out' mechanism. The demonstrator itself [⊣ SEMBLOG-DEMO] of course is one such route but it does not by itself allow people freedom to experiment with their own semantic blog.

Relevant tools and developments

We have been tracking recent semantic blogging related developments, some of which are briefly described below.

- **Platypus** (⊣ http://platypuswiki.sourceforge.net/) is a semantic wiki . Like other wikis, it offers a simple user interface to create a web page but enables the user to add RDF metadata. The idea of extending semantic blogging ideas to other tools is one that has also occurred to us, and in fact we are using a simple semantic wiki in our department. The key (just as in the blog) is to provide the user with a low cost (ideally zero cost) way to enrich content with useful metadata. This is achieved here through a wiki like syntax (in Platypus, the convention namespace:pagename).
- **Semblog** is designed to be a semantic publishing suite. A recent (2004) paper [⊣ SEMBLOG] summarises the project to date while an an associated blog (⊣ http://www.semblog.org/) and wiki (⊣ http://www.semblog.org/wiki/?semblog) provide updates on progress (some in English, some Japanese). The suite includes a site builder (RNA) which provides customisable views over RSS feeds, and an RSS aggregator (glucose) with group propagation capabilities. Using this base the authors have experimented with collaborative recommendation, co-citation search and social network construction.
- **Compendium** (⊣ http://www.compendiuminstitute.org) is a tool from the KMI institute at the Open University for "collectively elicit[ing], organiz[ing] and validat[ing] information". Essentially it's a graphical hypertext system which allows a group to incrementally build a collaborative mind map of the discussion (Bachler et al [⊣ ONTOLOGICAL_MEDIATION] call this a *semantic group memory*) when used in a meeting scenario. Although this is not obviously aligned with semantic blogging, it illustrates a substantive research effort in the important area of informal, group based metadata creation. One could use it, for example, as a kind of visual semantic blog authoring tool.
- One of the key ideas of semantic blogging is the idea of an informal, shared, peer-based lightweight ontology. This enables semantic blogs (and wikis) to be aggregated in a useful way. A number of tools are working along the same lines - for example **del.icio.us** (⊣ http://del.icio.us/) for bookmarks and **Flickr** (⊣ http://www.flickr.com) for photographs. Both these tools use simple flat lists of tags - if tags match (syntactically) between users, they are taken to refer to the same thing. This is clearly a convenient simplification, for example many of the Flickr photographs with the tag 'bristol' were not taken in the UK. A slightly more sophisticated approach was taken by the semantic blogging demonstrator which expressed blog categories in a simple ontology, TIF [⊣ TIF]. Dan Brickley has done a similar thing using the SWAD-E standard SKOS [⊣ SKOS] for a different blogging platform, WordPress (see ⊣ http://lists.w3.org/Archives/Public/public-esw/2004Aug/0001.html for details).

- **Rotten Tomatoes** (⊣ http://www.rottentomatoes.com/movies/) allows users to blog comments about movies; these comments are then aggregated and presented alongside a movie. There is no evidence that this application uses RDF; however the concept of an information item as a first class entity, which can be used to aggregate blog entries about that item, is very close to the thinking behind the semantic blog.

In addition to these tracked developments, some discussions have been ongoing both with HP groups (eg KM group, consulting, Research library) and external (eg University of Warwick, Communications of the ACM, Open University, Elsevier). There is no shortage of possibilities for developing the semantic blogging theme, and it is likely to remain an active research area for us (and others) long after the official SWAD-E project completion.

4 Semantic portals

4.1 Semantic portals - summary

We use the term Semantic Portal to refer to an information portal in which the information is acquired and published in semantic web format and in which the structure and domain model is made explicit (e.g. in the form of published ontologies).

There are several advantages to using semantic web standards for information portal design. These are summarized in Table 1.

Traditional design approach	Semantic Portal
Search by free text and stable classification hierarchy.	Multidimensional search by means of rich domain ontology.
Information organized by structured records, encourages top-down design and centralized maintenance.	Information semi-structured and extensible, allows for bottom-up evolution and decentralized updates.
Community can add information and annotations within the defined portal structure.	Communities can add new classification and organizational schemas and extend the information structure.
Portal content is stored and managed centrally.	Portal content is stored and managed by a decentralized web of supplying organizations and individuals. Multiple aggregations and views of the same data is possible.
Providers supply data to each portal separately through portal-specific forms. Each copy has to be maintained separately.	Providers publish data in reusable form that can be incorporated into multiple portals but updates remain under their control.
Portal aimed purely at human access. Separate mechanisms are needed when content is to be shared with a partner organization.	Information structure is directly machine accessible to facilitate cross-portal integration.

