

# Rethinking Web Design Models: Requirements for Addressing the Content

Arouna Woukeu, Leslie Carr, Gary Wills, Wendy Hall

Intelligence, Agent, Multimedia Group

Department of Electronics and Computer Science

University of Southampton, SO17 1BJ, UK

Tel: +44 (0)23 8059 3255

{[aw1](#), [lac](#), [gbw](#), [wh](#)}@ecs.soton.ac.uk

Technical Report Number: ECSTR-IAM03 -002

ISBN: **0854327940**

© University of Southampton

## Abstract

The objective of hypermedia design models is to produce a well-organised web site. The organisation is undertaken at the level of a particular building-block – an abstract data unit which may match a frame, paragraph or region on a Web page. The increasing sophistication of these models allows the designer to deal with interaction and personalisation, but precludes one of the basic features of hypertext – the text itself. This paper argues that this oversight remains a fundamental problem because the component of content production for many web sites is not an abstract data unit but the concepts embedded in the paragraphs, sentences and words of the content regions. Consequently there is a gap between the organisation of material and the origination of material that is not well-addressed by current design methods. The paper considers the problem of concept modelling in the Semantic Web, its implementation in various hypertext environments and whether this approach can inform the current generation of hypermedia design models.

## General Terms

Documentation, Design.

## Keywords

Web Design Methods, Hypermedia Design, Open Hypermedia, Semantic Web.

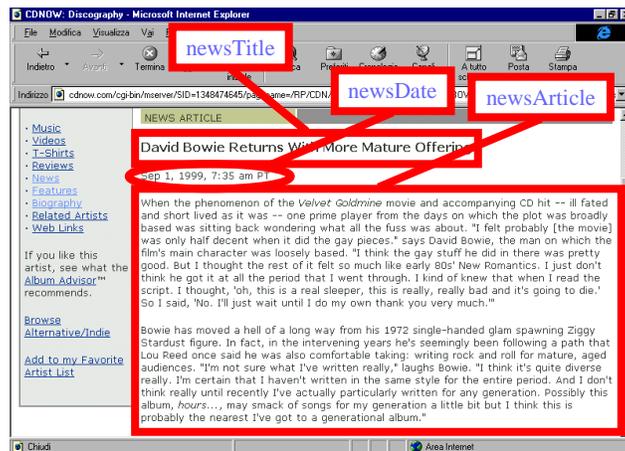
## 1 Introduction

Hypermedia design models have been principally aimed at public, data-rich web sites [9] which are part of the so-called *deep web*, that is, those sites which are visible manifestations of large databases. Although these have been typically characterized by e-commerce-style catalogs such as Amazon, CD-Now or Google, it is the case that content- and community- oriented digital library and portal sites are increasingly being dealt with [10].

These kinds of sites provide different kinds of information for the user in the form of more complex content (*e.g.* articles, mail messages) rather than easily-processed data. However, even the ‘data-oriented’ Web sites show the challenge of content – consider the web page shown in

figure 1 taken from (an original version of) the WebML tutorial<sup>1</sup>. It shows how the design method helps to identify the *type* of the information displayed to the user, but it stops short of modelling the significant concepts and knowledge fragments, which compose the information itself (e.g. the named entities, events and activities). Consequently the designer can only take advantage of the fact that this page contains a “news article”, and cannot assist the visitor to the page who is interested in the fact that it is *about* David Bowie and may want to follow up some of the facts that are mentioned. (What is the connection with Lou Reed? What was New Romanticism?)

As well as considering their (deep web) public information provision, organisations have also amassed large intranets (private, multimedia Web sites) in an attempt to capture their corporate knowledge [23] and are increasingly concerned with effective information access as a part of the imperative of knowledge management [39]. These collections have typically been gradually amassed in an unsystematic fashion, potentially from many parts of an organisation over a long period of time (relative to the structure of the organisation itself). Consequently, they will not be the subject of a single, disciplined design process, and may exhibit unconstrained and inconsistent metadata, vocabularies and indexing terms. The technologies which support these intranets are *document management systems* – databases of un-structured or semi-structured documents rather than semantically well-structured material.



**Figure 1: WebML Page Decomposition**

Rather than a Web site that forms a well-defined interface at the boundary of the organisation (to a customer or trading partner), the information contained in a document management system is related to the internal working of the organisation. The use to which the information is put after publication varies with the role of each user within the organisation and the type and context of the information that has been assembled. As the intranet grows in size and complexity, it becomes impractical to build and use it in the present ad hoc and labour intensive fashion.

<sup>1</sup> WebML in a Nutshell, <http://xerox.elet.polimi.it/webml/readings/webmlnutshell.html>

Consider the following example: a manager writing a policy statement is required to draw together information held in a number of business documents: corporate vision statements, corporate strategy documents, departmental policy documents, management summaries, financial reports, public relations statements *etc.* While reading the content of those documents, the manager will also want to know their purpose (*e.g.* the intended audience) and authorship (*e.g.* the authors' role and position of influence) in order to be confident about any inferences made from the documents.

