Adaptive Personal Information Environment based on the Semantic Web

Thanyalak Maneewatthana, Gary Wills, Wendy Hall

Intelligence, Agents, Multimedia Group
School of Electronics and Computer Science
University of Southampton, SO17 1BJ
United Kingdom
{tm03r, gbw, wh}@ecs.soton.ac.uk

Personalised information systems aim to give the individual user support in accessing, retrieving and storing information. In order to support knowledge workers during their tasks of searching, locating and manipulating information, a system that provides information suitable for a particular user's needs, and that is also able to facilitate the sharing and reuse information, is essential. This paper presents Adaptive Personal Information Environment (a-PIE); a service-oriented framework using Open Hypermedia and Semantic Web technologies to provide a personalised web-based system. a-PIE models the information structures (data and links) and context as Fundamental Open Hypermedia Model (FOHM) structures which are manipulated by using the Auld Linky contextual link service. a-PIE provides an information environment that enables users to search an information space based on ontologically defined domain concepts. The users can add and manipulate (edit, delete, comment, etc.) interesting data, or parts of information structures, into their information space leaving the original published data, or information structures, unchanged. a-PIE facilitates the shareability and reusability of knowledge according to users' requirements.

1. Introduction

Personalised information systems aim at giving the individual user support in accessing, retrieving and storing information. Knowledge management and the associated tools aim to provide an environment where people may create, learn, share, use and reuse knowledge, for the benefit of the organisation, the people who work in it, and the organisation's customers. However, instead of helping users, many systems are just increasing the information overload. Adapting text and sets of relationships of information, or contents, to the needs of individual users greatly enhances navigation and comprehension of information spaces.

The Semantic Web can be used to organise information in concept structures, while web services allow the encapsulation of heterogeneous knowledge and modularization of the architecture. In addition, web services also support dynamic and

shareable frameworks for automated adaptation [De Bra, *et al.* 2004]. An ontology can be used to enrich the semantics of data and information structures to aid the process of information searching (use and reuse of information).

In this paper we propose an Adaptive Personal Information Environment system (a-PIE), a service-oriented framework for reusability and shareability of information. a-PIE aims to provide an adaptive web-based system in which members of the community, or organisation, are able to browse information tailored to their needs, store content of interest into their own information repository, and are able to augment with metadata for reuse.

The background of related technologies; Open Hypermedia, Fundamental Open Hypermedia Model, Auld Linky, Semantic Web, and web services are briefly described in the next section. Then, a system overview of a-PIE is presented, focusing on the support for personalisation, reusability and shareability of information. Finally, some conclusions and future work are presented.

2. Background

In this section we present the motivation for this research, and briefly describe the technologies used. A brief scenario helps explain the motivation behind the design of a-PIE. A user in a virtual community is looking / searching for information with a keyword from a community web site. The user finds interesting pieces of information, and would like to keep it for further reference in his personal repository. On occasions, he would like to add annotation to particular information snippets, and would like to record the context before storing it in the personal repository. The user would also like to share the annotation with others in the community. As this user is an expert in this domain, he adds the annotation to the information snippets, and marks these as being useful for experts, intermediate, or beginners. So other people, who have authorisation by this user, will see information with annotation according to their users' profile (knowledge background, expertise).

In Open Hypermedia Systems, links are considered as first-class objects. These entities are manipulated separately from hypermedia documents and stored independently in link databases (linkbases). Links and data are then added to hypermedia documents by means of a link service. The advantages of the link service approach are that links can be created, added, and edited without affecting the original document. By moving hyperlinks out of documents and into link databases, the relationships between documents are separated from the document content [Berners-Lee, *et al.* 2001]. Therefore, the collection of documents becomes more maintainable, quicker to produce and easier to reuse. Changes in target documents only require changes in linkbases. Using external linkbases enables different sets of links, which can represent different contextual dimensions, to present link personalisation based on the user preference and their levels of expertise [Longpradit, *et al.* 2006]. In addition, individuals and groups can maintain their own personal link databases.