Table 1 - contrast semantic portals proposal with typical current approaches

To illustrate these advantages in practice and to build a functioning demonstrator we needed to pick a domain for the demonstrator portal. As the demonstrator domain we chose to develop a directory of UK environmental, wildlife and biodiversity organisations. We termed this specific demonstration service the Semantic Web Environmental Directory, abbreviated to SWED.

The idea is that each organization wishing to appear in the directory provides their organization description as RDF data, using a web-based data entry tool, and then hosts the data at their own web site (similar in style to FOAF [⊣ FOAF]). A portal aggregates the RDF data and provides a faceted browse interface to allow users to search and browse the aggregated data. Annotations to this data can be created by third parties and hosted by the suppliers or by an annotation server. These annotations permit new classification schemes and relational links to be added to the data. In particular, the ability to add new links is seen as opening up exciting opportunities to capture and visualize the complex relationships between environmental organizations.

For more background on the limitations of the existing directory solutions and ways in which this test domain is a good match to the semantic portals approach see the requirements document [⊣ SEMPORTALS-REQUIREMENTS].

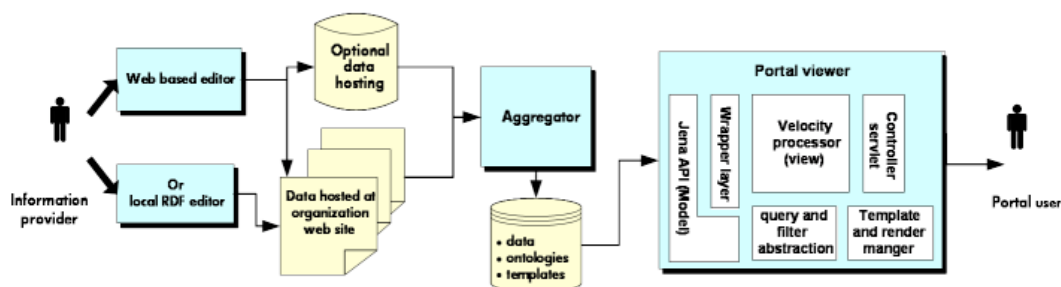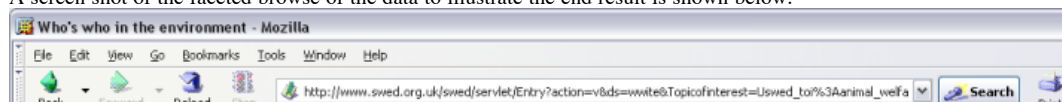An overview of the demonstrator is illustrated in figure 2 below:

Figure 2- block architecture of the demonstrator

The organisions use a web-based data entry tool to create an RDF description, publish that description on their web site and notify the portal. The portal aggregator scans such sites for new or changed data and uploads a copy into the portal database. The portal web viewer provides a rich faceted browser for navigation of the aggregated data. This viewer is constructed as a general purpose template-driven RDF web viewer which can be customized and repurposed to provide different views and navigation over information encoded in RDF.

A screen shot of the faceted browse of the data to illustrate the end result is shown below:

Figure 3- screen shot of SWED demonstrator results page

For more information on the semantic portals demonstrator and the software design see [⊣ SEMPORTAL-REPORT]. The demonstrator itself may be visited at: ⊣ http://www.swed.org.uk.

4.2 Semantic portals - lessons learnt

The detailed lessons learnt from the semantic portals demonstrator are detailed in [⊣ SEMPORTAL-REPORT]. Here we shall summarize and highlight just those lessons that seem most generally applicable.

Decentralization - benefits and costs

One of the core features of the semantic portals approach which distinguishes it from a typical portal design is the distributed nature of the data ownership. In the specific SWED demonstrator the data is provided by the organisations who wish to register in the directory - they maintain and publish their own information. This local ownership of data is perceived as a key benefit. It helps with sustainability of the overall system. If the central portal ceases to operate all of the raw data remains available and anyone can run a new portal harvesting the published data. The organisations can update and modify their own data at any time. It is easy for sub-groups in the community to augment the portal with new information such as classifications (e.g. classify organisations according a taxonomy of the species they work with) or relations (e.g. to link organsations to projects that have spun off from them). It is easy for a third part to run a parallel, value added portal, for example to offer an alternative user interface or to combine the organisation data with other resources such as publication information.