Hypermedia design methods help to identify the kinds of information needed to provide appropriate navigational access; document management systems help to collect metadata and provide classification and querying support to locate relevant information. However, managers do not often have sufficient time for unbounded browsing and searching to evaluate the appropriateness of supplementary documentation. What they could reasonably ask of a semantically-enriched support environment is to identify *relevant material* from *appropriate documents*, based on the context in which new material is being written.

The above scenario is not well supported by ad-hoc searching, but neither is it easily implemented with current web and hypermedia design models. Such models address the relations between information assets to provide site design and navigation features at the level of the document, unit or Web page, but fail to identify the connections between related information fragments for example an institution's three critical success factors and three section headings in the middle of its corporate strategy document..

Not only should such 'legacy' knowledge be accessible to the user of such a system, but new documents should be published in a form that facilitates reuse of the new knowledge embodied within them, providing explicit (hypertext) references to the sources of any reused knowledge.

The Semantic Web [5] augments the Web with explicit statements of document semantics, allowing the Web to be used as more than a human-browsable repository of information. The meaning of the published documents, knowledge about their authors and the reasons for their publication are all used to infer contextually appropriate associations, *i.e.* knowledge. This paper discusses the possibility of using Semantic Web techniques to improve hypermedia design models to support the kind of scenario developed above.

The paper continues by discussing current hypermedia and web design models (section 2) and their limitations (section 3). In this context, it introduces the Semantic Web (section 4) and some hypertext systems which incorporate some of its techniques and technologies (section 5). Finally we consider whether these techniques can be successfully applied to solve the shortcomings of Web development methods.

## **2 Engineering Web Design**

In this section we briefly present several well-known hypermedia or web-based design models and methods, focusing on their similarities and highlighting the gap that is to be addressed by our approach. Most of the currently available design methods (*i.e.* those described here) are model-driven and focus on the design stage of the hypermedia applications development life cycle or framework as proposed by Lowe & Hall [31]. All emphasize the need for an incremental and interactive development process [27], and generally consist of several orthogonal modelling dimensions. The typical modelling layers used in the process of designing

an application include the conceptual or structural level (information domain structure and design), the hypertext level (composition and navigation structure of the application), the presentational level (user interface or application look-and-feel design), personalization level (customization design), and the implementation level. The extent of coverage of these layers varies from one design model to another, and most of them formally focus on three layers, that is the conceptual, hypertext and presentational levels. At the conceptual level, the information domain is captured and modeled using three main design techniques:

- Entity-Relationship: information objects and data structure described by means of entities and relationships,
- Object-Oriented: information objects modeled as objects/classes.
- Ontology-based: information objects modeled as ontology classes.

The concept of views or perspectives is used at the hypertext level to enable the modeling of different applications, providing static views over the same conceptual model. A few methods such as WSDM, OOHDH and WebML provide support for more flexible personalization features (content, link, structure or context customizations) [14, 37, 9]. The compositional and navigational structure of an application is built on nodes (pages, navigation units, content units, slices or cards) and different types of links (perspective, structural, application link, etc) between them. The navigation units (nodes units) are mapped to conceptual units (entities or classes) to display the information or data at rendering/presentation time.

Although the design methods and models share at a higher level the similarities (common approach) described above in modelling web applications, they have several differences, including their main application domains, the level of coverage of the design process, the level of support provided at different stages.

HDM (Hypermedia Design Model) is an early E/R-based design model proposed by Garzotto *et al.* [18] to define the structure and interactions in large scale and read-only hypermedia systems. The model is suitable for domains with a high level of organisation, modularity and consistency. It focuses on hierarchically describing the information objects in terms of entities made of components containing units of information as well as the navigation structure, independently of their implementation. This navigation structure comprises perspective links between units, structural links between components, application links between entities, components or units, index and guided tours. The presentation design consists of slots (unit of information) and frames (grouping of slots).

RMM (Relationship Management Methodology) [26] is E/R-based, suitable for structured hypermedia applications and its design process consists of seven steps: entity-relationship design; slice design (grouping entities' attributes as node/presentation units called slice or M-slice); navigational design (access methods: link, menus, index, guided tour, indexed guided tour); protocol conversion design (converting design components into physical objects); user interface design (screen layouts); run-time behaviour design and construction and testing.

OOHDH (Object Oriented Hypermedia Design Model) is an OO-based design model that allows the specification of hypermedia applications as navigational views over the conceptual model [37]. Its design process consists of four main dimensions (See Table 1) and has recently been extended to formally cover requirements gathering [21] and personalisation modelling

[36]. Navigation units or nodes are mapped to conceptual classes, and the design and generation of OOHDM-based read-only web sites is supported by a CASE tool called OOHDM-Web [38]

The Enhanced Object-Relationship Model (EORM) [28] is an OO-based methodology whose major characteristic is the representation of relationships between objects (links) as separate objects. Links are therefore first class objects stored in reusable libraries, facilitating the mapping of relations into link class. The method is based on three frameworks: Class framework, Composition framework, GUI framework. EORM has early prototyping of user interfaces, a richer typology of links and a CASE tool (ONTOS Studio) to support the modelling process.

