Adaptive Personal Information Environment based on 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

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 an adaptive web-based system. a-PIE models the information structures (data and links), context and behaviour 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 (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

Knowledge management and the associated tools aim to provide an environment in which 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 [1].

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 a system in which members of the community or organisation are able to browse information tailored to their needs, store relationships to content of interest in their own information repository, and are able to augment the relationships 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 adaptation, reusability and shareability of information. Finally some conclusions and future work are presented.

2. Background

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 [4]. Therefore the collection of documents becomes more maintainable, quicker to produce and easier to reuse. Changes in target documents only require changes in linkbases. So using external linkbases enables different sets of links over the same content for different audiences and tasks. In addition, individuals and groups can maintain their own personal link databases.

The Fundamental Open Hypermedia Model (FOHM) [2] 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*: 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.

FOHM also provides two modifier objects, which can be attached to any part of the FOHM structure, these are Behaviour and Context. Behaviour objects are used by client applications, whereas Context objects define conditions for the visibility of particular objects. Context objects 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. The use of Web Services can be described as loosely coupled, reusable software components, which can be orchestrated on the fly [3].

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 [4]. 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 [5]. 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 enable the sharing and reuse of structures and data. a-PIE further enhances these functionalities by using ontologies to define the associations and facilitate interoperate between knowledge components.

a-PIE consists of several services. The domain concept service provides the relevant concept. The user model service updates user model. The data item and association 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 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 through 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, adaptation, reusability and sharing of information.

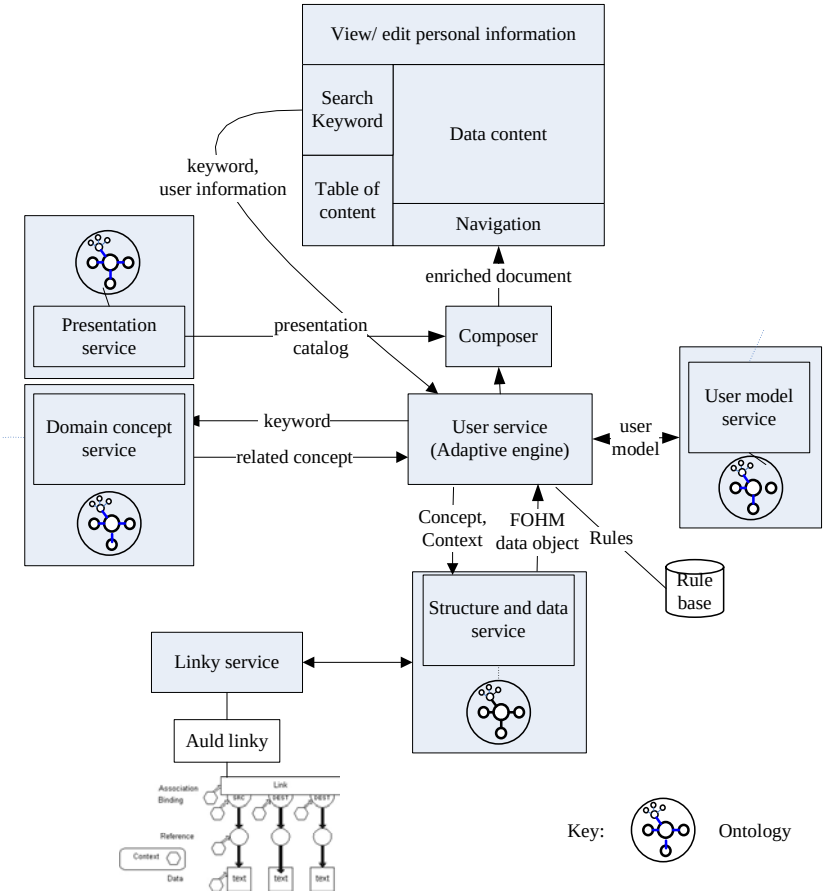

Figure 1 a-PIE System Architecture

3.1 Data creation, storage and adaptation

In order to promote reusability and sharing of information, the system separates these components; domain concept, data, structure, presentation, user information, context and behaviour. Each model has an ontology as a means for interoperation.

- *Domain concept*: represents the basic structure of concept schemes. Simple Knowledge Organisation System (SKOS) [6] is used to express the basic structure of concept.