These potential benefits were clear going into the demonstrator, though the appeal of them does seem to have born out. For example, the National History Museum are forming a collaborative group to create an infrastructure for sharing descriptions of collections of natural history resources. Inspired by the SWED demonstrator they would like to explore semantic web technology for this - the ability to have each partner host and maintain their own data, yet have it integrated for the community, was seen as a good benefit.

However, these benefits do come at a cost. For a user of the portal it is not necessarily clear where the information they are seeing has come from and how definitive it is. At the RDF level the different information elements are mixed together so that a portal page displaying information on one topic (on organisation in the SWED case) may be derived from multiple sources. If we make the source of each separate data item explicitly visible then the user interface can become complex and confusing. If we hide the source by default then the user might be accidentally fooled into trusting information from a third party source believing it to be from the organisation itself. In the SWED demonstrator we provide an interface which does highlight classification and links which come from sources other than the rest of the page and do show the source of the main page data. However, this is certainly not a complete user interface solution. Furthermore we do not attempt to show the user which information is definitive. It is up to them to decide whether to trust the data for their purpose. The whole notion of how we can model and convey trust in a decentralized information system is one requiring more research - particularly in terms of user interface paradigms.

Open ended extensions - a challenge

Related to the above issue is that of extensibility of semantic web information models. The semantic web makes an open world assumption - you must assume that a given resource may have other properties that you are not yet aware of. This permeates the entire notion of vocabularies within the semantic web. Instead of a schema approach in which resources are treated as if they were objects with a fixed number of property values the RDF vocabulary languages just express constraints on the properties themselves. This makes it very easy to create new vocabulary terms and use them to attach additional information and links to exist resources.

The benefits of this extensibility are clear. Without it the decentralization discussed above would not be so useful. With it then sub-communities can agree extensions to the vocabularies and use them without their extensions interfering with existing tools and data or with other parallel extensions. It also makes it possible incrementally enrich

the data model over time without breaking backward compatibility with existing applications.

However, again the cost of this freedom is significant.

First, consider the problems of data viewing. It is hard to make the tradeoff between a general or tailored viewing solution. It we adopt a generic approach to data visualization that sticks closely to the RDF model then the information is not very intelligible to a typical end user. In the SWED application if we the view the RDF description of an organisation as a table of properties (as is indeed supported as *view raw* in the demonstrator) or as a property graph then all the information can be seen but it is very hard for a non-expert to read. Conversely if we format the information in a way closely tailored to the specific application and end user then we can arrive at a very legible presentation but the result is brittle, data extensions will be missed. We provided partial solutions to this within the demonstrator. We adopted a template-driven approach in which the templates can be as generic or tailored as the designer wishes. We use late-binding to select the right template to use based on the RDF type of the resource being shown and display context in which it is embedded - this allows new templates to be added and discovered at visualisation time to tailor the display to new information types. We group properties together (using subproperty hierarchies and meta-classes of properties) to make it possible to write templates that can cope with structured extensions (e.g. putting a new relation link in the relations display or a new address property in the contact information box). These are all useful engineering work-arounds but this area of display approaches for open ended data models seems worth further investigation.

Second, a related problem concerns the data model itself. One cost of the open world, property-centric approach is precisely that we lack any form of schema language for RDF. Whilst extensibility is useful there are also times when an information provider is actually trying to conform to a known vocabulary and would like stronger data validation than is possible via RDFS and OWL, a more conventional schema solution would be valuable for such situations. Another related cost is that it is very hard to capture this extensibility as part of the formal domain model. For example, it is not possible to convey that use of a given vocabulary element is optional.