SOHDM (Scenario-based Object-oriented Hypermedia Design Methodology) is another OO-based approach focusing on process-oriented hypermedia systems to support organisational processes [30]. Scenarios are defined during the domain analysis and serve as the basis for object modelling and navigational design. Different types of OO views can be generated from the domain object model to compose a new application. It consists of six phases (See Table 1)

The Web Site Design Method (WDSM) [14] is a user-centred approach, as the application model is based on the user model, identifying user classes and their preferences and views. The design process for a read-only web site comprises three main stages (See Table 1). The conceptual design consists of both the object modelling (which can be E/R or OO based) and navigational design.

OntoWebber is an ontology-based approach to building read-only web sites, focusing on integrating heterogeneous data sources to build data-intensive web portals [43, 44]. A Domain Ontology serves as a reference ontology for data integration and content modelling, and as the starting point for the entire web site design. The site view modelling or site view graph consists of the navigation, content and presentation models; and further steps include personalisation and maintenance models (See Table 1). Nodes or navigation units are called cards which are mapped to ontology classes via the content model and the overall design is represented by an XML-based meta-schema using RDF and DAML+OIL [43]

WebML (Web Modelling Language) is a recent high level, model-driven, and E/R-based (compatible with UML class diagrams) design approach allowing a conceptual specification and automatic implementation of data-intensive web sites [9]. Four main orthogonal dimensions cover a web site specification (See Table 1) and navigation units (content units) in the hypertext model are mapped to relevant entities in the structural schema. A web site description is represented as a platform-independent XML meta-schema, and every concept derived from the specification has an associated graphical notation and XML representation. WebML extensions [6] allow interactive content management with entry units to update the site content, and the model has a CASE tool called WebRatio.

Overall, most current web design models provide users with model-driven approaches for the systematic design of high-level, read-only, well-organized, and easy to maintain web applications in different domains. Their coverage of the application life cycle focuses on the design stage with different levels of support provided at different orthogonal dimensions in the design process. Some are sustained by CASE tools for automatic implementation and generation of web sites. However, several limitations persist with regard to content modeling and management and the resulting linking capabilities.

	<b>Design process</b>	<b>Navigation structure:</b> Nodes and navigation units (nu)	<b>Interactivity</b>	<b>Modelling Technique</b>
<b>HDM</b>	1. Authoring-in-the-large 2. Authoring-in-the-small	Entities, components, units (nu)	Read-Only	E-R
<b>RMM</b>	1. E-R design 2. Slice design 3. Navigational design (slice diagrams) 4. Conversion protocol design 5. IU screen design 6. Runtime behaviour design 7. Construction and testing	outlines, slices (nu)	Read-Only	E-R
<b>WebML</b>	1. Structural model (data design) 2. Hypertext model Composition model Navigation model 4. Presentation model 5. Personalisation model	Pages, Content units (nu)	Read-Write	E-R / OO
<b>OOHDM</b>	1. Conceptual design 2. Navigational design navigational class schema navigational context schema 3. Abstract interface design 4. Implementation	Navigation classes (nu)	Read-Only	OO
<b>EORM</b>	1. Class framework 2. Composition framework 3. GUI framework	Navigation classes (nu)	Read-Only	OO
<b>SOHDM</b>	1. Domain analysis 2. OO modelling 3. View design 4. navigational design 5. Implementation design 6. Construction	classes	Read-Only	OO
<b>WSDM</b>	1. User modelling 2. Conceptual design Object modelling Navigational design 3. Implementation design (presentation) 4. Implementation	Navigation classes (nu)	Read-Only	E-R / OO
<b>Onto-Webber</b>	1. Domain ontology 2. Site View modelling, navigation model content model presentation model 3. personalisation model 4. maintenance model	Pages, Cards (nu)	Read-Only	Ontology

**Table 1 Web Design Model features**

### 3 Shortcomings of Web Design

Apart from WebML, which provides extensions enabling content management features, most of the methodologies have a static (data) view over the web site content and allow the modelling of read-only web sites. The resulting applications are largely built to present/publish the data, but not to manage the content. Consequently, many of the methodologies and models described in the previous section take a simple layered approach, separating the design issues so as to allow independence for:

- Mapping the domain, in terms of its structure, content, work flow, *etc.*
- Analysing the associations and relation in that domain
- Presenting the information to appropriate users

A common weakness with these approaches is the lack of ‘cement’ connecting the layers and the missing means of mapping between the different layers [32], *i.e.* in practice the result of one activity does not feed into the next.

At the hypertext level, navigation units (cards, slices, content units, navigation classes, slots, *etc.*) are generally mapped to information units (entities, units, classes, *etc.*) in order to present the content in web pages, but the level of granularity of these units does not allow authored links to reach the real text inside units. Automated links are restricted to navigation units or groups of navigation of units, and any link to/from the inside of the units have to be manually added.