The Fundamental Open Hypermedia Model (FOHM) [Michaelides, *et al.* 2001] is a protocol for open hypermedia with additional context-awareness features. It is a data model for expressing hyperstructure by representing associations between data. Auld

Linky is a context based link server which supplies links from specified linkbases by parsing FOHM structures. The four essential components of a FOHM structure are Data Objects, Associations, References and Bindings:

- *Data objects*: Data objects are wrappers for any piece of data that lies outside of the scope of the FOHM model.
- *Associations*: Associations are structures that represent relationships between Data objects.
- *References*: Reference objects are used to point at Data objects or at parts of Data objects.
- *Bindings*: Bindings specify the attributes of the connection between Association and Data objects through Reference objects.

FOHM also provides Context objects, modifier objects, which can be attached to any part of the FOHM structure. Figure 2 illustrates an example of the FOHM structure.

- *Context objects* define conditions for the visibility of particular objects. These objects can also define the description type, and are used by Auld Linky to distinguish which bindings should be returned to the user.

Web Services are software systems which provide standard ways to interoperate between various existing applications run on heterogeneous resources or frameworks. Web Services have been designed to wrap existing applications, and expose them using an interface described in machine-processable format: Web Services Description Language (WSDL). Other systems can interact with Web Services using Simple Object Access Protocol (SOAP) messages, or XML-RPC protocol, for example. The use of Web Services can be described as loosely coupled, reusable software components, which can be orchestrated on the fly [Booth, *et al*. 2004].

The key idea of the Semantic Web is to have data defined and linked in such a way that its meaning is explicitly interpretable by software processes rather than just being implicitly interpretable by humans [Berners-Lee, *et al.* 2001]. The Semantic Web can represent knowledge, including defining ontologies as metadata of resources. An ontology is a means to describe formally a shared understanding, and capture knowledge for a particular domain [Gruber 1993]. It will be necessary to annotate web resources with metadata to provide some indication of its content.

Community portals are information portals designed to support and facilitate a community of interest. They allow members of a community to contribute information either by submitting, or posting the information to the system.

3. System Overview

The Adaptive Personal Information Environment (a-PIE) aims to provide a system in which members of the community are able to browse information suitable to their particular needs, identify and store FOHM structures in their own information repository, which users may enhance prior to reuse; thereby, enabling the sharing and reusing of structures and data. In addition, a-PIE enhances these facilities by using

ontologies to define the associations and facilitate the interoperability between knowledge components and system.

a-PIE consists of a number of service-oriented components. The domain concept service provides the relevant concept. The user model service updates user model. The structure and data service manipulates data and structures (associations) as FOHM objects from linkbases through the contextual link server (Auld Linky). The user service or adaptive engine provides the facilities for reconciling the data content, FOHM structures, and user model, to present the individualised document to the user through a web browser.