A specific example which serves to illustrate the issue arose in developing the organisation ontology for describing "prorgs" (projects/organisations). We wished to reuse the existing `vcard:EMAIL` element from the vcard schema [⊣ VCARD]. This is defined to have an unconstrained domain, it could be applied to any class of resource, including our prorg class without problems. We do not require an organization to have an email address nor constrain it to a single address. So in fact there are no schema declarations to add. However, this means that the user of the schema has no idea than an email address might be permitted and that we would prefer them to use `vcard:EMAIL` to capture it, other than natural language documentation. There is no way to express the notion that a class has a expectation of a property value but that the property is optional. The nearest one might do would be to add a domain declaration saying that `vcard:EMAIL` has domain `swed:prorg` but this is clearly incorrect (all other uses of `vcard:EMAIL` would then incorrectly be deduced to refer to `swed:prorgs`). In this and similar cases we made use of subProperty hierarchies, we would declare a parent property (`swed:has_contact_details` in this case) for the optional properties. However, this is still not correct (it still implies additional domain constraints on the optional properties which are not globally valid) and is just a convention. A schema solution which makes it possible to distinguish between unanticipated extensions and anticipated but optional property use would be valuable.

Data entry is as important and hard as data viewing

At the start of the semantic portals demonstrator our focus was on the aggregation and display of the information. We knew that substantial work would be required to develop a sufficiently general and reusable viewing engine to enable us to create a usable portal directly from the RDF descriptions of the organisations. However, we had less awareness of just how costly the engineering of the data capture side of the problem would be. To enable the organisations to create their entries without needing to write RDF we created a web-based entry tool. Making this tool general (so that it could be adapted to capture of other information) yet usable was time consuming.

The basic issues of creating web forms for capturing of property values are tractable. Though there is the challenge of how to map the form information to RDF. For example, idiosyncrasies in the RDF/XML syntax makes it hard to use XForms for data capture directly into RDF/XML [⊣ XFORM-EXPT]. In our case we overcame these issues by using conventional forms and with server side support for mapping between the form and RDF.

The harder issues were designing the form user interface to enable entry of classification and relational data. In the case of classification data we wanted to allow data providers to select appropriate terms from the predefined thesauri. This was complicated by the need to allow multiple selections, the need for scope notes for each term and the preference to make all of the selections from a single form to encourage all providers to enter classifications. After using various open source tree controls for thesaurus display we found we had to develop our own JavaScript controls to get the right balance of usability and performance and used compression to enable us to embed the thesauri in the data entry form without excessive page load times.

The above are engineering issues. In some ways a more fundamental was the problem of relation capture. From our earlier studies [⊣ SHAB] we believed that capture of the relationships between the various projects and organisations would be a benefit of the portal. However, the user interface to allow relationships to be captured was hard. Even though that part of the form itself looked relatively simple the conceptual model of referring to another organisation via an identifying feature such as its primary URL was alien and in fact the whole notion of adding relational links at the same time as adding properties was confusing and attracted adverse user feedback. In some ways it would be better to treat relational link entry as a separate data entry phase and use some more graphical user interface paradigm to support this entry. However, that would likely mean that such information would not be entered at the same time as the basic organisation information and might get forgotten.

We were able to arrive at a usable data entry solution for the SWED demonstrator and the tools we developed are sufficiently flexible and schema-driven to be likely to reusable However, the general lesson remains that the design of the data capture user interface for rich RDF data is challenging and that more tool solutions to reduce the engineering cost of implementing these designs would be useful.

5 Overall lessons learnt

The experience of developing the two open demonstrators for this work package has generated several valuable lessons. Many of these relate to specific issues of software design, user interface design or domain modeling related to the specific applications and have already been covered above or in the individual reports on the two demonstrators. In this section we try to distill a few of the more general observations.

Decentralization and integration

The core purpose of the semantic web is to enable exchange and integration of data across application, enterprise, and community boundaries. This can be exploited in different ways. When data already exists in disparate locations, perhaps for specialist uses, then the semantic web offers new tools to enable that data to be exported and integrated to solve new problems. Conversely when architecting a new application based on semantic web technologies it becomes

easier to deliberately adopt a decentralized architecture with the attendant benefits (sustainability, scaling, timely updates) and risks (consistency, trust and user interface issues).

In both demonstrators we deliberately adopted a decentralized solution. In the case of the semantic blogging this is a natural part of the blogging architecture and we are merely exploiting this to publish structured data. In the case of the semantic portals demonstrator this decentralized approach is not normal but as we discussed [⊣ above] the benefits of this in terms of sustainability are viewed very positively, though we also noted the challenges of user interface and trust that this architecture raises.

