



Open Research Online

The Open University's repository of research publications
and other research outputs

Semantic web technology to support learning about the semantic web

Book Chapter

How to cite:

Dzbor, Martin and Motta, Enrico (2007). Semantic web technology to support learning about the semantic web. In: Luckin, Rosemary; Koedinger, Kenneth R. and Greer, Jim eds. Artificial Intelligence in education - building technology rich learning contexts that work. Frontiers in Artificial Intelligence and Applications, 158. IOS Press, pp. 25–32.

For guidance on citations see [FAQs](#).

© [not recorded]

Version: [not recorded]

Link(s) to article on publisher's website:

<http://www.booksonline.iospress.nl/Content/View.aspx?piid=6753>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

Semantic Web Technology to Support Learning about the Semantic Web

Martin DZBOR and Enrico MOTTA

Knowledge Media Institute, The Open University, Milton Keynes, UK

Abstract. In this paper we describe ASPL, Advanced Semantic Platform for Learning, designed using the Magpie framework with an aim to support students learning about the Semantic Web research area. In particular, we describe the evolution of ASPL and illustrate how we used the results from a formal evaluation of the initial version of the system to re-design the functionalities provided to the users. The second version of ASPL semantically interprets the results provided by a non-semantic web mining tool and uses them to support various forms of semantics-assisted exploration, based on pedagogical strategies such as performing *lateral steps* and *query filtering*.

1. Introduction

Education, like many other disciplines, takes advantage of Web to provide learning resources speedily and easily, and to tailor them to the specific needs of a learner. However, education has always relied on a strong *interpretative component*. In addition to recalling knowledge from knowledge bases, searching document repositories or retrieving from information warehouses, education requires also analysis and synthesis – both on the level of individual learners and at group level. Interpretation, in general, comprises the ability to link otherwise independent information sources, to make statements about them, and to make inferences from the available knowledge. Above all, education is a social, interactive activity, which centers on the learners and expects their active participation.

The size of databases and other repositories of resources that are suitable for learning is no longer the greatest obstacle in harnessing the Web in educational practice. Already in 1940s Bush [3] pointed out there was more information published than it was possible for humans to process. This processing bottleneck has not been overcome yet; in learning it often takes the form of relating one chunk of knowledge to another and applying it in new situations. Similarly, Bloom [1] argued that information processing goes beyond its recall and that advanced cognitive processes, such as synthesis and judgment lead to longer-lasting knowledge. These processes share one feature: they are *relational*; i.e., they comprise associations between separate pieces of information.

Although humans rely on associations to make sense of information, the challenge of supporting associative thinking is far from resolved. For instance, in [9] the authors point out that as (i) it is hard to formally capture all subtleties of a learning task in a tutoring system, and (ii) learner modeling is only approximate, tutoring systems tend to be over-constrained closed worlds. To address the constraints, it is more valuable for the learner to see *what can be done* with a given knowledge, rather than merely following a prescribed workflow for a learning task.

In this paper, we present an approach to designing a semantically enriched application, the *Advanced Semantic Platform for Learning* (ASPL), to facilitate exploratory learning on the Web. We summarize results of evaluating a more data retrieval centred ASPL-v1, and justify design amendments for ASPL-v2. We conclude by situating this work in the wider context of supporting learning and exploration on the Web, and we highlight a number of principles about the effective use of Semantic Web technologies in this context.

2. Evaluating a Semantic Web application for learning support

Fig. 1 shows an initial version of our experimental system (ASPL-v1) loaded with an ontology about the Semantic Web research and community and applied to the web page of W3C reports library (<http://w3.org/TR>). ASPL is based upon our *Magpie* framework (for details see e.g. [7]), which offers a generic solution to partially address a known knowledge acquisition bottleneck [11]: namely, how to apply ontologies to the task enrichment (in our case, to the selection and navigation through learning resources), and how to accomplish this in the absence of extensive semantic annotations of such resources.