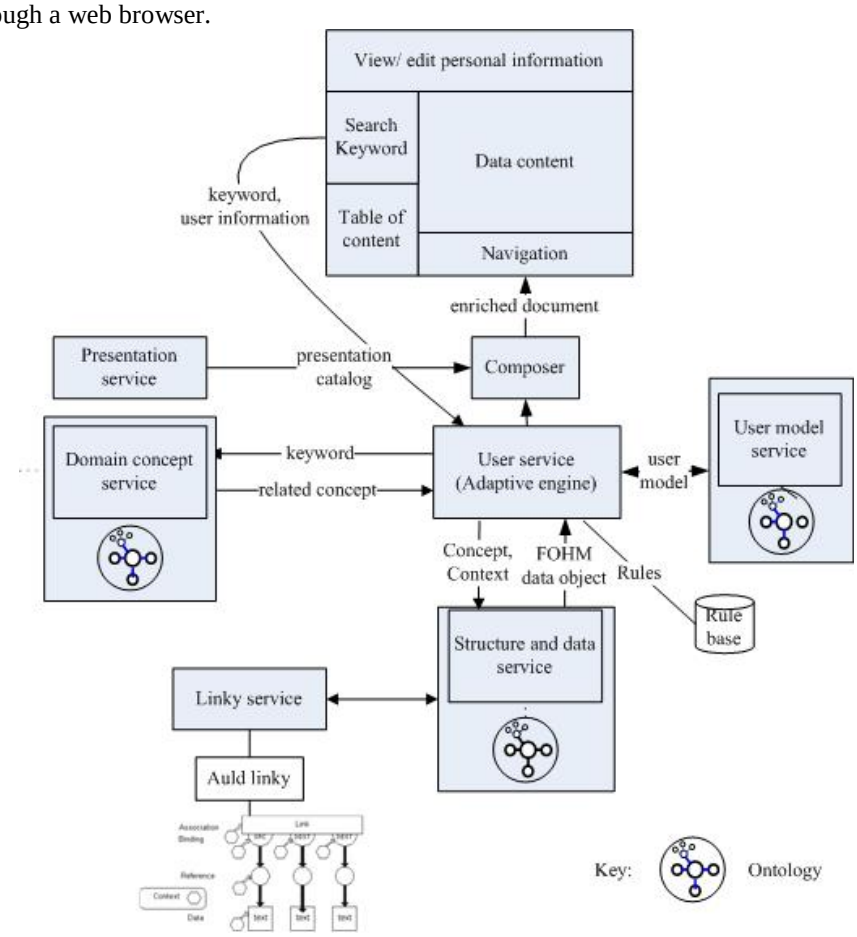

Figure 1 a-PIE System Architecture

Figure 1 illustrates the system architecture of a-PIE. The functionality of the system is made available to software agents through a Web Service interface (WSDL), and to end-users by means of a Web browser. The input into the system is a collection of data objects from a user, and the output is an enriched document customised to a particular user's needs.

The following section presents how a-PIE supports data creation, storage, personalisation, reusability and sharing of information.

3.1 Data creation, storage and personalisation

In order to promote reusability and sharing of information, a-PIE separates the main system components; domain concept, structure and data, user model, and presentation.

- *Domain concept service*: manipulates domain concept used on organisation web sites. Simple Knowledge Organisation System (SKOS) [SKOSCoreGuide 2005] is used to express the basic structure of concept.
- *Structure and data service*: manipulates structure and data represented on the web. Data are represented in the form of FOHM Data objects, while structures connects FOHM data objects defined in XML format as a series of FOHM Association structures. Contexts, represented by FOHM Context objects, can be attached to the Data object, or Association object, for describing the context in which the data item, or association, is visible (or hidden) from the user.
- *User model service:* manipulates user profiles, such as background knowledge, preference, information about user, set of topics, or concepts, where the user has expressed an interest.
- *Presentation service*: provides the display and machine-related information, such as the colour schemes for resource presentation.
- *User service*: is an intermediary engine to integrate every service to work in collaboration.

Structures that can be represented by FOHM are:

- *Navigational link*: The navigational link is an association with typed data items, source, destination or bi-directional locations.
- *Tour*: The tour is an association that represents an ordered set of objects.