By contrast, Open Hypermedia Systems (OHS) promote links to first class objects that are stored and managed separately from multimedia data. The advantage of these systems is that they allow links to be added to the multimedia content in a way that is appropriate to the user and to the document contexts. Early OHS like Microcosm [12] and Intermedia [41] have influenced the XLink Web standard [14] which allow links to be added to Web documents *independently of their storage.*

Navigation can be viewed as a combination of hypermedia linking, information retrieval and document management [40]. While navigation design is covered by all the major models, none directly address the issue of hyperlinks in the content; some (those based on HDM) even stipulate that links *should not* be placed in the content. This position arises concerning links embedded at design time, clearly these embedded links can become invalid when the context in which the webpage is used changes [19]. Many other models restrict user navigation between pages/containers by the use of buttons or links contained in toolbars or sidebars. These can be more flexible, and can be changed at run time. However, usability studies show that “*when they arrive on a page, users ignore navigation bars and other global design elements: instead they look only at the content area of the page*” [34]. In other words, links should not be completely ruled out of the content that the user is viewing.

Allowing links to become first class objects that are only embedded in documents at the time of viewing allows only those options appropriate to the users to be displayed. In practice, this may be simply achieved by swapping different linkbases in and out of use, thereby creating different paths through the same documents. In addition, the choice of linkbases can be deferred to an agent thereby making it adaptive to the user [2] and dependent on the context in which the user is browsing the information space [15].

These open hypermedia techniques allow the addition of links to material to good effect. These links can either be extracted from databases or computed dynamically; it is this freedom that should be subject to the discipline of the design process, to allow the rationale and the mechanisms for choice to be clearly expressed and tested for effectiveness. To enable the design process to inform a more complete linking activity, a structured approach is needed that would enable the microstructure (low level) of information objects (or documents' content) to be addressed and modelled.

## 4 Semantic Web

Hypertext is just one example of the use of a family of techniques that are intended to transcend the limitations of static, sequential presentations of text [35]. Hypertext uses computer effects (such as linking, indexing and interaction) to improve familiar textual communication for human beings; it is the practice of human communication augmented by computer-manipulated media, databases and links. By contrast, the Semantic Web is an application of the World Wide Web aimed at computational agents, so that *programs*, and not just humans, can interpret the meaning of the information stored in the WWW hypertext [5]. The basis of this interpretation is an ontology, a structure which forms the backbone of the knowledge interpretation for an application.

In Knowledge Management (KM) an ontology is “*a specification of a conceptualization*” [20]. Gruber explains that a common ontology defines the vocabulary with which queries and assertions are exchanged among agents (people or software). The ontology sets out all the entities (objects or concepts) that we are interested in and the relationships that bind these entities together. This is intended to be a *pragmatic* definition, *i.e.* it defines the vocabulary that is actually *in use*, and the concepts that are *useful* in problem-solving. It does not give the deep underlying philosophical vision of the ultimate entities in the field. Hence, in KM, an ontology is a tool whose quality is entirely dependent on its usefulness.

The World-Wide-Web Consortium (W3C) describes an ontology as defining the terms used to describe and represent an area of knowledge (usually called a *domain*), for use by people, databases, and applications to share information [24]. An ontology merely specifies *one* way of understanding the world, and different ontologies will be useful for different things. Hence there could be two or more ontologies that describe the same phenomena but are very different to each other – yet both could be, for their own purposes, correct. An immediate Web application of ontologies is in searching – otherwise a purely syntactic activity matching patterns of letters. Ontology-augmented searches can determine that a page about “yetis” is relevant to a search for “monsters”, because a yeti is a specific subtype of monster, even where the sequence of letters *m-o-n-s-t-e-r* does not appear.

To link the computer-accessible semantics contained in an ontology with the human-oriented semantics contained in the ‘content unit’ of a web page, a process of annotation is required. Formal statements in a standard Web language (currently RDF [29]) refer directly to concepts in the ontologies and to some content on the Web, enabling a program to determine that a particular string (*a-b-o-m-i-n-a-b-l-e -s-n-o-w-m-a-n*) in a particular document refers to a Yeti.

An ontology is a formal model that allow reasoning about concepts and objects that appear in the real world and (crucially) about the complex relationships between them [17]. It seems reasonable to imagine that some kinds of complex structures may be required for discussing and

exploring inter-relationships between objects when we make hypertext statements about those interrelationships. Normal hypertext design practice (above) is to analyse the texts themselves in order to devise a suitable hypertext infrastructure. By contrast, ontologically-motivated hypertexts derive the structuring of their components from the relationships between objects in the real world.

## 5 Qualifying Systems

Many Semantic Web developments have focused on the issues related to knowledge modelling and knowledge publishing (ontologies, knowledge-bases, inferencing) and as a result, tend to sideline the role of complex, user-centred documents. However, ontologies have influenced a number of hypertext developments in recent years, some of which bridge the gap between the (human-readable) Web and the (machine-processable) Semantic Web.

