

ScholOnto: An Ontology-Based Digital Library Server for Research Documents and Discourse

SIMON BUCKINGHAM SHUM, ENRICO MOTTA, JOHN DOMINGUE

Knowledge Media Institute, The Open University, Milton Keynes, MK7 6AA, U.K.

Email: {S.Buckingham.Shum, E.Motta, J.B.Domingue}@open.ac.uk

WWW: <http://kmi.open.ac.uk/projects/scholonto/>

Phone: +44 1908 655723

Fax: +44 1908 653169

Key words: scholarly discourse – scientific publishing – ontologies – knowledge-based systems – argumentation – visualization – eprint servers – internet digital libraries

Abstract. The internet is rapidly becoming the first place for researchers to publish documents, but at present they receive little support in searching, tracking, analyzing or debating concepts in a literature from scholarly perspectives. This paper describes the design rationale and implementation of *ScholOnto*, an ontology-based digital library server to support scholarly interpretation and discourse. It enables researchers to describe and debate via a semantic network the contributions a document makes, and its relationship to the literature. The paper discusses the computational services that an ontology-based server supports, alternative user interfaces to support interaction with a large semantic network, usability issues associated with knowledge formalization, new work practices that could emerge, and related work.

1 Introduction

It is becoming standard practice for researchers to publish their documents on the internet (or intranets), via personal, institutional and discipline-specific servers. As older documents are gradually digitized, and scientific publishers reconceptualize their roles, we move towards the highly desirable situation (for researchers) of instant access to scholarly documents. What remains largely unaddressed is the key challenge of managing this ocean of information. *Publishing* documents is the only first step. Obtaining the *best* documents is the next, harder step, and *interpreting* them the next, and most complex to support computationally within a digital library. In this article we present work on a digital library server which overlays on conventional documents and metadata a semantic web of scholarly claims, discourse and perspectives. We seek to augment individual *analysis* of literatures for significant conceptual structures, provide more effective discovery of *relevant documents*, and enable *structured discourse* between researchers.

The paper begins by considering tasks that face scholars in analysing a document or literature (§2). We argue that many metadata approaches to document description, and current technological infrastructures provide poor support for these interpretive tasks. After outlining current technologies for scholarly publishing and discourse on the internet (§3), we present a representation scheme to assist scholars in articulating, interpreting and contesting perspectives (§4). Specifically, these perspectives take the form of *networks of contestable claims* about ideas. We propose that researchers enrich their texts with nodes and links which they add to an evolving semantic network reflecting the relevant literature. We detail how such a model is being implemented as a digital library server (§5), and following a worked example (§6), discuss the services it will provide to scholarly communities, and issues this work raises for scientific publishing practices (§7).

2 The scholarly work of interpretation

Contextualising ideas in relation to the literature is a fundamental task for authors and readers—are they new, significant, and trustworthy? Scholars start by bringing to bear their own knowledge of the field, which often leads to *commentary and discourse* of various kinds. Such discourse can range from private annotation of a document, to personal email and discussion groups, to formal peer review of conference/journal submissions and published reviews of literatures and books.

Our point of departure to assume that the author wishes to persuade the reader to accept his/her perspective, which constitutes a set of claims about the world. Usually, the author has

some new ideas that s/he is contributing, and asserts particular relationships between these and existing ideas already published in order to demonstrate both the reliability of the conceptual foundation on which s/he is building, and the innovation and significance of the new ideas. The reader's task is to understand which ideas are being claimed as new, and assess their significance and reliability.

Faced with the different task of literature search and analysis, the scholar has ideas or work that s/he is trying to locate in the literature, for instance, how is such work described?; or perhaps these ideas already exist as islands that have yet to be bridged? The interpretive task includes formulating the ideas of interest to uncover relevant documents, reading the documents and then interpreting them to characterize any patterns that appear to emerge. A common scenario is that of a newcomer to a scholarly community (e.g. a student; librarian; lecturer or researcher from another discipline) who wants to know, for instance, what the seminal papers are, or if there are distinctive perspectives on problems or classes of technique that define that community. More specifically, researchers are typically interested in the following kinds of phenomena (exemplified with queries):

- The *intellectual lineage* of ideas: e.g. where has this come from, and has it already been done? (“Are there any arguments against the framework on which this paper builds?”)
- The *impact* of ideas: e.g. what reaction was there to this, and has anyone built on it? (“Has anyone generalised method M to another domain?” “Has anyone extended Language L?”)
- *Perspectives*: are there distinctive schools of thought on this issue? (“Has anyone proposed a similar solution to Problem P but from a different theoretical perspective?”)
- *Inconsistencies*: e.g. is an approach consistent with its espoused theoretical foundations?; is there contradictory evidence to a claim? (“Are there groups building on Theory T, but who contradict each other?”)
- *Convergences*: are different streams of research mutually reinforcing in interesting ways? (“Who else uses Data X in their arguments?”)

Currently, researchers have very little, if any, way to articulate such questions in a library, analogue or digital (citation analysis offers limited support for assessing impact). Current metadata initiatives are focused on the encoding of *primary content attributes* to improve retrieval and interoperability. Inconsistencies and interpretations in encoding at this level are considered undesirable in order to assist machine processing. In contrast, these are precisely the features that a system needs to support the *interpreted, knowledge level*, as researchers contest the significance of work, and the concepts which it underpins. The present work seeks

to address the fundamental requirement for an ontology capable of supporting scholarly research communities in interpreting and discussing evolving ideas: overlaying *interpretations* of content, and supporting the emergence of (possibly conflicting) *perspectives*.

3. Current scholarly publishing and discourse technologies

3.1 Automated e-print servers

E-print servers automate the submission, archiving and dissemination of author-provided documents, both unreviewed preprints, and reviewed reprints (e.g. of conference and journal of papers). A current initiative that consolidates the trend towards self-archiving of research publications is the Open Archives Initiative [46] comprising individuals and organizations who have initiated such servers in their respective communities. E-print servers typically support standard document formats, use hierarchical taxonomies and keyword schemes for indexing and search, and send e-mail alerts of new documents. The system described below, ScholOnto, adds a semantic layer to the conventional metadata requested by e-print servers (see Related Work, §7).

3.2 Evolving the Web beyond simple linking

As the first global hypertext system to achieve large scale uptake, the Web provides a rudimentary infrastructure for publishing interlinked documents and discourses. However, the Web provides little support for *structuring, searching* or *analysing* scholarly concepts, documents or discourses. Early, pre-Web hypertext systems have already demonstrated (on a small scale) the power of features such as semantically typed nodes and links, bi-directional links, composite nodes which represent more complex structures, and structural searching. It is widely recognised that the Web now needs to exploit such features [4]. ScholOnto implements a number of these features, most importantly, semantic links.