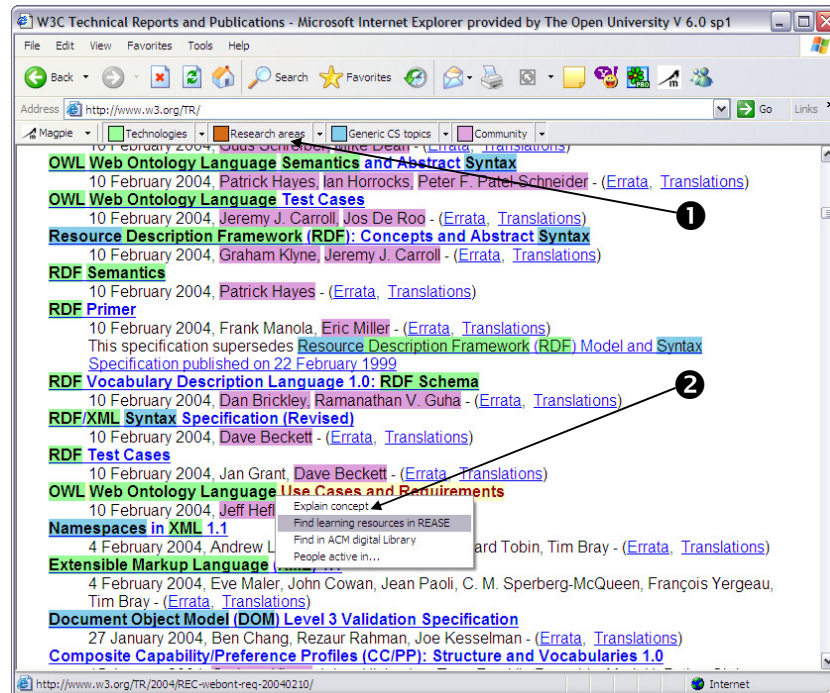


Fig. 1. A screenshot showing a Magpie-enhanced web browser on a visited web page, which is annotated using the lexicon semi-automatically acquired for the Semantic Web domain

In terms of learning, ASPL intends to offer learners the access not only to atomic resources (e.g. publications) but also to relational associations. Our aim is to show that semantically associated entities enable the learner to draw more abstract analytic and/or synthetic conclusions, which are, in turn, beneficial to support an open-ended task of analyzing and reviewing the state-of-the-art in a given research domain.

ASPL as a Magpie-based application takes a user-selected lexicon, and uses it to annotate web pages, as the user browses the Web. Lexicons are organized into top-level categories, shown by marker ❶ in Fig. 1, which form the basis for ASPL annotations, so-called semantic layers (colour highlights in Fig. 1) over standard, non-semantic resources. To each annotated item the Magpie's web browser plug-in associates a list of web services that may facilitate e.g. a particular learning task. Hence application development with Magpie consists of: (i) designing and populating an ontology; (ii) selecting the key classes in the ontologies for the purpose of supporting semantic navigation – as e.g. indicated by marker ❷ in Fig. 1; and (iii) associating web services to these classes, to provide ontology-driven functionalities.

The ASPL-v1 experimental system was bootstrapped with ontologies describing the Semantic Web research and community. The 'Community' and 'Research areas' categories were *automatically* populated using CORDER – a tool for mining and capturing of entities from text corpora [16] in the role of a Named Entity Recognizer [12] to identify typed entities in web pages and, by large-scale web crawling and taking in account several factors beyond mere term co-occurrence (e.g. the distance between entities in a resource), we obtained accu-

rate relational knowledge about these entities. In our user study we aimed to identify how people explore the domain of retrieved scientific materials applicable to the task of preparing a literature review on a given subject. To prevent confounding the study with our views on how participants should carry out the task, we constrained the services largely to information retrieval. That is, for instance, publications were found and displayed, with no further guidance on what to do next or inference provided.

2.1 Formative evaluation and requirements gathering

The evaluation with users took place in November 2005 at four different universities in Europe, referred below as Site 1 to 4. Our aim was (i) to see how semantic annotations and services are used by the students during an information gathering task, and (ii) to identify ways of supporting more complex parts of the task of preparing the literature review using the Semantic Web. Group A comprised participants using standard search tools such as Google; thus this was our control group. Group B involved participants using ASPL-v1.

The participants were asked to compile ten references, which is their view addressed the task they were given. The interactions with the ASPL-v1 were screen and voice recorded, and responses were marked by a domain experts based on their subjectively perceived suitability for a literature review. A participant's score was the sum of all the marks, and is shown in Table 1. The 'Mean quality' reflects only the marks from the experts; whereas the 'Overall mean score' takes into account bonuses/penalties (mostly to do with time limits for the task), which were converted to points.