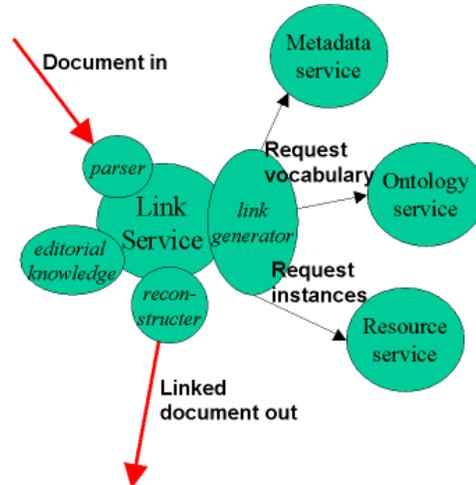
### 5.1 COHSE

The COHSE project (Conceptual OHS Environment, [8]) produced an experimental ontological hypermedia system by combining an existing open hypermedia link service with an ontological reasoning service to enable documents to be linked via the concepts referred to in their contents. COHSE was particularly concerned with the authoring process, tackling the problem that the manual construction of hypertexts for non-trivial Web applications (where documents need to be linked in many dimensions based on their content) is often inconsistent and error-prone [16]. Attempts to improve the linking through simple lexical matching had serious limitations due to the uncontrolled method of adding links: many keywords turn up in many contexts and there is no simple lexical basis for discriminating important terms and significant links. The aim of the COHSE project therefore was to combine the OHS architecture with an ontological model to provide linking on the *concepts* that appear in Web pages, as opposed to linking on *simple uninterrupted text fragments*.

Ontologies are used to describe the interrelationships between concepts embedded in the documents to provide a new ‘catalogue of internal knowledge’ [3]. An ontological hypertext environment needs to have some mechanism for interpreting the ontology and exposing these concepts and relationships in the real world as links (or other artefacts) in the hypertext. COHSE used a standard Web browser controlled by an adapted *link service* which in turn used three independent services to manipulate the exposed Document Object Model (DOM) of the Web page, resulting in the effect of ontologically-controlled hypertext.

In Figure 2, the **ontology service** manages ontologies (sets of concepts related according to some schema) and answers specific queries about them. The ontologies are internally represented using DAML+OIL [4] and queries are satisfied using the FACT reasoner [25]. The purpose of the service is to answer fundamental questions about the concepts in an ontology, for example: *what is the parent of this concept*, or *how is this concept represented in a natural language*, or *what concept does this string describe* or *are these two concepts similar or the same?* Unlike other Semantic Web systems, COHSE’s ontology server does not use *specific relationships* to answer ontology-specific (and hence domain-specific) questions (*e.g. who wrote this paper*, or *what kind of person manages an academic project* or *who can be a chartered engineer?*). The **metadata service** annotates regions of a document with a concept, rather than the familiar case of annotating a document with a simple piece of text. An XPointer is used to identify each region in the document; a fragment of RDF that corresponds to a DAML+OIL

statement identifies the concept. The **resource service** is a simple librarian that is used to lookup Web pages which are examples of a particular concept (*i.e.* that can be used to *illustrate* a concept).



**Figure 2 COHSE Architecture for Ontological Hypertext**

When a web page is loaded, the ontology service provides a complete listing of all the language terms that are used to represent the concepts in the relevant ontology. Each language term is searched for in the document, and, if found, its associated concept is looked up. The metadata service is also used to determine whether any regions in the document have been manually annotated (allowing concepts to be recognised even if the document does not use the ‘approved’ language terms). Having identified the significant concepts in the document, the resource service provides a list of documents that are about instances of this concept.

At this point, a number of potential link anchors and destinations have been identified for the page and decisions can be taken about whether the document contains too many or too few links. In those circumstances, alternative links may be chosen from the broader or narrower concepts in the ontology in order to expand or cull the set of link anchors. The decisions about link culling and presentation are controlled by behaviour modules which define the navigation and interaction semantics of the resulting ontological hypertext.

## 5.2 CREAM

CREAM (CREAting Metadata [22]), is an ontology-based framework for metadata and document creation. It is based on the Ont-O-Mat tool, a component-based annotation and authoring system built around a document editor and ontology browser. CREAM supports Semantic Web knowledge creation by annotation both during and after authoring. Annotation can be achieved by filling out knowledge templates under the control of the ontology browser (either by typing values or by dragging and dropping literal strings from the document editor). More interestingly, documents (or content fragments) can be built by a process of reverse-annotation; entries from the ontology or knowledge-base are used to create text (*e.g.* Leslie Carr is a researcher who works at Southampton University) which may retain links back to the knowledge base.

The major concern of the CREAM framework is to create knowledge that can be used in Semantic Web applications (*e.g.* querying, inference and structured portal generation). It therefore uses ontologies mainly for annotation purposes and achieves limited support for content authoring as a by-product of the annotation activities. However, this support is significant because in embryonic form it imposes a principled knowledge framework on otherwise free-form textual material as it is being created.

### 5.3 ONTOPORTAL

Ontoportals are a generic application framework for building ontology-based web portals [44]. It shows how a semantic meta-layer of ontology concepts and relationships can be instantiated or projected over existing weakly interlinked web resources to generate a web portal meaningfully describing and linking the resources and their relations. The framework provides facilities such as: exploration (browsing an ontoportal); knowledge capturing (content creation or update); thread discussion (on themes around the resources being browsed); and searching (keywords search over the stored metadata); corresponding to four main modes of interaction with users.