3.3 Document discussion and annotation systems

The extension of online document environments to support annotation and discussion begins to support the interpretive process. We have reported elsewhere [10], [55] on a tool [15] that extends a document-centric use of the Web with threaded, textual discourse facilities for peer review in an electronic journal [32]. Many other initiatives into document discussion and annotation are developing standards [62] and tools [59]. Text is a rich medium in which to publish and discuss ideas with subtle nuances, with the corresponding cost that it takes a long

time to read, and is hard to analyse computationally. The complementary approach described in this paper, with different goals, focuses on the conceptual models implicit in textual documents and discourse; this is a move from textual documents and discourse to network representations, whilst maintaining the focus on scholarly discourse. The goal is to provide a *summary* representation of ideas and their interconnections, in order to assist literature-wide analysis. We propose that this has advantages over textual media for tracing the intellectual lineage of a document’s ideas, and for assessing the subsequent impact of those ideas, that is, how they have been challenged, supported and appropriated by others.

3.4 Automatic information analysis and retrieval

Information retrieval research using statistical and text analysis techniques to analyse document corpora is finding increasing application on the Web. Techniques with particular promise for analysing scholarly literatures include the clustering and visualization of documents based on semantic similarity (e.g. using latent semantic indexing [14]), autonomous citation indexing [36] and the analysis of the context surrounding in-text citations [5], and identification of basic types of inter-document relationship (e.g. “cites”, or “summarizes” [1]). Automatic techniques clearly have the advantage that they can be run on large text corpora with little human effort. However, from the perspective of scholarly interpretation, it is extremely hard to automatically identify more complex kinds of *scholarly relationships* between documents that are implied by the interpretive tasks listed in §2. For instance, automatic techniques can usefully cluster documents on statistical indices, but are unable to distinguish whether they are in vehement disagreement, or form a coherent perspective with respect to some common problem, method, or theory. We are therefore using *human-encoded* document descriptions to express scholarly claims, argumentation and conceptual structures. This has the added advantage that the key stakeholders promoting that work—namely the authors—take ownership of, and have control over, the way in which their work is represented, rather than losing this control to automated classification. An obvious tradeoff is in the effort required to codify new kinds of information, and in the discussion we consider how our approach tackles the ‘capture bottleneck.’

4 An ontology for representing scholarly claims

4.1 Ontologies

An ontology in philosophy refers to a conception of what can exist or ‘be’ in the world [11]. The artificial intelligence community has appropriated the term to mean the construction of *knowledge models* (e.g. [24], [41], [43]) which specify concepts or objects, their attributes and

inter-relationships. A knowledge model is a specification of a domain, or problem solving behavior, which abstracts from implementation-specific considerations and focuses instead on the concepts, relations and reasoning steps characterizing the phenomenon under investigation. Our application of knowledge modelling in this project is to specify and implement a semantic network which reflects important aspects of the web of ideas and perspectives that often remain (from a computer's perspective) *implicit* in textual documents, and (from other researchers' perspectives) *implicit* in the minds of a scholarly community's members. As a semantic network of *individual* perspectives is constructed, it is possible that this could reflect meaningful aspects of the *collective perspective* of its stakeholders.

4.2 A discourse-oriented ontology

The role of an ontology is to reflect a community's consensus on a useful way to conceptualise a particular domain. It might appear paradoxical, therefore, to propose the use of ontologies to support scholarly communities in managing their knowledge, since conflicting worldviews, evidence and frames of reference lie at the heart of research and debate. Any attempt to impose a 'master ontology' on a research field is unlikely to succeed.

The key issue is in what is represented. We observe that whilst the scope and subject matter of a research field is in flux, the *way* in which it evolves is relatively stable. It is hard, for instance, to envisage when researchers will no longer need to make claims about or contest the nature of a document's *contributions* (e.g. "this is a new theory, model, notation, piece of evidence"), or its *relationships* to other documents (e.g. "it applies, modifies, predicts, refutes..."). Our approach provides researchers with a way to make *claims* about the *significance* of ideas and concepts—a focus on *discourse* and *argumentation*. Moreover, separating the representation of concepts from claims about them is critical to the co-existence of multiple perspectives.

The kind of representational scheme we are focusing on, intended to be understandable and customizable by a wide range of disciplines, is summarized in Figure 1. This ontology is designed to support scholars in making *claims* by asserting relationships between concepts, or between concepts and other claims. Other scholars may in turn contest these claims using a variety of argumentation relations.

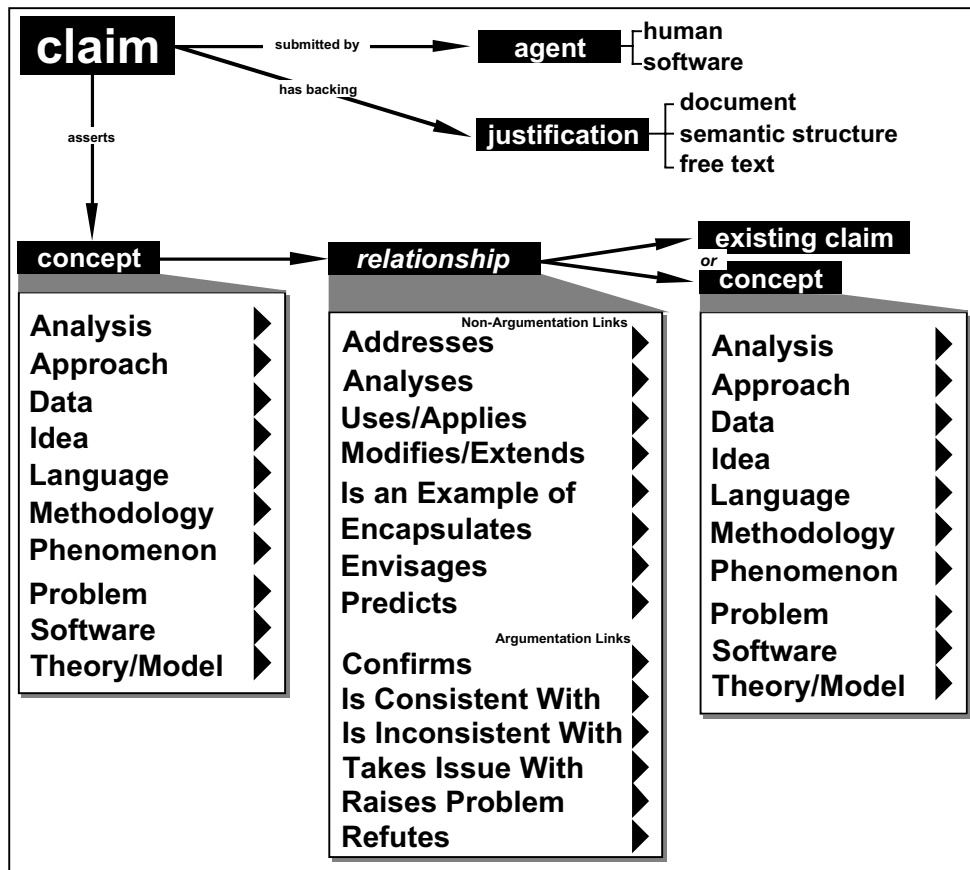


Fig. 1. The structure of a scholarly *Claim* in the ontology. All claims are owned by an agent, and have some form of justification. Claims assert new relationships with other claims, or between concepts.