Table 1. Overall and per-location results of ASPL-v1 evaluation

	Mean quality		Overall mean score		Significance
	A	B	A	B	t-test (overall)
Overall	10.85	9.79	19.11	18.19	0.2851003
Site 1	13.67	10.29	24.11	20.68	0.06526177
Site 2	12.00	8.00	20.25	13.82	0.06365032
Site 3	8.57	12.50	15.57	22.26	0.00469364
Site 4	8.86	8.00	15.57	14.71	0.39071656

Overall, the variance in performance was not statistically significant for the whole population (at $p=5\%$) and only 'Site 3' showed a performance rise for the ASPL users. As we suspected, skilled users of Web search engines did not find much value in Magpie-annotated web pages for the sole purpose of retrieving but not discriminating the results in a manner similar to a generic search engine. The outlier 'Site 3', however, comprised students with a prior tuition in writing literature reviews. They relied on semantic annotations more frequently and tried to interpret and explore the retrievals through Magpie services. The annotations helped them filter to the key references among retrievals. This outlier group spent more time navigating through the resources and exploring them in order to ascertain their quality, and used a range of navigation strategies (e.g., via authors, via topics, via related topics, etc.)

If we qualitatively generalize this – our outliers went beyond mere information retrieval, whether semantic or not. They showed expectations for more *investigative navigation*, which was what we aimed to capture and embed in the ASPL. For example, these outliers expected to obtain some guidance on what to do with partial solutions, assistance with query changes and re-formulations, etc. A flat list of semantically indiscriminate records provided in the initial demonstrator of ASPL was clearly insufficient in this respect.

2.2 Conceptual analysis of the learning task

In our view, people who managed to replicate some aspects of the investigative or *exploratory navigation* through the retrieved resources coped with the task better. The challenge for the

ASPL re-design was thus to facilitate more of such guided exploration in otherwise large and ill-defined problem space. If we look at the task our participants tackled, its outcome is essentially an argument comprising different publications and rationale, which can be conceptualized in terms of key activities it involves. *One such conceptualization* is shown in Fig. 2

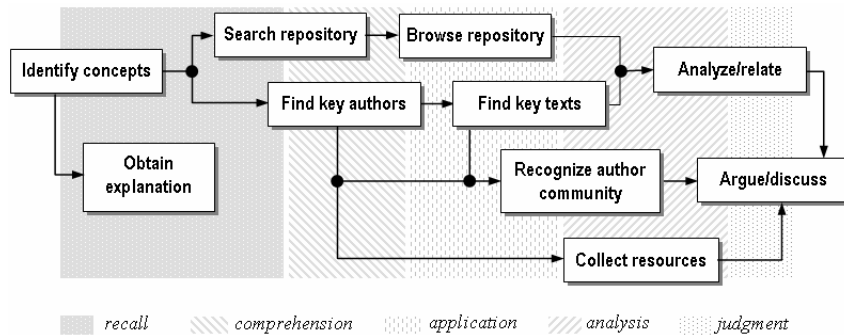


Fig. 2. Conceptual decomposition of the learning task (preparing literature review)

Although Fig. 2 contains paths linking activities, these should not be seen as a workflow in a traditional sense. The edges in fact describe shallow, typical conceptual dependencies, but one may not need to carry out all them to reach a judgment, for example. As our user study showed, there may be many configurations of this conceptual task model. We call these configurations *sensemaking paths*. A sensemaking path, along which learning tasks can be organized, is one way to elaborate and to formalize exploratory navigation. This form of exploration corresponds to the capability to refine an initial problem; e.g. in our case it helps reduce the number of retrieved papers if (say) teams of co-authors or broader communities of practices are explored as alternatives accessible from the original query.

The alternative paths differ in how the review task is interpreted – is it more about listing the prominent approaches or is it about comparing alternative approaches? Is it author- or topic-driven? Both interpretations are valid and accomplish the same aim. In our study, people tended to start with a literal (i.e. topic-driven) interpretation of the task. Users achieving better quality recognized that the search output was too large, coarse-grained and semantically generic, and opted for an alternative strategy; e.g.: (i) filtering the key authors and follow that trail to their works, or (ii) looking at topics conceptually close to the ambiguous original query and use them to alter the initial problem. The choice between alternative sensemaking paths is based upon pedagogic knowledge and guidance, which, in turn, enables a tutoring system to support the learner in performing a ‘lateral’ step from one sensemaking strategy to another; e.g. from a topic-driven to the author-driven one. It frames an open task, and it is orthogonal to the task decomposition that is common in many ITS.