Producing a new ontoportal (an Ontoportals-based web portal) involves creating and populating the domain ontology that can later be reused to generate other ontoportals in similar domains, and setting XML-based presentation rules for different display modes. Examples of applications built from this framework include Metaportal (web portal for the metadata research community) that can be found on the project web site at <http://www.ontoportals.org.uk/>; and two educational domain portals TPortal and XPortal<sup>2</sup> [45] (for teaching and learning purposes).

Ontoportals are an ontological hypertext system, therefore ontologies are used to improve the navigational facilities of resulting web portal applications. New types of links, that is conceptual links, are inferred from the underlying domain ontology structure to enrich the linking between resources and to enable complex queries to be answered by simply following ontological links (query by linking).

In the systems described in section 5, combinations of ontologies, knowledge-bases, document services and hypermedia services have been produced to create some sort of conceptual hypermedia system that supports the creation and linking of WWW documents at retrieval time (as readers browse the documents) or at authoring time (as authors create the documents).

None of these systems fully addresses the concerns of hypermedia authoring in the context of a web site; COHSE promotes the creation of links and CREAM promotes the creation of (metadata or) text.

## 6 Requirements of Improved Web Methods

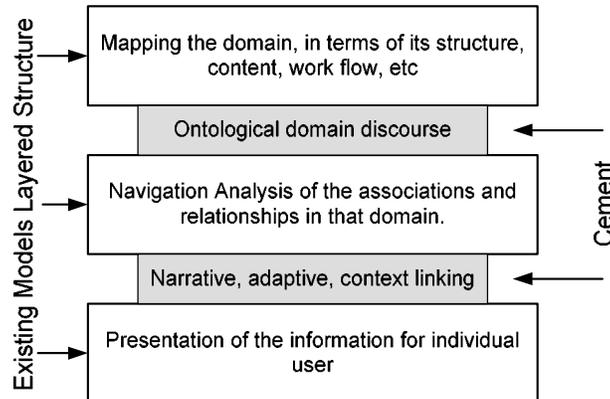
Existing web design models suffer from a lack of ‘cement’ (as described in section 3), in other words, they have no clearly defined way of moving from one stage to the next. While each of the models and methodologies described in section 2 have their own advantages and disadvantages [11], they all emphasise the imposition of organising principles on a collection of documents (by clustering, partitioning and decomposition). To inform the design of other types of information environments, we require a model that will also help expose the relationships

---

<sup>2</sup> TPortal and XPortal are used in the Department of Electronics and Computer Science, University of Southampton. Internal links are <http://pip.ecs.soton.ac.uk/tportal/cgi-bin/explore.cgi> and <http://pip.ecs.soton.ac.uk/xportal/cgi-bin/explore.cgi>.

within the content of the individual documents, *e.g.* that bullet point 1 of a Company policy document is expanded in paragraph two of the Departmental policy document.

We suggest that these parallel requirements could be satisfied by an interleaved model (Figure 2). Ontologies have a dual role in expressing both large-scale concepts/relationships and also discrete entities/specific instances, consequently they may be used as the ‘cement’ that maps between the domain analysis and hypertext navigation layers. Different ontologies would be used for each of the user groups or the tasks to be undertaken by the web site, so providing alternative perspectives and navigational paths through the information domain.



**Figure 3: Interleaved Models**

The existing models (white areas) examine the macro-structure of the collection (web site, intranet, repository *etc.*) which is used to design navigation and presentation strategies for the documents, and provide a ‘catalogue of assets’. The layers shown are independent of the exact design method used, and may work with either an object-oriented or entity-relational approach. The two greyed areas of missing ‘cement’ are needed to couple the organising principle of the web site with the semantics within the texts and with the objectives of the presentations.

Previous work on the use of knowledge tools to support hypertext by Nanard *et al.* [33] set out the requirements for a hypertext design environment, recommending that tools are required for generalising and instantiating knowledge models, to enable designers to alternate between bottom-up and top-down approaches, thus promoting both structuring and updating activities.

We suggest that the knowledge modelling work of the Semantic Web could be exploited by applying ontologies not only as an organising principle for documents, but also to describe the interrelationships between concepts embedded in the document content. The model could also expose these concepts (and the structure of their inherent relationships) whilst documents are being written or read.

We suggest that there are many benefits to be had from extending the scope of the design activity from simply dealing with the web site ‘templates’ into dealing with the semantics of the content. The availability of this design framework at:

- authoring time will support authors with appropriate knowledge for constructing texts (*i.e.* narrative and rhetorical material) and the relevant links between them
- reading time will support readers with adaptive and context-sensitive information delivery and linking techniques.

## 7 Concluding Remarks

The objective of hypermedia design models is to produce a well-organised web site. This paper has elaborated the problem of modelling the semantics within text using current design models and methodologies. While the hypermedia models are used to provide the macro-structure of a web site (provide site design and navigation features at the level of the document, unit or Web page), they fail to identify the interconnected semantic fragments contained within the text.

Many of the methodologies and models used in web design take a simple layered approach, separating the design issues so as to allow independent implementations. These fail to provide a method of mapping from one stage to the next, that is, they lack ‘cement’. Treating links as first class objects has provided a means of joining the navigational layers and presentation layers, through adaptive, contextual, and narrative linking.