The goal is to design a relatively small set of uncontroversial conceptual and relational types which are simple enough to understand without being simplistic, yet expressive enough that most researchers can express the key claims made in most documents. We hypothesise that most disciplinary communities can intuitively grasp the conceptual and relational types in Figure 1. A common core schema across disciplines will obviously assist interoperability, but we also envisage (like general metadata approaches such as Dublin Core [19]) that schema extensions may be required, which would be the subject of community-specific negotiation (e.g. a computer-science community might specialize *Language* into *Human-Readable* and *Computer-Readable*, and the latter into *Specification* and *Implementation Language*).

We are adopting a philosophy of ‘minimal ontological commitment’ [25] and incremental formalization [53], which reflects an emphasis on making explicit just enough structure to be usefully expressive and enable the provision of valuable computational services, but leaving the *document texts* to express the details and nuances of an author’s argument. This minimizes the effort required to submit a document description. Numerous more elaborate argumentation schemes have been proposed for computer-supported argumentation (e.g. [37], [49], [54],

[56]), but our analysis of this literature shows usability and literacy obstacles to their successful uptake (see discussion).¹ Moreover, whilst rigorous and carefully maintained argumentation networks enable more powerful computational analysis (a primary motivation for increasing a schema's expressiveness), such tools make assumptions about users' expertise and consistency of representation that are unlikely to hold in the context of an open, internet community.² The ScholOnto scheme therefore provides a relatively simple set of argumentation links to make it as easy as possible to add an argumentation link to a concept or claim; more elaborate schemes can be introduced as there is the demand.

To summarize, we propose that this represents a novel approach to the persistent problem facing any modelling effort, namely, that the world being described is typically dynamic, necessitating resource-intensive updating and restructuring of the domain model. The domain targeted by ScholOnto is the relatively consistent way in which researchers in a community present and contest new claims in a literature (whilst recognising that there are idiosyncracies between communities). Shifting the representational focus to the way in which researchers make new contributions to the literature avoids the problem of requiring commitment to discipline-specific ontologies that may become outdated and which may express a particular perspective. The representation of the domain in ScholOnto is only constructed in the context of authors' claims about their work, which are open to contest by others.

We should emphasise that we are not claiming that the richness of a literature, or the nuances of meaning in specific documents, can be reduced to a semantic network. Computational access to a literature's conceptual structure comes at the price of abstraction. What we seek to provide, firstly, is an infrastructure for explicit, summary descriptions to increase the chances of finding the right documents for more detailed reading. Secondly, this network could also become a new medium for engaging in structured, cumulative discourse that goes beyond simply mirroring material expressed in conventional documents. The network could become the cradle in which new ideas are born, exposed and tested.

This then is the rationale and vision; we describe next the implementation.

¹ An author may of course wish to represent their argument's structure at a fine-granularity, but initially, we suspect most will not have the 'literacy' skills or time to do so. There are, however, intriguing possibilities for the future role of 'non-linear argumentation' using 'native hypertext' structures, and we are following with interest the work of Kolb [34], [35], Carter [13] and others. We intend that our discourse ontology be extendible to include the kinds of new constructs and composite structures required as this strand of work develops.

² It has been suggested to us that some communities may wish to moderate their ScholOnto networks to maintain tighter control over quality. This might also create better conditions for imposing more rigorous coding and structuring of contributions. This is one of the social dimensions to the technology that we are curious to study (further discussion of how new practices might emerge can be found in §7).

5 Implementation

5.1 Knowledge modelling infrastructure

The ScholOnto server builds on a suite of robust knowledge modelling technologies developed and tested in other domains. The *OCML* modelling language [40] supports the construction of knowledge models through the specification and operationalization of functions, relations, classes, instances and rules. It also includes mechanisms for defining ontologies and problem solving methods [2], the main technologies developed in knowledge modelling research. Problem solving methods are specifications of reusable problem solving behaviours. OCML has been used in domains including medicine, geology, engineering design and organizational learning. As a result the language is now associated with a large library of reusable models. OCML provides the formalism for defining our ontology for scholarly debate and interpretation.

Knowledge engineers can collaboratively browse, edit and search OCML knowledge models over the Web using a server-Java applet architecture [18], [60]. A tool called *WebOnto* provides multiple visualizations of OCML knowledge models, a direct manipulation and forms interface for creating new knowledge structures, and a groupware facility which supports both synchronous and asynchronous joint model building. WebOnto's server is built on top of a customized web server *LispWeb* [48]. In addition to implementing the standard HTTP protocol, the LispWeb server offers a library of high-level Lisp functions to dynamically generate HTML pages, a facility for dynamically creating image maps, and a server-to-server communication method. Further details on these tools and our approach to enriching documents with ontologies can be found in [42]. A prototype system has been implemented using this infrastructure, as detailed next and in the worked example (§6).

5.2 Representing scholarly claims

Figure 2 shows the class structure of the ScholOnto ontology once specified in OCML and loaded into WebOnto, reflecting the scheme in Figure 1.