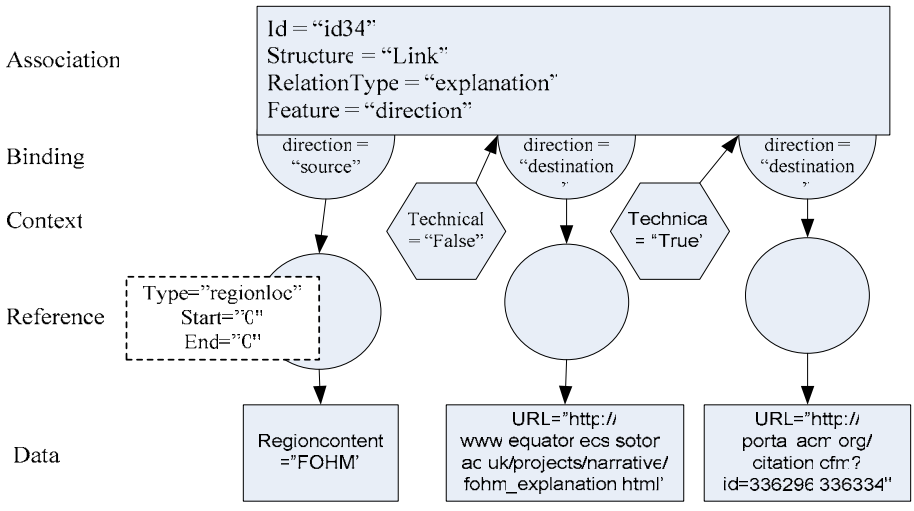

Figure 2 A simple FOHM Navigation Link

Personalisation
According to Brusilovsky's taxonomy [Brusilovsky 2001] of adaptive hypermedia techniques – adaptive navigation support and presentation, FOHM structures can be combined to implement a range of these techniques [Bailey, *et al.* 2002]. An example of FOHM structure is shown in Figure 2. This is a link with one source (the location with the word "FOHM") and two destinations (with urls). Both destinations explain "FOHM", the first with none technical information, while the second with technical detail. If the structure was loaded into Auld Linky and queried using this context, then Auld Linky would remove the inappropriate destination.

a-PIE provides adaptive hypermedia support through the use of Auld Linky [Michaelides, *et al.* 2001] as a contextual link server to integrate FOHM structure and data objects according to the context. The context stores value-pairs of concept and level of difficulty of the particular FOHM objects. Therefore, the system can produce the information suitable to users' needs as a personalised web-based system. Once data is made available by the organisation and published on the web site, other users can then use, reuse, or browse, based on user preferences specified in user profiles.

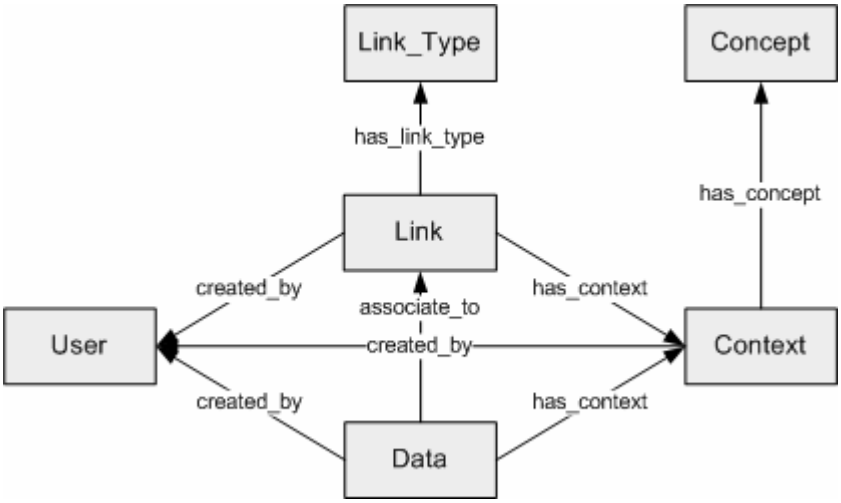

Figure 3 Ontology of a-PIE

3.2 Sharing, reusing and enriching the information

Hypermedia link-types are knowledge relations [De Troyer and Casteleyn 2003], and a set of hypermedia link types may be represented as an ontology. In a-PIE, the FOHM models are used to represent hypermedia contents and structures. The ontology of link types has created based on relationships suggested by Bieber and Yoo [Bieber and Yoo 2000].

Sharing and reuse of information are integral aspects of the Semantic Web. In

a-PIE, the ontology is based on Semantic Web technology standards (RDF [Klyne and Carroll 2004] /OWL [McGuinness and Harmelen 2004]), and is the backbone of the system. The ontologies represent relationships of domain concepts. The ontologies are also used to enrich links and data content, and to enable other users to share and reuse the content, or structure. The user can add data (Association / Data object) and their supplementary information, and add the context (Context object) to particular data objects existing on the Web. An individual user can add the context, and define whether the context can be seen, shared, or reused (i.e. edited, or deleted) by other users. The user who has permission can then view the context, as specified by the original user (user's preference). The users might also use their own domain concept and context for categorising, or describing, the information. The ontology of the a-PIE is shown in Figure 3.