When discussing the semantic web with technical strategists within commercial enterprises we find this decentralization message can be negatively perceived. Enterprises do indeed face problems with many disparate tools and local solutions locking data up in specialist silos and perceive great benefits for better integration. Typically this is approached by top down mandate whereby the local tools and data stores are phased out and replaced with more general, central, centrally maintained datastores. A technology which purports to facilitate continued decentralization is not welcome in such circumstances. Thus in looking to apply the semantic web in enterprise settings it is important to focus on the integration aspects in situations where the data is distributed for good reasons and not regard decentralization as necessarily appropriate. Though even in cases where the right solution is to centralize the data with a single imposed information model the semantic web technologies may still be relevant to assist with migration, to publish the centralized information model for future reuse and to provide machinery for backward-compatible evolutionary extensions to the model.

One compelling example of the semantic web as a common integration technology resulted from combining the two demonstrators to show how they can work together. We took the two separate demonstrators and integrated them into a demonstration knowledge management application which couples the highly personal decentralized blogging with centrally maintained document knowledge bases. The portal tools were used to create an integrated view onto centrally indexed documents, project and personnel information and classified news items published over personal semantic blogs. This integration experiment was very easy due to the common use of the semantic web standards and was a convincing illustration of the potential benefits of adopting a common standards-based metadata infrastructure.

Ontology reuse challenges - domain models are context dependent

The decentralization story for the semantic web applies to vocabularies as much as to the data itself. The semantic web does not include or mandate a standard upper ontology to link vocabularies together. Applications are free to develop or reuse any appropriate vocabularies, the semantic web just offers standards for how those vocabularies can be published. On the other hand if all applications generate their own vocabularies then little has been achieved, the more vocabularies can be reused the greater the integration that is possible.

At the start of the demonstrators we expected to reuse vocabularies whereas in the event we largely developed our own. This seems like an important learning and worth expanding on here at the risk of repeating points made in the individual demonstrator reports.

The core issue is that vocabularies for describing real world objects and concepts, and thesauri for classifying them, are always approximations to the reality. We are typically imposing fixed discrete categories, relationships and constraints on concepts that are in fact continuous, overlapping and context dependent. Even seemingly fundamental dichotomies such as male/female are not simple (in some contexts one may need to account for notions of mental rather than physical gender, for hermaphroditism, surgical alteration, asexual single celled organisms etc.). Thus the right approximations to make depends upon the context of use.

When thinking about structured ontologies this is fairly clear, the features of what makes up useful description of a *person* or an *organisation* can vary wildly. For example, an ontology which uses legal distinctions to partition the notion of *organisation* into different categories (registered charity etc) is not relevant to applications with more informal notions of emergent social structures or which are set in countries with an incompatible legal framework.

Rather than expect to be able to find and reuse the perfect ontology for such concepts it seems more appropriate to define stable ontologies for smaller component concepts such as names, dates, contact location or web presence. Even these are hard to do (a cross-culture model for the structure of a person's name is, for example, pretty challenging) and the ones we did use (notably [⊣ VCARD]) are not always represented in the most obvious ways. However, this does re-raise the issue that it is hard capture the constraints on their reuse in any validatable way [see discussion ⊣ above].

Another technique which can be useful is to at least mark the concepts within the custom developed ontologies as being related to some more general shared set of concepts. One approach to this, which is exploited by the *foaf* [⊣ FOAF] vocabulary is to relate the concepts to terms in a linguistic resource such as wordnet. This doesn't of itself provide interoperability but it does provide some hints for future manual or semi-automatic schema mapping tools.

Similar considerations of context-dependence and reuse apply to the design of classification thesauri. For example, within the semantic portals demonstrator a user searching the portal wants a thesaurus that is rich but navigable and doesn't mind overlap between branches, a data provider wants minimal overlap and as simple a set of terms as possible to reduce classification costs, an external integrator wants maximum precision for reuse without caring about issues of labeling cost and navigation. Thus even where industry standard vocabularies exist (e.g. the ISIC classification of industrial activities [⊣ ISIC]) they were found to be ill-suited to the user interface needs of either data provider or portal user within our specific domain even though they would have been ideal form the point of view of reusable publication of data.

One solution to this, which was adequate for the demonstrator, is to develop "colloquial" thesaurus structures which are small and navigable enough to meet the end user needs but to link them to the terms in the larger richer standards for the purposes of information integration. Whether this is possible depends on the details of whether the inheritance structures are compatible.

Tool requirements - the next challenges

One of the originally stated purposes of the demonstrators was to give feedback to tool developers on the limitations of the existing tools.