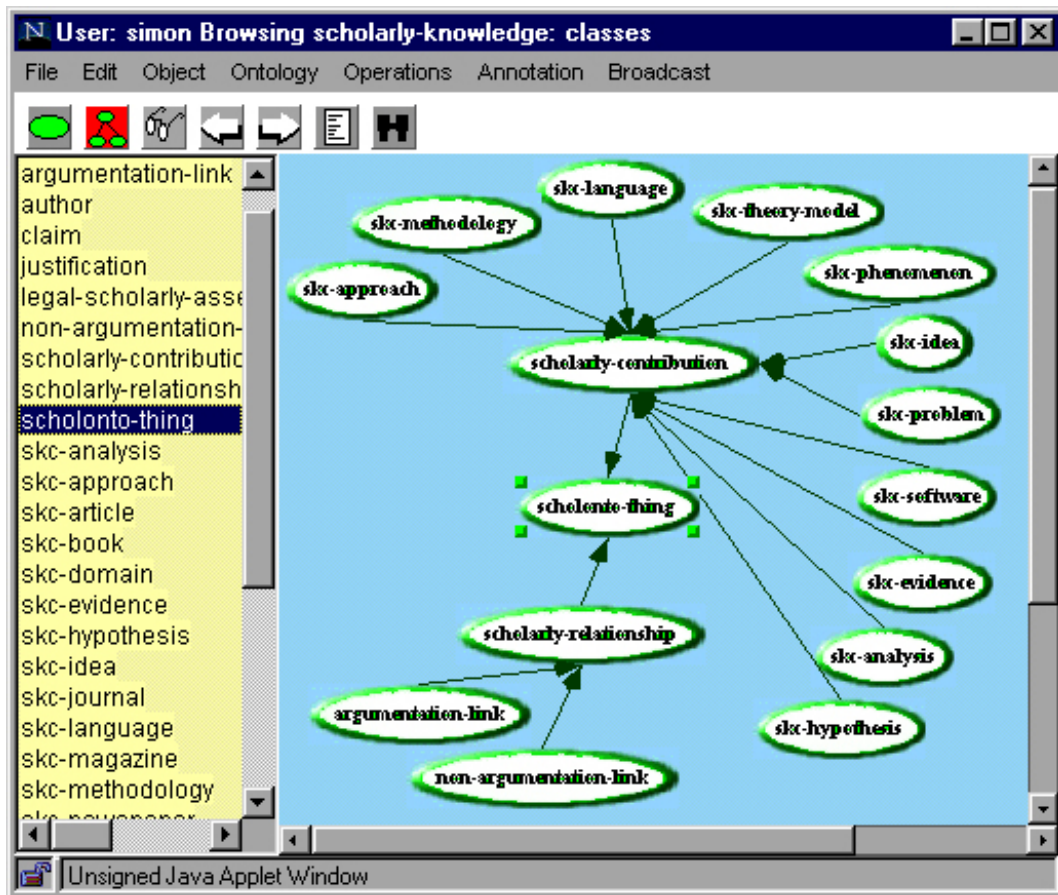


Fig. 2. Class structure of the ScholOnto ontology (whilst being edited in *WebOnto*).

Figure 3 shows how the class definition for *THEORY-MODEL* is specified in OCML.

```

(def-class skc-theory-model (scholarly-contribution)
  ;; non-argumentation relationships
  (addresses :type skc-problem)
  (analyses :type (or skc-data, skc-idea))
  (uses-applies :type (or skc-analysis, skc-approach,
    skc-data, skc-idea, skc-language, skc-methodology,
    skc-phenomenon, skc-software, skc-theory-model))
  (modifies-extends :type skc-theory-model)
  (is-an-example-of :type (or skc-theory-model, skc-data,
    skc-idea, skc-phenomenon))
  (encapsulates :type skc-theory-model)
  (envisages :type scholarly-contribution)
  (predicts :type scholarly-contribution)

  ;; argumentation relationships
  (confirms :type scholarly-contribution)
  (is-consistent-with :type scholarly-contribution)
  (is-inconsistent-with :type scholarly-contribution)
  (takes-issue-with :type scholarly-contribution)
  (raises-problem :type problem)
  (refutes :type scholarly-contribution))

```

Fig. 3. OCML provides the schema specification language underpinning ScholOnto. This extract shows how legitimate relationships for the THEORY-MODEL class of SCHOLARLY-CONTRIBUTION are defined, based on the schema summarised in Figure 1.

As introduced schematically in Figure 1, a **claim** is formally defined (Figure 4) as a relation between an agent, who makes a legal-scholarly-assertion, with some justification. A legal-scholarly-assertion is a statement instantiating a scholarly-relationship (e.g. **addresses**, **predicts**, **refutes**) between two elements (e.g. methodology X addresses problem Y). The justification may be free text supporting a claim, but more substantively, either a document (which may have its own associated conceptual structure) or a specially created structure to serve as backing.

```

(def-class legal-scholarly-assertion (assertion) ?x
  :iff-def (and (assertion ?x)
                (== ?x (?a ?b ?c))
                (scholarly-relationship ?a)))

(def-relation claims (?X ?Y ?Z)
  :constraint (and (set ?X) (every ?x agent)
                  (legal-scholarly-assertion ?Y)
                  (justification ?Z)))

```

Fig. 4. Defining a legal scholarly assertion and a claim in OCML. Separating claims from scholarly relationships makes possible multiple (possibly competing) claims by different authors.

OCML's environment makes it easy to construct rules that could, for instance, check for instances where an author is a member of two sets of authors who have made conflicting assertions (Figure 5). To both `support` and `refute` a particular claim is usually a sign of inconsistency, or perhaps, of a position that has changed from an earlier paper.

```

(def-relation inconsistent-position (?auth ?assertion)
  :constraint (and (author ?auth)
                  (legal-scholarly-assertion ?assertion)))
  :sufficient (and (member ?auth ?x) (supports ?x ?y ?z)
                  (member ?auth ?x2) (refutes ?x2 ?y ?z2)))

```

Fig. 5. An OCML rule for an agent to check for positions that both `support` and `refute` a particular claim. This might reflect an `inconsistent-position` or at least, claims meriting closer examination.

Whilst not yet implemented, ScholOnto should enable researchers to encapsulate such rules as agents that could be reused and customized from a library.

5.3 From claims to perspectives

The incremental submission to ScholOnto of document descriptions would over time create a complex network of concepts and relationships, which we have been referring to as *claims*. This complexity, on the one hand, should provide the richness that will make ScholOnto a helpful reflection of the literature and ongoing debate, but the system must also provide support for managing this complexity. In the discussion we discuss intelligent services that

the knowledge model can support, and so we focus here on one particularly important ‘higher order’ structure that could emerge from the proposed ontology.

The vague notion of a ‘school of thought’ has at least one specific representational form within ScholOnto, which we refer to as a *perspective*. A perspective is defined as a common set of concepts on which one (but usually several) researchers build their work, typically, the common set of THEORY-MODELS, perhaps the associated METHODS they deploy, or the body of EVIDENCE to which they appeal. Conversely, the concepts that they collectively TAKE-ISSUE-WITH may represent a different perspective. The ontology thus provides a structural approach to detecting the emergence of ‘camps’ within a literature, in which the basis for one cluster of documents is challenged by another cluster. A ScholOnto agent could therefore be set to monitor the network for *contrasting perspectives* defined as structures where ≥ 3 documents support ideas of type A and challenge ideas of type B, and ≥ 3 documents do the converse.

The detection of structural corollaries of perspectives would also make use of *link families*, codified within the ontology. A natural family exists, for instance, concerning relationships that indicate a *difference of opinion* of some sort, varying in strength from RAISES-PROBLEM, TAKES-ISSUE-WITH or IS-INCONSISTENT-WITH, to the very strong REFUTES (we anticipate that authors will want to exercise control over how their work is linked to others’ work). Thus, an agent could be set to (in essence) *show me any papers that ‘have a problem’ with this theory*. In addition, if a query is posed to the system that implicates the milder IS-INCONSISTENT-WITH relationship, it is reasonable for the system to infer that the stronger REFUTES relationship will also be of interest, and use this to retrieve potentially relevant material. Our current search engine provides an explicit *rationale* to explain why an item has been returned, using relational types to show and navigate the lineage (see Figure 10). Heuristic rules operating on link families thus enable the computation of inferred links that have not been explicitly codified.