Benefits of Semantic Web to a user’s interaction with learning resources may be seen in terms of improved interpretation of the navigation in a multi-dimensional space. At any point, there are alternative directions to take; each step triggers a specific outcome, has specific strengths and weaknesses. Some are about *deepening knowledge in one direction* (this is conceptually close to faceted views [15]) and other steps are ‘lateral’; i.e. moving from one view of the domain to another (similar to horizontal navigation [2]). Unlike specialized faceted or navigational tools, Semantic Web technology may tackle both alternatives. Hence, an outcome is likely to be a more robust and flexible support for the learner.

3. How to realize different sensemaking strategies

The notion of exploratory navigation is not novel [3, 4, 9]. What our work brings to the state of the art is an *open, extensible framework* and the capability to achieve *exploration by combining the specialized services*. The information retrieval services are rarely conceived as exploratory; mostly their aim is to access material needed for a given task. The ASPL frame-

work offers several entry points enabling the exploration by linking such simple services *and* interpreting their outcomes semantically. For example, entities such as authors or research areas in ASPL provide means to get started with the literature review. The role of ASPL is then to enrich these ‘gateways’ by offering services realizing the filtering, deepening and lateral dimensions, as mentioned in the previous section.

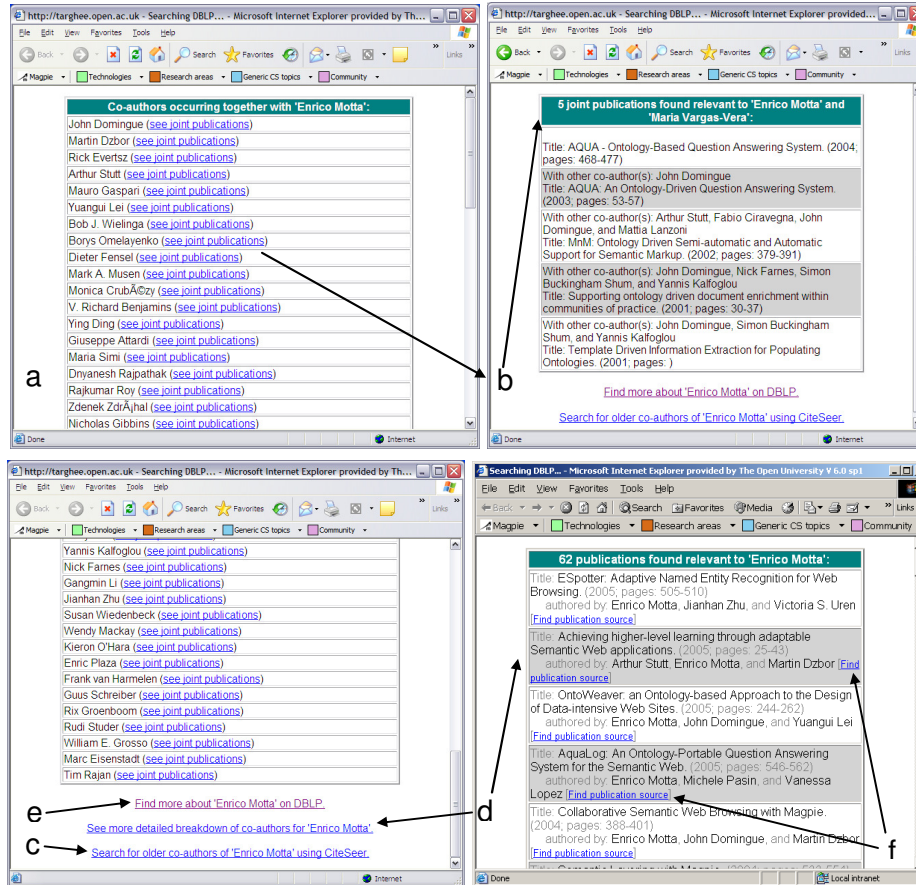


Fig. 3. Screenshots with different forms of ‘path narrowing’ as illustrations of sensemaking strategies

3.1 Strategy #1: deepening acquired knowledge

The filtering strategy was implemented as a refinement of the response to the query about publications satisfying a given keyword. On the cognitive level, what these refinements aimed at was to support ‘spotlight navigation’ – continuously narrowing a path towards something that can be further analyzed. This path narrowing strategy is not hierarchic or taxonomic; it lacks the prescriptive nature such structures. For example, if the learner requests a list of co-authors publishing with Enrico Motta (see Fig. 3a), the ASPL-v2 sensemaking support now offers to refine the initial list of co-authors (Fig. 3a) in several directions:

- filter Enrico’s publications to those co-authored with a specific person (Fig. 3b)
- look at this co-author community over time; e.g. older vs. newer clusters (Fig. 3c)
- view co-authorship in terms of publication clusters vs. individuals (Fig. 3d)
- explore the provenance and hence credibility of the co-author list (Fig. 3e)
- formulate the query for finding the source of a particular publication (see Fig. 3f)

As shown in Fig. 3 these alternative ways of continuing the exploration of the retrieved resources use different semantic relationships. For instance, for step in Fig. 3b we use a relation of co-authorship (applicable to a person); in Fig. 3c we deploy the property of publication year (applicable to papers). None of these is sophisticated on its own, but such independent chunks can be easily mined in documents or in databases. The added value of Semantic Web is inferring meaningful associations among such independent chunks.

3.2 Strategy #2: lateral navigation

On the other hand, the strategy of lateral navigation can be seen as opposite to the ‘spotlight browsing’ – a serendipitous and divergent exploration of different aspects of a particular collection. The purpose of modelling this strategy was not to narrow the exploratory path, but on the contrary, to open it up in a loosely related but different direction. This learning strategy responds to the question of how else a similar task can be achieved.

For example, assume the learner starts with retrieving publications using the keyword “hypertext”, and gets thousands of references assembled purely on the basis of containing given keyword. In ASPL-v1, the collection used in this manner was ACM Portal¹, which offered 40,000+ hits in response to such a query. What could be offered to the user at the point of receiving too large number of results are alternative approaches to the problem:

1. try expanding or changing “hypertext” e.g. with *related topics*, or
2. try shifting to *leading authors* in “hypertext” and possibly search for the publications of those towards the top of the list, or their co-authoring communities

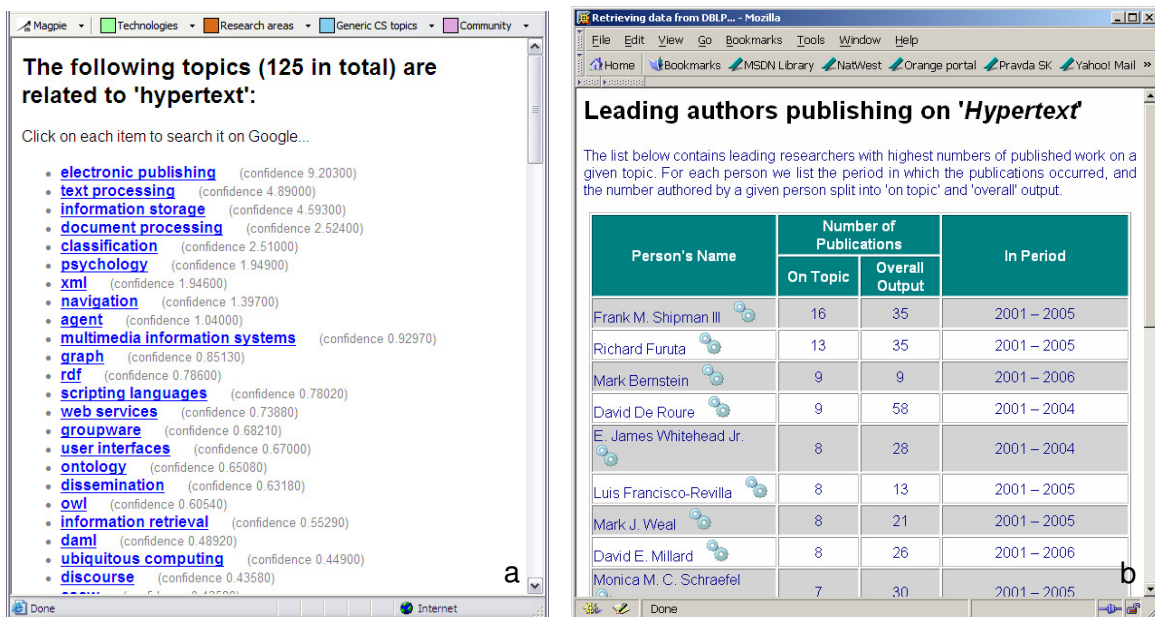


Fig. 4. Lateral navigation from a topic-based query: (a) to related sub-topics, and (b) to experts