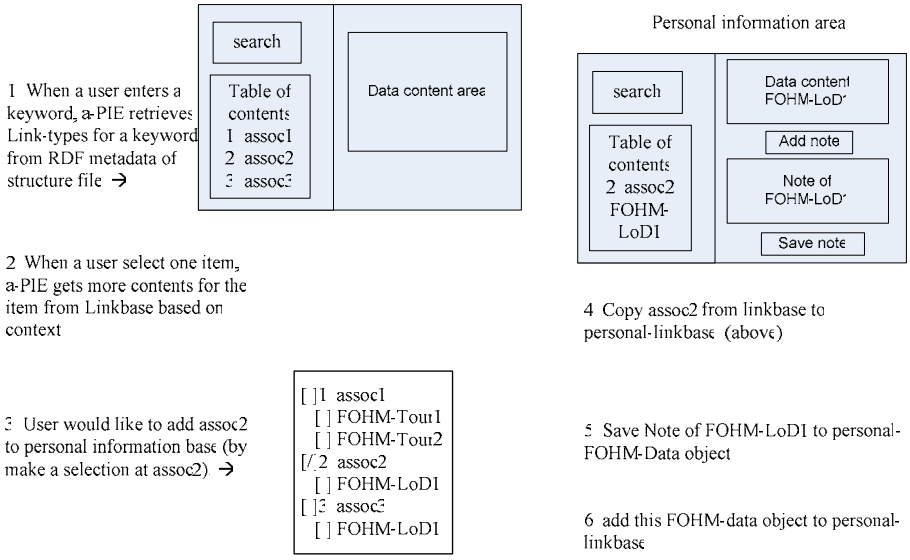

Figure 4 Simple processes for adding note to particular data item

Figure 4 illustrates a simple scenario on how information can be used and reused in a-PIE. The browser is divided, apart from menu and navigation areas, into three main areas; search, table of contents and data content areas. The data content area displays the data content stored as FOHM-data objects. In this scenario, a user initiates a query, the results are then returned (after the query processing) to the user according to their profile (the level of detail and types of user). The users add notes, or comments, to a particular piece of information.

The processes for browsing, adding structures the user is interested in, and adding more data content (such as notes or comments) to a particular item, are described in the following sequences of operations (see Figure 4).

1. A user enters a keyword into a search box. Then, the system will find the associations that have relationships related to this concept from the metadata of FOHM associations stored in a RDF file, and then will show the results as table of contents.

2. When the user selects the association they are interested in, the system will get more structure of contents from the linkbase, represented as FOHM associations.
3. The data content stored in FOHM data object will be displayed in the data content area if the user clicks on any item in table of contents. In addition, the user is able to store any information to their personal information area, simply by selecting a particular structure of information, and the appropriate option to include it in their personal information area.
4. The system will manage the storing of the selected association to the personal linkbase.
5. The user is then able to add more data such as notes, or comments, to the data content, which exists on the Web, in the personal information space, while the original data content remains unchanged.
6. The system will save all modification to the linkbases, and the data content that is represented in FOHM data object, in the personal information space.

3.3 Implementation

a-PIE is prototyped in Java as a web application under Apache Tomcat, using Auld Linky version 0.72 and Apache Axis. Apache Axis[1] is a java platform for creating and deploying Web Services applications. Java offers various advantages in comparison to other languages, in that it can be used under different operating systems. Another important consideration is the availability of one of the most advanced frameworks to build Semantic Web applications, including a rule-based inference engine which is Jena[2]. a-PIE uses Jena, an open-source project, to manipulate RDF models, and for a set of limited reasoning features.

The metadata of FOHM associations and FOHM data objects are represented in RDF, while the ontologies for concepts are represented in RDF/OWL.