6 A Worked Example

We now return to our opening examples of literature and document interpretation, and use a worked example to clarify how our tools could support scholarly work. Within the hypertext research literature, one of the landmark articles is the summary of the *Dexter Hypertext Reference Model* by Halasz and Schwartz [28], which specifies both semiformaly (and in a fuller report formally using the Z notation), abstract properties of hypertext systems, enabling comparison of existing systems, and specification of theoretically possible future systems. Following publication, the Dexter model was extended and refined by many other researchers, and formed the basis for a variety of prototype system (e.g. [12], [16], [23]).

We are exploring a variety of user interfaces to make interaction with a ScholOnto server as natural as possible. Firstly, there are conventional forms and menus. Figure 6 shows a user interface design for submitting the Dexter article's description to the repository. The *KNote* tool already generates forms with contextual menus for adding new instances to an OCML ontology [42], but needs to be refined to generate forms that non-knowledge engineers can use. The forms interface would guide users through the schema using dynamic menus, and enable them to browse and search for existing concepts to assist their reuse. Some domain concepts are simple to reference (e.g. the name of a specific software system, framework or methodology), whilst others are less concrete, e.g. finding the name(s) used to describe a domain problem ("user disorientation"), an idea ("a global hypertext system"), or an empirical phenomenon such as a piece of evidence ("low ability students benefit most from physics simulations"). A design challenge is to actively support users in locating relevant registered concepts for reuse, and assisting them in linking new concepts into the network.

An alternative interface that we are considering is direct annotation of the abstract/document with ScholOnto concepts (assuming there is an online version), as illustrated by the annotation tool in Figure 7. This provides the opportunity to link to specific points in the document, and is particularly useful when the concepts one wishes to register/cite within ScholOnto are already expressed in a succinct form in the text (e.g. the names of theories, datasets, methods); these can be registered directly with no further editing. Other, more complex concepts will need to be summarised by the author with a succinct label, but via a 1-N link, a concept can point to multiple points in the document if desired.

This document concerns:

▼ **Design Domain** Select from ACM CCS keywords (optional)
 Hypertext/hypermedia
 New...

H.1 MODELS AND PRINCIPLES
 H.2 DATABASE MANAGEMENT (E.5)
 H.3 INFORMATION STORAGE

▼ **ACM CCS**
 ▼ **H: Information Systems**
 H.0 GENERAL
 H.1 MODELS AND PRINCIPLES
 H.2 DATABASE MANAGEMENT (E.5)
 H.3 INFORMATION STORAGE AND RETRIEVAL
 H.4 INFORMATION SYSTEMS APPLICATIONS
 H.5 INFORMATION

Citation:
 Halasz, F. and Schwartz, M. (1994). Comm. of the ACM, 37 (2), 30-39

URL: www.acm.org/pubs/citations/journals/cacm/1994-37-2/p30-halasz/

This article describes a new:

▼ **Theory/Model**
 Dexter Hypertext Reference Model
 Submit...

▼ **<none>**
 Submit...

New...

▼ **Theory/Model**
 ▼ **Design Domain**
 Hypertext/hypermedia
 Dexter Hypertext Referenc

Relationships to other articles/concepts:

Dexter Hypertext Reference Model

▼ **Analyses** ▼ **Software**
 Augment
 Concordia
 Hypercard

▼ **Predicts** ▼ **Software**
 Theoretically possible Dexter compliant syst

▼ **Uses/Applies** ▼ **Specification Language**
 Z

▼ **Language**
 ▼ **Design Domain**
 State Transition Networks
 Task Action Grammer
 User Action Notation
 YDM
 Z

▼ **Problem**
 ▼ **Design Domain**
 Hypertext/hypermedia
 Commercial hypertexts
 ...absence of
 Educational hypertext
 ...weak evidence
 Navigation
 ...disorientation
 Standards
 ...absence of

Analyses
 Addresses
 Uses/Applies
 Modifies/Extends
 Predicts
 Supports
 Raises Issues With
 Refutes

Software
 Methodology
 Language
 Theory/Model
 Phenomenon
 Idea
 Problem

Sub-Problem of...
 Variation on...
 New...

Fig.6. User interface for submitting claims about a document's contributions (mockup). The top section displays standard bibliographic metadata of the sort which might already exist in a digital library. The middle section asks for the novel contributions of the documents, and the lower section the key relationships to the existing literature. These additional layers of description give the document its presence in the semantic network.