The alternative lateral modifications of the original problem are depicted in Fig. 4: on the left we see a loosely related list of themes related to “hypertext”, while on the right is an orthogonal list of authors active in “hypertext”. Fig. 4a can be interpreted as: *Continue exploring main topic “hypertext” in the context of sub-topic (e.g.) “navigation”*. Moreover, one can also carry out more analytic or synthetic tasks, such as for example: *Compare the outcomes of “navigation” vs. “text processing” contexts of the main theme “hypertext”*. Similarly, by switching from topic-based navigation to authors, one actually reformulates the original retrieval problem using semantically close entities, which may, in turn, lead to different outcomes – a shortcut to these queries is expressed using the ‘cogs’ icon metaphor in Fig. 4b.

3.3 Benefits of an automated discovery of associations

Intelligent Tutoring Systems [10] and other tools supporting learners are active applications – mostly in the sense that they lead the learner through their repositories. Yet they largely operate on a closed set of resources and tend to use manually defined abstract relationships between concepts and resources [9]. The most popular link of this kind is ‘requires’ – as in

¹ <http://portal.acm.org>; both ‘Guide’ and ‘Digital Library’ showed a similar lack of discrimination.

“*Study of ontology-based annotation requires knowledge of RDF.*” Applying the ‘requires’ link transitively, it is possible to compute user paths through the resources and ensure each user follows a prescribed learning task. However, manual annotations are not scalable; they assume one path fits all user needs. Unfortunately, as the number of links increases, this approach reduces the feasibility of this type of systems.

Rather than tying the learner into one specific learning task, we see learning tasks as an optional element in a semantic system supporting the learners. Many learning tasks can be achieved by following several, often very distinct, *paths through the space of (learning) resources*. It is nearly impossible to formalize any one of these paths as the ideal execution of the learning task. Instead, different paths can be triggered by associations that happen to be useful at a particular moment. The relationship between tasks and paths is many to many – one task can be achieved by following several paths, and vice versa.

4. Learning and exploration: related research

The notion of exploratory user interaction with learning resources has been around for some time. In the pedagogical domain, for example, Laurillard [13] characterizes education and learning as a conversation of the learner with the problem and available resources – a conversation that facilitates exploration of the relevant knowledge space and creates a rich set of different kinds of associations between the chunks in the explored knowledge space.

Similar insights have been made by cognitive scientists studying skill acquisition in engineering or architectural design. For example, Schön [14] talks about reflective conversations of a practitioner with a design situation. These allow multiple ways of interpreting the situation and rely on associative frames. The theory of problem framing [6, 14] was developed on a rather abstract, cognitive level; nevertheless, its core is in recognizing and selecting associations to make sense of a situation. This, in turn, can be considered as the basis for a more operational, technical approach.

The Web community attempted to address the issue of how dynamic associations can be made in standard hypermedia. E.g. COHSE [4] has re-introduced the serendipity into navigating through the web documents using hyperlinks. Its hyperlinks were not restricted to hard-coded ones; they were to some extent independent of the actual web pages and could be inserted into a page as and when a need arose. This approach became known as open hypermedia. In educational hypermedia, benefits of horizontal (non-hierarchical) navigation in digital textbooks were analyzed; e.g. in [2]. They notice *lack of support for the horizontal links* because this mode of navigation is more resource-intensive than standard classification into a vertical taxonomy. Vertical content-based links represent the order, the plan; the horizontal, associative links are more serendipitous, interpreted, and may be subjective to a learner.

Nonetheless, one obstacle has always been the problem of specifying meaning. In the standard Web, this is normally implicit inside the mind of the resource provider. Semantic Web technologies, such as RDF or OWL², take a complementary stance by assuming that each relation or association can be committed to a formal semantic interpretation. Once an explicit semantic commitment (or annotation) is made, the relation acquires a specific meaning, which in turn enables distinguishing between (say) a person being active in an area and one area being similar to another. In short, the Semantic Web extends the notion of associative navigation by the fact that associations are not only established, but more importantly, interpreted.

5. Conclusions

Our study pursues the view of a learner’s interaction with resources on the Semantic Web as more than annotation, retrieval and subsequent browsing of semantic metadata. In order to

² See W3C recommendations <http://www.w3.org/TR/rdf-schema> and <http://www.w3.org/TR/owl-ref>, respectively.

apply semantic knowledge the re-designed ASPL-v2 used an exploratory approach to interact with distributed learning resources. Specifically we implemented two distinct modes of exploratory learning: (i) convergent, ‘spotlight’ browsing of semantically enriched resources [5], and (ii) divergent, ‘serendipitous’ browsing into an open web space [8].