In the event the tools for manipulating, storing and querying semantic web data (RDF and OWL) were entirely adequate for these demonstrators. We used HP's Java toolkit [⊣ JENA] for both demonstrators and so are, perhaps, biased in this assessment, but we had no real difficulty with functionality, stability or performance of the tools. There are a few areas where improvements would be helpful. For example, the provenance tracking needed for the semantic portals work made use of an experimental MultiModel API. That experimental API met the need but should be folded back into the main codebase for Jena and more efficient implementations of it would be useful for larger scale applications.

Whilst we anticipate continued incremental improvements to these tools our experience suggests that core requirements for storage, query and access APIs have largely been met. Rather it seems like the most pressing tool needs are elsewhere. In particular, for applications that interact with people (not true of all semantic web applications) there is a need for richer tools for the development of data creation and data visualization of semantic web information. As part of the demonstrator projects themselves we have generated early versions of some such tools - both our data creation and web visualization tools are designed to be quite customizable and reusable. However, these

are but early attempts and experience will shown better ways to structure the problem and generalize and package the solutions. Similarly our experience during the demonstrators with use of scripting (template scripting for the visualizations, Jena rules for scripting small data transformations and inferences) suggest that a scriptable semantic web workbench would be a valuable tool to help with data ingestion and exploration.

Value of demonstrators

Finally, the basic approach of using demonstrators as a means of outreach does seem to have been very successful. Both of the demonstrators have community groups interested in extending the demonstrators or building upon the ideas. It was particularly gratifying to find examples where groups who had avoided the semantic web up till now found the demonstrators convincing illustrations of both the utility and practicality of semantic web. The reaction of the Natural History Museum representatives to the SWED demonstrator and their exploration of reusing the technology for collection-level descriptions of natural world resources is a good example of this. Similarly our experience of talking to HP customers and partners suggests that people can generalize from the specific applications shown in the demonstrators to see the more general relevance of the technology.

6 Conclusions

In summary, we have in this workpackage produced two live demonstrators of the application of semantic web tools and standards. Each of these is grounded in specific applications (bibliography management, organizational directory) but illustrates the semantic web features in ways that people can appreciate and generalize from. The demonstrators are perceived as sufficiently practical and convincing to inspire groups to consider applying the semantic web who were not previously minded to do so. They have enabled us to learn some general lessons on developing semantic web applications and produce some spin off tools for capture and display of semantic web information that we hope will prove reusable resources for the community. Indeed we have already seen bids made to develop other applications which will reuse the specific tools or the broad approach.

References

[BLOGTALK]
BlogTalk, - A European Conference On Weblogs
BlogTalk 2003: http://2003.blogtalk.net/
BlogTalk 2004: http://www.blogtalk.net/
[BLOG-VISION]
Semantic Blogging Demonstrator - The Vision
http://jena.hpl.hp.com/~stecay/downloads/blogTalk.pdf
[BLOJSOM]
Blojsom, a java based blogging platform
http://blojsom.sourceforge.net/
http://wiki.blojsom.com/wiki/display/blojsom/blojsom+Quickstart
[BROWNSAUCE]
Brown Sauce : an introduction
Steer, Damian
HP Labs Tech Report 2003
http://www.hpl.hp.com/techreports/2003/HPL-2003-10.html
[EPERSON]
The ePerson Snippet Manager: a Semantic Web Application
Banks, Dave; Cayzer, Steve; Dickinson, Ian; Reynolds, Dave
HP Labs Tech Report 2002
http://www.hpl.hp.com/techreports/2002/HPL-2002-328.html
[FOAF]
The *friend of a friend* project.
http://www.foaf-project.org/
FOAF Vocabulary Specification (RDFWeb Namespace Document 16 August 2003; Dan Brickley & Libby Miller): http://xmlns.com/foaf/0.1/
[FRBR]
Functional Requirements for Bibliographic Records
A project of the OCLC [OCLC]
http://www.oclc.org/research/projects/frbr/
[GRAPH_VISUALIZATION]
Node-centric RDF Graph Visualization
Sayers, Craig
HP Labs Tech Report 2004
http://www.hpl.hp.com/techreports/2004/HPL-2004-60.html
[IBIS]
Issue-Based Information Systems (IBIS) Terms, Danny Ayers October 2002
http://ideagraph.net/xmlns/ibis/
[ISIC]
International Standard Industrial Classification
http://unstats.un.org/unsd/cr/registry/regct.asp?Lg=1
[JENA]
Jena, a semantic web toolkit
http://jena.sourceforge.net/
http://www.hpl.hp.com/semweb/jena.htm
[JOSEKI]
Joseki, the Jena RDF Server
http://www.joseki.org/
[LESSONS_LEARNT_SEMBLOG]
Semantic Blogging - Lessons Learnt SWAD-Europe deliverable 2003
http://www.w3.org/2001/sw/Europe/reports/demo_1_report/