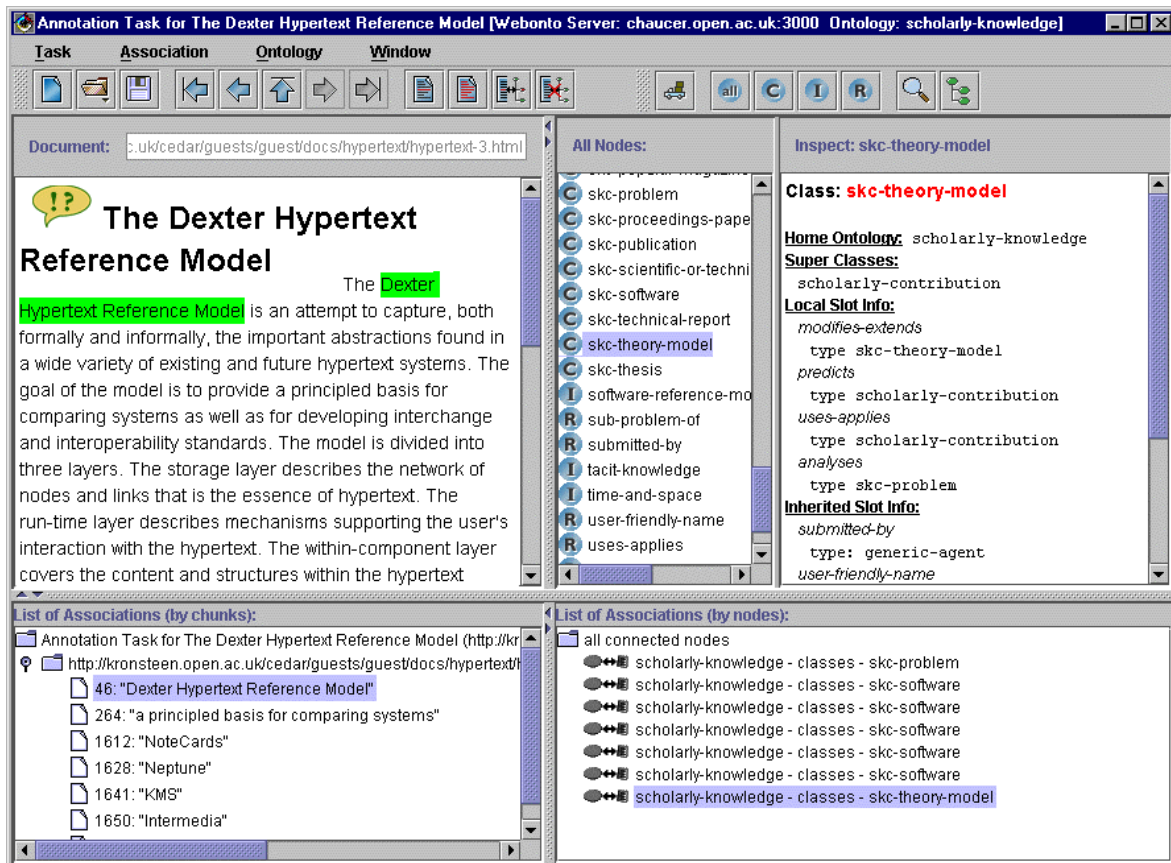


Fig. 7. Annotating a document directly with ScholOnto concepts (annotation tool built by Enrich Project [21]).

A third mode of user interaction, given that the *semantic network* is the dominant conceptual model for the system, is for users to express claims by drawing them as maps. Depending on individual cognitive style, and the complexity of the structure being built, a well designed visual notation can provide affordances that forms and direct annotation support poorly (e.g. multiple links between multiple nodes). Figure 8 shows a visual description of the Dexter article using a hypermedia mapping tool.³

³ We have adapted the *Compendium* approach to conversational modelling and group memory construction [51]. This provides a rapid modelling environment for prototyping ScholOnto structures, as well as a possible future client user interface.

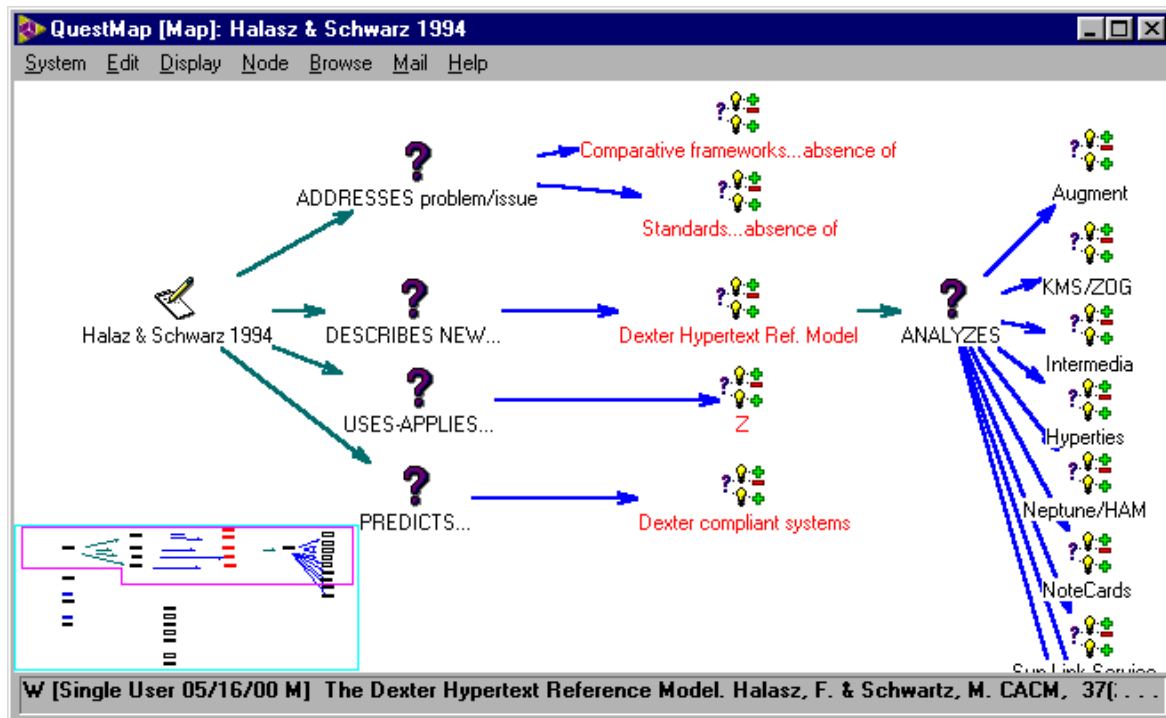


Fig. 8. Expressing a document's semantics using a concept mapping interface.

Once the description is submitted, the server would generate an OCML entry in the ontology, as shown in Figure 9.

```

(def-instance dexter-htxt-ref-model-article article
  ((describes-scholarly-contribution-element
    dexter-htxt-ref-model
    (concerns-domain hypertext-hypermedia)
    (has-author halasz-f schwartz-m)
    (has-title "The Dexter Hypertext Reference Model")
    (publication-details "Communications of the ACM, 37
      (2), 30-39")
    (has-url "www.acm.org/pubs/articles/journals/cacm/
      1994-37-2/p30-halasz/")
    (acm-ccs "I.7.2" "H.1.1" "H2.1" "H3.2" "H5.1")))

(def-instance dexter-ht-ref-model theory-model
  ((addresses absence-of-standards
    (analyses notecards augment hypercard hyperties
      intermedia kms-zog neptune-ham)
    (envisages theoretically-possible-dexter-compliant-
      systems)
    (uses-applies Z)))

```

Fig. 9. The OCML entry for the Dexter article, declaring its contributions to the literature (`dexter-ht-ref-model`, which is a `theory-model`, which predicts `dexter-compliant-systems`), and its relationship to other concepts (it `analyses` several existing systems, and `uses-applies` the `Z` notation).

The article is now registered in the knowledge base, enabling users to formulate queries that would help answer *What motivated the Dexter Hypertext Reference Model, and what impact has it had?* The forms-based *Lois* search interface guides the user through the construction of such queries, and currently generates an HTML interface which enables the user to navigate by following concepts associated with search ‘hits’ (Figure 10).