- *Domain data model*: represents the data created and stored by individual or organisations on their own web sites, in the form of FOHM Data objects.
- *Structure model*: connects FOHM Data objects defined in XML format as a series of FOHM Association structures. FOHM Association structures can be used to model several hypermedia structures; Navigational Link, Tour, Level of Detail, and Concept.
- *User model:* represents user-related information, such as background knowledge, preference, and information about user.
- *Context model*: represented by FOHM Context objects which can be attached to a Data or Association object for describing the context in which the data item or association is visible (or hidden) from the user.
- *Behaviour model*: represented by FOHM behaviour objects which can be attached to a Data or Association object. Behaviour objects describe an action that occurs as a result of an event.
- *Presentation model*: display and machine-related information, such as the colour schemes for resource presentation.

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.
- *Level of Detail*: The Level of Detail structure represents an ordered sequence of objects, where each object represents similar conceptual information with increasing of complexity.
- *Concept:* The concept is a collection of objects, such as text, image, or audio, which represent the similar conceptual information.

a-PIE provides adaptive hypermedia support through the use of Auld Linky [2] as a contextual link server to integrate FOHM structure and data objects according to the context. From Brusilovsky's taxonomy [7] of adaptive hypermedia techniques; adaptive navigation and presentation, FOHM structures can be combined to implement a range of these techniques [8]. 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.

Therefore, this system can produce the information suitable to users' needs as an adaptive web-based system. Once data is made available by the organisation and published through the web site, other users can use, reuse, or browse. This means that anyone with an Internet connection can use and reuse the information.

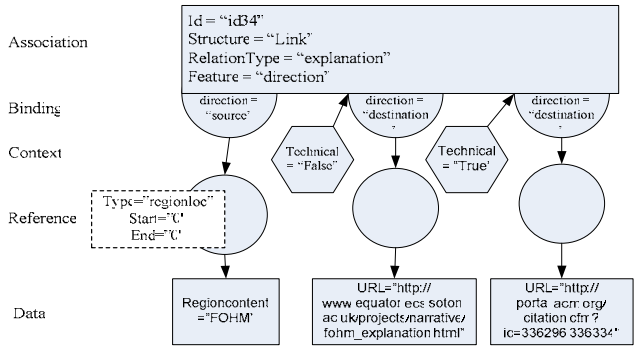

Figure 2 A simple FOHM Navigation Link

3.2 Sharing, reusing and enriching the information

In Open Hypermedia links are first-class objects and manipulated independently. Open Hypermedia makes links (associations) between different pieces of information (e.g., images or text). In a-PIE, the FOHM models are used as a protocol from open hypermedia. Hypermedia link-types describe the associations. Therefore, hypermedia link-types are knowledge relations [9], and a set of hypermedia link types may be represented as an ontology. In a-PIE, an ontology of link types has created based on relationships suggested by Bieber and Yoo [10].

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 [11]/OWL [12]) 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 or organisations to share and reuse the content or structure of the FOHM representation. In particular, each user in a community or organisation can browse and search the site. In addition, the user can; add data, add additional specialist information, and change the context or behaviour of the FOHM structures. The users might also use their own domain concept, context or behaviour for categorising or describing the information.

Figure 3 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, after processing the results are returned 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 sequence of operations.

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 metadata of

FOHM associations stored in a RDF file, and then show the results as table of contents.

2. When the user selects the association they are interested in, the system will get more s*tructure 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 (as FOHM structures), simply by selecting a particular structure of information and the appropriate option to add it to 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 in the personal information space, while the original remains unchanged.
6. The system will save all modification to the linkbases and data content which is represented in FOHM data object, in the personal information space.