4. Related work

The Semantic Community Portal approach, provided by the Semantic Web Environmental Directory (SWED) [SWED 2004], aims to creating and maintaining web-based community information resources. The portal provides the system which enable third-party to reuse the information. The outstanding point for this directory management is the separation of data creation and storage from the publication. The Semantic Web technologies are used to enrich the basic SWED data records by various organisations. The data itself, written in RDF, is created and stored by the organisations on their own web sites. The directory organisation harvests the RDF files of organisations that are relevant to their particular area of interest. The directory organisation is able to add some additional specialist information themselves, or they might use their own vocabulary for categorising, or describing, the information. Therefore, although this approach supports the reusability of information for the

[1] http://ws.apache.org/axis/
[2] http://jena.sourceforge.net

knowledge workers in organisations, it does not provide the reuse of information for an individual user, and there are no adaptation supports.

De Bra *et al* propose a *modular adaptive* hypermedia architecture [Chepegin, *et al.* 2003] for the next step of adaptive Web-based systems. In this approach, four main system components are separated. These different levels are the domain model, application model, user model and adaptation model. Each system component is implemented as a Web service and communicates through service invocations. Ontologies are applied to define and unify the system's vocabulary, and describe knowledge of each system service. It enables standardization and formalization of meaning, and enables reusability and interoperability. The modular adaptive architecture is based on existing Web service framework like the IRS-II [Motta, *et al.* 2003]. In this approach, the service uses UPML to specify reusability in knowledge-based systems. An advantage of the modularized architecture is that it enhances very high degree of flexibility [De Bra, *et al.* 2004].

The purpose of COHSE [Carr, *et al.* 2002] is the integration of an open hypermedia architecture, especially the Distributed Link Service, with ontological services to provide an architecture for the Semantic Web and the linking, based on the concepts that appear in Web pages. Links are separated from the documents and manipulated separately from documents content. The documents are linked according to the metadata annotated to documents. The COHSE approach annotates the documents based on description logic and augments the documents with the annotations at browsing or reading time. However, COHSE is not proposed to be an AH system. Therefore, there is no adaptation of the links and contents, nor engagement of user profiles to support personalisation or adaptation of the presented contents and links.

5. Summary

This paper describes a semantic adaptive information environment approach to support knowledge workers. The adaptive Personal Information Environment (a-PIE) based on the Semantic Web is proposed. The advantages of this approach are not only providing the adaptive information for particular users' needs, but also the reusability and shareability of information for an individual. a-PIE provides an information environment for users in a community, or organisation, where the user can browse, or search information, based on domain concepts defined by ontologies. The users are also able to manipulate their own information space by adding, or deleting data, or parts of information structures, into their own information space. In addition, they can add personal information such as comments, or notes, to the existing data, or information structures; while the original data, or information structures, published remain unchanged. Moreover, it is possible to provide multiple aggregations and views of the same data in different contexts.

The adaptation, reusability and shareability of knowledge components in this system is achieved by using Semantic Web technology, and by storing separately the data, information structures, domain concept, context, presentation and user information models. The data, information structure and context are represented by FOHM object models, and manipulated by Auld Linky, a contextual link service.

FOHM object models represent hyperstructure with context explicitly. Ontologies are used to define common explicit relationships for domain concepts and to enrich data, information structures, presentation and user models. The ontology is able to facilitate sharing of information chunks between several users and reuse of information chunks in different contexts. Moreover, the service-oriented framework is used to provide the loosely coupled and reusable software components.

The current stage of a-PIE is in the integration phase. Future work will be looking at a formal evaluation of the a-PIE in order to establish what is the extent and limits of this framework proposed.

6 Acknowledgement

Our thanks to The Royal Thai Government for funding this work.

7 References

[Bailey, *et al.* 2002] Bailey, C., Hall, W., Millard, D., and Weal, M., 2002 "Towards Open Adaptive Hypermedia," In Proceedings of the second International Conference on Adaptive Hypermedia and Adaptive Web-Based System, Malaga, Spain, pp. 36-46.