As a source of information, knowledge and guidance, more and more used to support learning both formally and informally, the Web needs tools with the capacity to create semantic associations and intelligently use such associations e.g. to filter more familiar user interaction processes, such as search or data retrieval. The main advantage of the ASPL – annotation tied in with relationship composition is ultimately in reducing information overload and in increasing the smartness of the systems supporting learners.

Applying Semantic Web to construct multiple exploratory paths and attending to different aspects of the exploration, rather than to the individual nodes of the semantically enriched space, has several side effects. For instance, from the user experience viewpoint, the application becomes more flexible. A semantically enriched application does not confine its user to one specific activity or role. Another side effect is the dynamics of the semantic application. Ontology-driven solutions are often brittle; often based on closed worlds that enable reasoning solely about the known concepts. Linking the association discovery to the presentation overcomes this brittleness, and also avoids the knowledge acquisition bottleneck.

Acknowledgements

This work has been supported by the KnowledgeWeb and NeOn projects funded by the IST Framework VI of the European Commission’s grants no. FP6-507482 and FP6-027595.

References

- [1] Bloom, B.S., *A Taxonomy of Educational Objectives Handbook 1: Cognitive Domain*. 2nd ed. 1965, New York, US.
- [2] Brusilovsky, P. and Rizzo, R., *Map-Based Horizontal Navigation in Educational Hypertext*. *Journal of Digital Information*, 2002. **3**(1): p. 156.
- [3] Bush, V., *As we may think*. *The Atlantic Monthly*, 1945. **176**(1 (July)): p. 101-108.
- [4] Carr, L., Bechhofer, S., Goble, C., et al. *Conceptual Linking: Ontology-based Open Hypermedia*. In *10th Intl. WWW Conf.* 2001. Hong-Kong.
- [5] Collins, T., Mulholland, P., and Zdrahal, Z. *Semantic Browsing of Digital Collections*. In *Proc. of the 4th Intl. Semantic Web Conf.* 2005. Ireland. p. 127-141.
- [6] Dzbor, M. and Zdrahal, Z. *Design as interactions between problem framing and problem solving*. In *15th European Conference on AI (ECAI)*. 2002. Lyon, France. p. 210-214.
- [7] Dzbor, M., Motta, E., and Domingue, J. *Opening Up Magpie via Semantic Services*. In *Proc. of the 3rd Intl. Semantic Web Conf.* 2004. Japan. p. 635-649.
- [8] Dzbor, M., Motta, E., and Stutt, A., *Achieving higher-level learning through adaptable Semantic Web applications*. *Int. J. of Knowledge and Learning*, 2005. **1**(1/2): p. 25-43.
- [9] Eisenstadt, M., Price, B.A., and Domingue, J., *Software Visualization As A Pedagogical Tool*. *Instructional Science*, 1983. **21**: p. 335-365.
- [10] Frasson, C., Gauthier, G., and McCalla, G.I. *Intelligent Tutoring Systems*. In *Intl. Conf. on Intelligent Tutoring Systems (ITS)*. 1992: Springer-Verlag, Berlin.
- [11] Hayes-Roth, F., Waterman, D.A., and Lenat, D.B., eds. *Building Expert Systems*. 1983, Addison-Wesley: Massachusetts, US.
- [12] Hirschman, L. and Chinchor, N. *Named Entity Task Definition*. In *Proc. of the 7th Message Understanding Conf. (MUC-7)*. 1997.
- [13] Laurillard, D., *Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies*. 2nd ed. 2002, London, UK: RoutledgeFarmer.
- [14] Schön, D.A., *Reflective Practitioner - How professionals think in action*. 1983, USA: Basic Books, Inc.
- [15] Schraefel, M.C., Smith, D.A., Owens, A., et al. *The evolving mSpace platform: leveraging the Semantic Web on the Trail of the Memex*. In *Proc. of the Intl. Conf. on Hypertext*. 2005. Austria.
- [16] Zhu, J., Goncalves, A.L., Uren, V.S., et al. *Mining Web Data for Competency Management*. In *Proc. of the IEEE/WIC/ACM Conf. on Web Intelligence*. 2005. France.