Similarly, Semantic Web knowledge models (ontologies) can be used to ‘cement’ the domain and the navigational analyses, not only by providing the relation between different content units but also by making explicit the relationships between the semantics within the text. As a result, designers can extend their influence into the texts to influence the production of semantic content.

This approach to extending the existing design models was born out of the experience of using the models described in section 2 in large industrial hypermedia projects, in running a Web Design undergraduate course, and in the authors’ research activities into the Semantic Web [1]. Elements of a large website for a current project (the Virtual Orthopaedic European University (VOEU), <http://voeu.ecs.soton.ac.uk/>) have been built using this model. Further work is underway to instantiate the abstract model (Figure 3) and requirements listed here into a generalised Web design model supported by appropriate tools, with a full case study.

## 8 Acknowledgements

This work has been funded in part by the Writing in the Context of Knowledge project funded in the UK by the Engineering and Physical Sciences Research Council under grant number EPSRC GR/R91021/01.

## 9 References

- [1] AKTors, Advance Knowledge Technologies, <http://www.aktors.org>,
- [2] Bailey, C. and Hall, W. (2000) An Agent-Based Approach to Adaptive Hypermedia Using a Link Service. In Brusilovsky, P. and Stock, O. and Strapparava, C., Eds. *Proceedings of the First International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH2000)*, pp. 260-263, Aug, 2000
- [3] Bechhofer S., Carr L., Goble C., Hall W. (2001) Conceptual Open Hypermedia = The Semantic Web? Position Paper. In *Proceedings of The Second International Workshop on the Semantic Web - SemWeb'2001*
- [4] Bechhofer, S. and Goble, C. (2001) Towards Annotation using DAML+OIL. *K-Cap Workshop on Knowledge Markup and Semantic Annotation, Victoria, B.C., Canada*
- [5] Berners-Lee, T., Hendler, J., Lassila, O. (2001) The Semantic Web, *Scientific American*, May 2001 34-43

- [6] Bongio, S. Ceri, P. Fraternali, A. Maurino (2000). Modelling data entry and operations in WebML. *WebDB: 87-92*
- [7] Buckingham-Shum S, Motta E, Domingue J. (2000) ScholOnto?: An Ontology-Based Digital Library Server for Research Documents and Discourse *International Journal on Digital Libraries 3 (3), 237-248*
- [8] Carr, L., Bechhofer, S., Goble, G., Hall, W., De Roure, D. (2001) Conceptual Linking: Ontology-based Open Hypermedia *Proceedings of The 10th International WWW Conference*. May 2001
- [9] Ceri S, Fraternali P, Bongio (2000) A: Web Modelling Language (WebML): a modelling language for designing Web sites. In *Proceedings WWW9 Conference, Amsterdam, May*
- [10] Ceri, S (2002) Conceptual modeling of data-intensive Web applications. *IEEE Internet Computing 6(4) July*.
- [11] Christodoulou SP, Styliaras GD, Papatheodorou. (1998) Evaluation of Hypermedia Application Development and management Systems. In *Proc. Hypertext 98 The ninth ACM Conference on Hypertext and Hypermedia*, Pittsburgh June 20-24. pp1-10.
- [12] Davis HC, Hall W, Heath I, Hill G, Wilkins R. Towards an Integrated Environment with Open Hypermedia Systems. in *Proceedings of the ACM Conference on Hypertext, EHCT'92*, Milian, Italy, December 1992 pp 181-190. ACM Press 1992
- [13] DeRose, S, Maler, E. and D. Orchard. (2001) XML linking language (XLink) *version 1.0. W3C Recommendation, 27 June 2001*.
- [14] De Troyer O. and Leune C. (1997) WSDM: A user-centered design method for Web sites. In *Proceedings of the 7th International World Wide Web Conference, 1997*.
- [15] El-Beltagy, S., Hall, W., DeRoure, D. and Carr, L. (2001) Linking in Context. In *Proceedings of HT01, the Twelfth ACM Conference on Hypertext*, pp 151-160 August 2001
- [16] Ellis, D., Furner, J., and Willett, P. (1996) On the creation of hypertext links in full-text documents - measurement of retrieval effectiveness. In *Journal of the American Society for Information Science, Vol. 47, No. 4, 287-300*.
- [17] Fensel D., van Harmelen, F., Horrocks, I., McGuinness, D.L., Patel-Schneider, P.F. (2001) OIL: An Ontology Infrastructure for the Semantic Web, *IEEE Intelligent Systems*, Vol. 16, No. 2, 38-45.
- [18] Garzotto F., Paolini P. (1993) HDM- A Model-Based Approach to Hypertext Application Design. In *ACM Transactions on Information Systems, Vol. 11, No1 January p1-26*.
- [19] Garzotto F., Luca Mainetti L., Paolini P (1996). Information Reuse in Hypermedia Applications. In Proc. The Seventh ACM Conference on Hypertext, HYPERTEXT '96 Washington DC March 16-20 1996. pp 93-104
- [20] Gruber T. R. (1993). A translation approach to portable ontologies. *Knowledge Acquisition, 5(2):199-220*. On-line at [http://ksl-web.stanford.edu/KSL\\_Abstracts/KSL-92-71.html](http://ksl-web.stanford.edu/KSL_Abstracts/KSL-92-71.html)