[MEME]
Semantic Blogging : Spreading the Semantic Web Meme
Cayzer, Steve
HP Labs Technical Report 2004
┤ http://www.hpl.hp.com/techreports/2004/HPL-2004-48.html

[MODS]
Metadata Object Description Schema
┤ http://www.loc.gov/standards/mods/mods-overview.html

[N3]
N3, an RDF language for the Semantic Web
┤ http://www.w3.org/DesignIssues/Notation3.html

[NETAPI]
RDF Net API. W3C Member Submission 2 October 2003
Graham Moore and Andy Seaborne
┤ http://www.w3.org/Submission/2003/SUBM-rdf-netapi-20031002/

[NETWORK-INFERENCE]
Network Inference
┤ http://www.networkinference.com/index.html

[OCLC]
Online Computer Library Center
┤ http://www.oclc.org/

[ONTOLOGICAL_MEDIATION]
Ontological Mediation of Meeting Structure: Argumentation, Annotation, and Navigation
Bachler, M., Buckingham Shum, S., De Roure, D., Michaelides, D. and Page, K.
1st International Workshop on Hypermedia and the Semantic Web August 30, 2003, Nottingham, UK
┤ http://www.ecs.soton.ac.uk/~dem/workshops/htsw2003/submissions/htsw03-bachler.pdf

[OWL]
OWL Web Ontology Language Reference
W3C Candidate Recommendation 18 August 2003
┤ http://www.w3.org/TR/owl-ref/
Other OWL resources available at ┤ http://www.w3.org/2001/sw/WebOnt/

[REDLAND]
REDLAND RDF Application Framework, Dave Beckett
┤ http://www.redland.opensource.ac.uk/

[RDF-LANGUAGE]
Resource Description Framework (RDF) Model and Syntax Specification
W3C Recommendation 22 February 1999
┤ http://www.w3.org/TR/REC-rdf-syntax/
RDF/XML Syntax Specification (Revised)
W3C Proposed Recommendation 15 December 2003
┤ http://www.w3.org/TR/rdf-syntax-grammar/
Resource Description Framework (RDF): Concepts and Abstract Syntax
W3C Working Draft 10 October 2003
┤ http://www.w3.org/TR/2003/WD-rdf-concepts-20031010/
Other RDF resources available at ┤ http://www.w3.org/RDF/

[RDF-MODEL]
RDF Semantics
W3C Proposed Recommendation 15 December 2003
┤ http://www.w3.org/TR/rdf-mt/

[RDF-SCHEMA]
RDF Vocabulary Description Language 1.0: RDF Schema
W3C Working Draft 10 October 2003
┤ http://www.w3.org/TR/rdf-schema/

[RSS10]
RSS 1.0 Standard
┤ http://web.resource.org/rss/1.0/
RSS1.0 Specification: ┤ http://web.resource.org/rss/1.0/spec
RDF Rich Site Summary - history and resources: ┤ http://www.oasis-open.org/cover/rss.html

[RSS20]
RSS 2.0 Standard
┤ http://backend.userland.com/rss

[RSS-MODULES]
RDF Site Summary 1.0 proposed modules
┤ http://web.resource.org/rss/1.0/modules/proposed.html

[SEAL]
Semantic portal - The SEAL approach,
Maedche, A., Staab, S., Stojanoic, N., Studer, R., Sure, Y.,
in *Creating the Semantic Web* Fensel, D et al (eds.) MIT Press, MA, Cambridge, 2001.