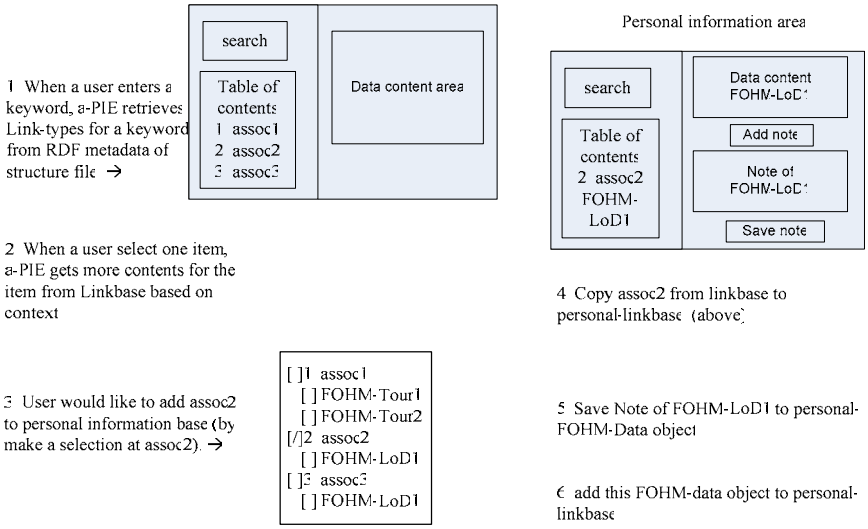

Figure 3 Simple processes for adding note to particular data item

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. Firstly, it can be used under different operating systems. Another important consideration is the availability of one of the most advanced frameworks to

[1] http://ws.apache.org/axis/

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 enriched documents are built by transformation from patterns catalogue through the presentation service. The metadata of FOHM associations and FOHM data objects are represented in RDF, while the ontologies for every service are represented in RDF/OWL.

4. Related work

The Semantic Community Portal approach, provided by the Semantic Web Environmental Directory (SWED) [13], aims to overcome the limitations and problems with existing approaches 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, this approach supports the reusability of information for the knowledge workers in organisations, not for an individual, and no adaptation supports.

The purpose of COHSE [14] is the integration of an open hypermedia architecture, especially the Distributed Link Service, with ontological services to support an architecture for the Semantic Web and provide the linking based on the concepts that appear in Web pages. Therefore, links are separated from the documents and links are 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, there is no adaptation of the links and contents in this approach. In addition, COHSE does not provide the reusability of links for knowledge workers.

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 sharebility of information for an individual. a-PIE provides an information

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

environment for users in a community or organisation in which to 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, behaviour, presentation and user information models. The data, information structure, context and behaviour are represented by FOHM object models and manipulated by using Auld Linky, a contextual link service. Ontologies are used to define common explicit relationships for domain concepts and to enrich data, information structures, presentation and user models. Moreover, the service-oriented framework is used to provide the loosely coupled and reusable software components.

Future work will focus on developing the ways for each community to create each model; data, structure, context, presentation and behaviour based on a service-oriented framework.

6 Acknowledgement

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

7 References

[1] P. De Bra, L. Aroyo, and V. Chepegin, "The Next Big Thing: Adaptive Web-Based Systems," *Journal of Digital Information*, vol. 5, 2004.

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

[3] D. Booth, H. Haas, F. McCabe, E. Newcomer, M. Champion, C. Ferris, and D. Orchard, "Web Services Architecture W3C Working Group Note," vol. 2005: W3C, 2004.

[4] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific American*, vol. 279, pp. 35-43, 2001.

[5] T. R. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition*, vol. 6, pp. 199-221, 1993.

[6] "SKOS Core Guide," vol. 2005, A. Miles and D. Brickley, Eds.: World Wide Web Consortium, 2005.

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

[8] C. Bailey, W. Hall, D. Millard, and M. Weal, "Towards Open Adaptive Hypermedia," presented at Proceedings of the second International Conference on Adaptive Hypermedia and Adaptive Web-Based System, Malaga, Spain, 2002.
[9] O. De Troyer and S. Casteleyn, "Exploiting Link Types during the Conceptual Design of Web Sites," *International Journal of Web Engineering Technology*, vol. 1, pp. 17-40, 2003.
[10] M. Bieber and J. Yoo, "Towards a Relationship Navigation Analysis," presented at Proceedings of the thirty-third Hawaii International Conference on System Sciences, 2000.
[11] RDFConcept, "Resource Description Framework (RDF) Concepts and Abstract Syntax," G. Klyne and J. Carroll, Eds.: World Wide Web Consortium, 2004.
[12] "OWL Web Ontology Language Overview," vol. 2005 June, D. L. McGuinness and F. v. Harmelen, Eds.: World Wide Web Consortium, 2004.
[13] "Semantic Web Environmental Directory," http://www.swed.org.uk, 2004.
[14] L. Carr, S. Kampa, D. De Roure, W. Hall, S. Bechhofer, C. G. Goble, and B. Horan, "Ontological Linking: Motivation and Analysis," *CIKM*, 2002.