- [21] Guell, N., Schwabe D., Vilain, P. (200): "Modeling Interactions and Navigation in Web Applications", *Lecture Notes in Computer Science 1921, Proceedings of the World Wide Web and Conceptual Modeling '00 Workshop, ER'99 Conference, Springer, Salt Lake City, 2000. ISBN 3-540- 41073-2.*
- [22] Handschuh, S and Staab S (2002) Authoring and Annotation of Web Pages in CREAM. In *Proceedings of WWW2002*. May 7-11 2002. Hawaii.
- [23] Heath I., Wills G., Crowder R., Hall W., Ballantyne J. (2002) A New Authoring Methodology for Large-Scale Hypermedia Applications. In *Multimedia Tools and Applications 12(2/3): 129-144, November 2000.*
- [24] Heflin J, Volz R, Dale J (2002) Requirements for a Web Ontology Language W3C Working Draft 08 July 2002 On Line at: <http://www.w3.org/TR/webont-req>
- [25] Horrocks, I. (2000) Benchmark analysis with fact. In *Proceedings of TABLEAUX 2000*, 62-66.
- [26] Isakowitz T, Stohr EA., Balasubramanian P. (1995) RMM: A design Methodology for Structured Hypermedia Design. In *Communications of the ACM Vol.38 No8 August pp 34-44.*
- [27] Koch, N. (1999) A Comparative Study of Methods for Hypermedia Development, *Technical report 9905, Luddwig-Maximilians-Universitat Munchen, November 1999.*
- [28] Lange DB. (1996) An Object-Oriented design Method for Hypermedia Information Systems. In *Journal of Organizational Computing & Electronic Commerce, Vol. 6 (3).*
- [29] Lassila, O., Swick, R. eds. (1999) Resource Description Framework (RDF) Model and Syntax Specification, *W3C Recommendation 22 February 1999*
- [30] Lee H., Lee C. and Yoo C. (1998) A scenario-based object-oriented methodology for developing hypermedia information systems. In *Proceedings of 31st Annual Conference on Systems Science, Eds. Sprague R., 1998.*
- [31] Lowe D, Hall W, *Hypermedia Engineering, the Web and Beyond, Wiley 1999.*
- [32] Lowe D, Webby R (1999) "A Reference Model, and Modelling Guidelines for Hypermedia Development Processes", *The New Review of Hypermedia and Multimedia, Vol. 5, pp 133-150, 1999 Taylor Graham Publishers.*
- [33] Nanard J, Nanard M. (1995) Hypermedia Design Environments and the Hypertext Design Process. In *Communications of the ACM, .38,(8), pp49-56*
- [34] Neilsen J (2000) Is Navigation Useful. *Alertbox, January 9, 2000.*  
<http://www.useit.com/alertbox/20000109.html>
- [35] Nelson, T. (1987). *Literary Machines. 87.1 edn. Computer Books.*
- [36] Rossi G, Schwabe D, Guimaraes R. (2001) Designing Personalized Web Applications. In *proceedings WWW10, Hong-Kong, May*
- [37] Schwabe D, Rossi G, Barbosa SDJ. (1996) Systematic Hypermedia Design with OOHDM, *Proc. Of the Hypertext' 96, Washington, March.*

- [38] Schwabe D, . de Almeida Pontes R, Moura I. (1999) OOHDWeb : An Environment for implementation of Hypermedia Applications in the WWW. SigWEB Newsletter, Vol. 8, #2, June de 1999
- [39] Shadbolt, N.R., and Milton, N (1999) From Knowledge Engineering to Knowledge Management. In *British Journal of Management*, 10, 309-322.
- [40] Wills G.B Design and evaluation of Industrial Hypermedia. *PhD thesis University of Southampton January 2000*
- [41] Yankelovich N, Haan B, Meyrowitz N, Drucker S (1988) Intermedia: the concept and the construction of a seamless information environment, *IEEE Computer* 21 (1), 81-96
- [42] Yuhui J, Decker S, Wiederhold G (2001). OntoWebber: Model-Driven Ontology-Based Web Site Management. In *1st International Semantic Web Working Symposium (SWWS'01)*, Stanford University, Stanford, CA, July 29-Aug 1, 2001.
- [43] Yuhui J, Sichun X, Stefan Decker. (2002) OntoWebber: A Novel Approach for Managing Data on the Web. In *the 8th International Conference on Extending Database Technology (EDBT 2002)*. Prague, Czech Republic, March 24-28, 2002.
  
- [44] L. Carr., S. Kampa, and T. Miles-Board. MetaPortal Final report: Building Ontological Hypermedia with the Ontoport Framework. Technical Report, IAM, ECS, University of Southampton, March 2001.
  
- [45] Woukeu Arouna., Wills G, Conole G., Carr L., Kampa S., Hall W. (2003). Ontological Hypermedia in Education: A framework for building web-based educational portals. In *Proceedings of ED-MEDIA 2003 – 12<sup>th</sup> World Conference on Educational Multimedia, Hypermedia & Telecommunications*. Honolulu, Hawaii, USA, June 23-28, 2003.