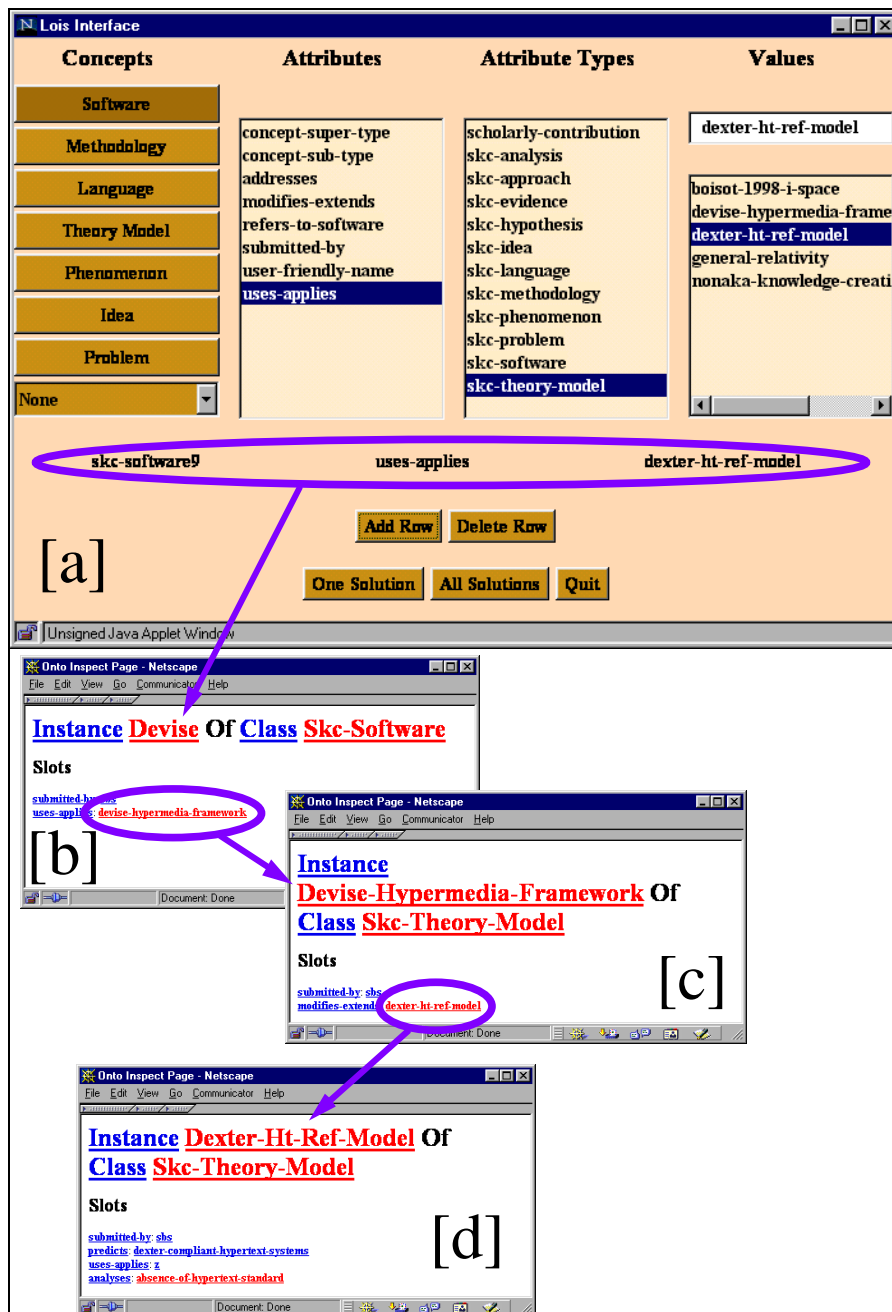


Fig. 10. [a] The *LoIs* search interface guides the formulation of structural queries in the ontology via active menus. This example shows a search for software which uses-applies the Dexter-Hypertext-Ref-Model. The following windows show how concepts associated with a search result can then be explored (at present via an HTML interface onto the OCML knowledge base). [b] A system called *DeVise* is found, which uses-applies the *DeVise-hypermedia-framework*. [c] Following this link displays the entry for this concept, revealing that it modifies-extends the Dexter Model, whose entry is opened in [d]. Since concepts *using* or *modifying* a concept specified in a search may also be of interest to a researcher, a search heuristic would ensure that the *DeVise* system was returned as a hit, even though (in this example [b]) its use of the Dexter Model was not explicitly codified.

In addition to textual views and navigation, we also envisage filtered visualizations of the literature (e.g. Figure 11) showing the Dexter Model’s motivation and conceptual roots (links to the left), and the nature of the work which has built on it since by the respective authors, or other researchers (links to the right).

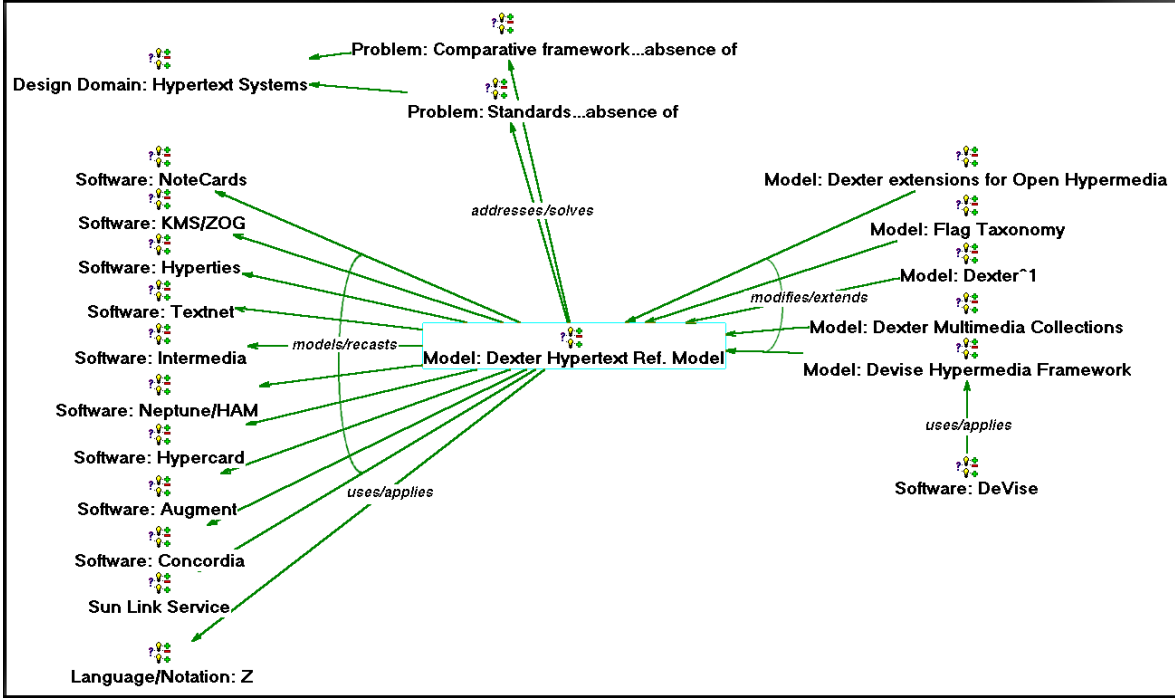


Fig.11. A semantic network model provides the basis for generating visualizations of the literature, a simulation of which is shown above. The motivation for the *Dexter Model* (central node) is shown by links to the left (the *problems* it sought to address, the *systems* it analyzed). A measure of the model’s impact is shown by links to the right, showing work that *modified/extended* the model, and implemented software systems based on these extended models (e.g. the *DeVise* system, lower right).