[SEMAVIEW]
Semaview
┤ http://www.semaview.com/

[SEMBLOG]
Personal Knowledge Publishing Suite with Weblog
Ohmukai, I., Takeda, H. and Numa, K 2004
in Glance, N., Adar, E., Hurst, M. and Adamic, L. eds., WWW 2004 Workshop on the Weblogging Ecosystem:
Aggregation, Analysis and Dynamics (2004)
┤ http://www-kasm.nii.ac.jp/papers/takeda/04/www2004ohmukai.pdf

[SEMBLOG-DEMO]
Semantic Blogging Demonstrator
┤ http://jena.hpl.hp.com:3030/blojsom-devt/blog/

[SEMBLOG-DESIGN]
Semantic Blogging Demonstrator - Design Notes
┤ http://jena.hpl.hp.com/~stecay/downloads/architecture.pdf

[SEMBLOG-KM]
Semantic Blogging and Informal Knowledge Management

Steve Cayzer
HP Labs Technical Report, in preparation

[SEMBLOG-REPORT]
12.1.3: Open Demonstrator 1: Semantic Blogging demonstrator
Steve Cayzer, 2003
http://www.w3.org/2001/sw/Europe/reports/demo_1_report/

[SEMPORTALS-REQUIREMENTS]
12.1.5 Semantic Portals - Requirements Specification
Dave Reynolds, Paul Shabajee
http://www.w3.org/2001/sw/Europe/reports/requirements-demo-2/

[SEMPORTALS-REPORT]
12.1.7: Semantic Portals Demonstrator- Lessons Learnt
Dave Reynolds, Paul Shabajee, Steve Cayzer, Damian Steer
http://www.w3.org/2001/SW/Europe/reports/demo-2-report/

[SHAB]
Shabajee, P. (2004) *Summary Report from SWARA Survey of Biodiversity/Wildlife Information in the UK*, HP (Hewlett Packard) Labs Technical Report, HPL-2004-58.
Available: http://www.hpl.hp.com/techreports/2004/HPL-2004-58.html

[SHIRKY-SYLLOGISTIC]
The Semantic Web, Syllogism, and Worldview
http://www.shirky.com/writings/semantic_syllogism.html
See summary of discussion at http://www.poorbuthappy.com/ease/semantic/

[SKOS]
Simple Knowledge Organisation System
A deliverable of [SWADE] workpackage 8.1 http://www.w3c.rl.ac.uk/SWAD/deliverables/8.1.html
SKOS core schema http://www.w3c.rl.ac.uk/2003/11/21-skos-core
SKOS mapping schema http://www.w3c.rl.ac.uk/2003/11/21-skos-mapping

[SWADE]
Semantic Web Advanced Development - Europe http://www.w3.org/2001/sw/Europe/

[SWADE-ANALYSIS]
Semantic Web Applications - Analysis and Selection HP SWADE Report 2002
http://www.w3.org/2001/sw/Europe/reports/open_demonstrators/hp-applications-selection.html

[SWADE-SURVEY]
Open demonstrators: Semantic web applications - analysis and selection. Appendix: Application survey
Dave Reynolds , Steve Cayzer, Ian Dickinson, Paul Shabajee
http://www.w3.org/2001/sw/Europe/reports/chosen_demos_rationale_report/hp-applications-survey.html

[SWADE-THESAURUS]
SWAD-E Thesaurus activity
http://www.w3c.rl.ac.uk/SWAD/thesaurus.html
SWAD-E [SWADE] Workpackage 8 description: http://www.w3.org/2001/sw/Europe/plan/workpackages/live/esw-wp-8.html
RDF Thesauri: http://www.w3c.rl.ac.uk/SWAD/rdfthes.html

[SWADE-XML]
Integration with XML Technology SWAD-E [SWADE] Workpackage 5
http://www.w3.org/2001/sw/Europe/plan/workpackages/live/esw-wp-5.html

[SWADE-USER-STUDY]
User Study on Bibliographic Management
http://www.w3.org/2001/sw/Europe/reports/open_demonstrators/hp-requirements-specification.html#sec-userStudy
Part of [SEMBLOG-RQMTS]

[TIF]
Thesaurus Interchange Format
Paper: "Modelling Thesauri for the Semantic Web" http://www.w3c.rl.ac.uk/SWAD/thesaurus/tif/deliv81/final.html
Schema: http://www.w3c.rl.ac.uk/SWAD/thesaurus/tif/tif.html
Now replaced by [SKOS]

[TOPIC-MAPS]
Topic Maps activity
http://www.topicmaps.org/

[VCARD]
Representing vCard Objects in RDF/XML
Renato Iannella, W3C Note 22 February 2001
http://www.w3.org/TR/vcard-rdf

[XFORM-EXPT]
An experiment in the use of XML forms for representation of RDF data.
Damian Steer, SWADE report, to be published.

B Changes

31-8-2004
Initial draft.