[Berners-Lee, *et al.* 2001] Berners-Lee, T., Hendler, J., and Lassila, O., 2001, "The Semantic Web," *Scientific American*, Vol. 279, No. 5, pp. 35-43.

[Bieber and Yoo 2000] Bieber, M. and Yoo, J., 2000 "Towards a Relationship Navigation Analysis," In Proceedings of the thirty-third Hawaii International Conference on System Sciences, pp. 60-66, http://portal.acm.org/citation.cfm?id=820317.

[Booth, *et al.* 2004] Booth, D., Haas, H., McCabe, F., Newcomer, E., Champion, M., Ferris, C., and Orchard, D., "Web Services Architecture W3c Working Group Note", http://www.w3.org/TR/ws-arch/, accessed in June 15.

[Brusilovsky 2001] Brusilovsky, P., 2001, "Adaptive Hypermedia," *User Modeling and User Adapted Interaction*, Vol. 11, pp. 87-110.

[Carr, *et al.* 2002] Carr, L., Kampa, S., De Roure, D., Hall, W., Bechhofer, S., Goble, C. G., and Horan, B., 2002 "Ontological Linking: Motivation and Analysis," In Proceedings of the Information and Knowledge Management (CIKM) Conference.

[Chepegin, *et al.* 2003] Chepegin, V., Aroyo, L., De Bra, P., and Houben, G.-J., 2003 "Chime: Service-Oriented Framework for Adaptive Web-Based Systems," In Proceedings of the Conferentie Informatiewetenschap, Eindhoven, pp. 29-36, http://wwwis.win.tue.nl/infwet03/proceedings/3/.

[De Bra, *et al.* 2004] De Bra, P., Aroyo, L., and Chepegin, V., 2004, "The Next Big Thing: Adaptive Web-Based Systems," *Journal of Digital Information*, Vol. 5, No. 1, http://jodi.tamu.edu/Articles/v05/i01/DeBra/#refs

[De Troyer and Casteleyn 2003] De Troyer, O. and Casteleyn, S., 2003, "Exploiting Link Types During the Conceptual Design of Web Sites," *International Journal of Web Engineering Technology*, Vol. 1, No. 1, pp. 17-40.

[Gruber 1993] Gruber, T. R., 1993, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition*, Vol. 6, No. 2, pp. 199-221.

[Klyne and Carroll 2004] Klyne, G. and Carroll, J. J., "Resource Description Framework (Rdf): Concepts and Abstract Syntax", http://www.w3.org/TR/rdf-concepts/, accessed in October 2005.

[Longpradit, et al. 2006] Longpradit, P., Bailey, C.P, Hall, W., and Wills, G., "Personalised Navigation System with Multidimensional Linkbases ", In Proceedings of the fourth International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, AH2006, Dublin, Ireland.

[McGuinness and Harmelen 2004] McGuinness, D. L. and Harmelen, F. v., "Owl Web Ontology Language Overview ", http://www.w3.org/TR/2004/REC-owl-features-20040210/ accessed in October 2005.

[Michaelides, *et al.* 2001] Michaelides, D. T., Millard, D. E., Weal, M. J., and De Roure, D. C., 2001 "Auld Leaky: A Contextual Open Hypermedia Link Server," In Proceedings of the seventh Workshop on Open Hypermedia Systems ACM Hypertext Conference, Aarhus, Denmark.

[Motta, *et al.* 2003] Motta, E., Domingue, J., Cabral, L., and Gaspari, M., 2003 "IRS-II: A Framework and Infrastructure for Semantic Web Services," In Proceedings of the second International Semantic Web Conference 2003 (ISWC 2003), Sanibel Island, Florida, USA.

[SKOSCoreGuide 2005] SKOSCoreGuide, "Skos Core Guide", http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20050510/, accessed in October 2005.

[SWED 2004] SWED, "Semantic Web Environmental Directory", http://www.swed.org.uk/swed/swed_technical_resources.htm, accessed in October 2005.