Modelling hypertext and artificial intelligence literatures using a tool built on such technologies follows our ‘bootstrapping’ approach to testing our own tools, but there are obvious dangers of reflexivity and limited scope. We are therefore modelling completely unrelated literatures validate the ontology’s expressiveness, and are designing ontology and user interface usability studies.

7 Discussion

We turn now to discuss some of the issues raised by ScholOnto, new possibilities it opens up, and contextualise our approach to related work.

7.1 Intelligent services

The existence of a formally represented knowledge model makes possible new kinds of knowledge management and reasoning services that will help make sense of the network. As a knowledge model support inference-based searching and alerting, the system allows us to write heuristics that could assist in finding relevant documents, e.g. “if `Method Y` extends `Method X`, and `Method X` is challenged, then `Method Y` may be challenged”. As described earlier in the context of detecting *perspectives*, ScholOnto’s heuristic knowledge and machine learning techniques will also support agents that make it possible, for instance, to discover a ‘European perspective’ on a particular issue, by analyzing the geographic spread of authors whose concepts contributed to a significant conceptual structure in the network.

The *impact* of a piece of work can be represented in the proposed ontology in various ways. ScholOnto’s ontology and rules provide the basis for ‘semantic citation analysis’, showing who has cited either a document, or a concept associated with it (or derived from it), and *why*. Other useful indices might include the level of *reuse* of a concept (directly, or via its intellectual descendents), or the number of different domains or problems to which it has been applied.

As with the internet in general, some notion of *authority* becomes increasingly important as we seek to filter out noise. In addition to quantitative measures of impact as just outlined, ScholOnto in principle could make additional qualitative criteria available to a researcher, who could declare that a particular research group, methodology or theoretical perspective should carry more weight, and serve as a filter on the literature.

7.2 Codification: the double-edged sword

Well designed codification schemes, apart from enabling useful computational services, can reduce cognitive overload by guiding users’ attention to relevant subsets of concepts, and by restricting interaction in appropriate ways. However, there is a critical tradeoff to negotiate. We are acutely aware that many schemes for providing structured descriptions founder on the ‘capture bottleneck’—the envisaged beneficiaries of the system simply do not have the

motivation, skill or time to invest in codifying shared resources to build a critical mass of useful material. We draw sobering lessons on the codification issue from our analyses of the computer-supported argumentation literature [8], corroborated by evidence relating to groupware [26], design rationale support [9], organizational memory systems [6], [50], and many other CSCW systems that require users to formalize information [52]. Specifically for ScholOnto, we need to address issues raised by the usage patterns of hypertext systems such as Textnet [57] for scholarly annotation and linking, and its descendants such as NoteCards [27] and Aquanet [39]. These provided rich schemas of semantic node-link types, but the limited adoption of these features led Trigg and others to conclude, correctly in our view, that rich taxonomies of node and link types often overwhelm users.

However, rather than eradicate all human-encoded semantics from our systems, more effective deployment may depend on finding the right mix of target domain, context of use and user community. For ScholOnto, these factors have unique characteristics lacking in many other domains where group memory, knowledge structuring and argumentation have been previously tested:

- *Academia*. Closer attention is paid in research and academic study, in contrast to many other domains, to coherence, argumentation and evidence, making the initial investment of effort more worthwhile. (Students might also be asked to analyse and build conceptual networks in a given literature—cf. Figure 8, and related work on educational concept mapping [33], [44].)
- *Strong motivation to disseminate work*. Making an impact through publications is a primary activity for academics. Having completed a new document, *if* the benefits outweigh the costs, we hypothesise that an author will want to maximise its ‘digital presence’ by taking time to encode its contributions and connections to the existing literature more carefully than a set of undifferentiated keywords.
- *Time to reflect on how to represent ideas*. In contrast to the synchronous group working context in which many argumentation/concept mapping systems have been proposed (but often failed), scholars will be describing their documents in an individual setting, with time to reflect on how best to construct a network description.
- *A simple semantic schema complementing text*. We hypothesise that most researchers will have no trouble in understanding the kinds of node and link types in Figure 1; they are building blocks in research discourse. Nor are we requiring an author to make explicit a document’s thesis at a fine granularity; the structural representation is a summary to assist the document’s discovery, not a substitute.
- *Benefits of internet scalability*. Previous research with pre-internet groupware has always focused on individual or small scale collaborative use. In a small team who already know each others’ work, it is often hard to justify the overheads of information

structuring in order to track documents and debates. This is in sharp contrast to the challenge of tracking and analysing developments in an international, evolving digital library. Using the Web as our collaboration and delivery medium has a second advantage: the size of a Web-based research community increases the chances of quickly building a critical mass of users, which will in turn improve the value of the services provided.

7.3 Emergent work practices

If a significant proportion of a research community adopted a digital library infrastructure of the sort proposed, such that it became a de facto standard to register new documents on the server and engage in debate via a semantic network, it is likely that it would effect a shift in working practices. The public face of an author's work would extend to include corresponding concept maps and libraries. Formulating the contributions of an article such that this map was succinct and to the point could in fact improve the authoring process. A research project or lab would most likely evolve a 'library' of conceptual elements and structures, which provided a resource for others to link to their work.

A research field has a small number of recognisable 'genres' of article which an active member of that field can usually recognise very quickly. Indeed, these are formally recognised by some conferences and journals in order to guide reviewers on appropriate review criteria for different kinds of submission (systems paper, theory paper, empirical paper, etc.). We can envisage the possibility of a journal establishing a set of templates for submitting and reviewing submissions in different categories, for example, "a paper of type-X should reasonably demonstrate in its conceptual network representation its relationships to concepts of type A, B and C, and we expect arguments of the form P, Q and R...". The system would also assist reviewers in locating related work and debates that the author had omitted.

7.4 Related work

Ontologies are beginning to be used in the context of digital libraries, although for different purposes to those presented here. Ontologies can assist the extraction of concepts from unstructured textual documents [20], by serving as a source of knowledge about the particular topic. Ontologies can also assist in managing document descriptions in large digital libraries [61].

We noted earlier the emergence of open e-print server architectures. A ScholOnto server as proposed here extends this infrastructure by adding a semantically enriched layer of document encoding, with associated services as discussed. A strategy for migrating from a conventional

e-print (or other digital library) server to a ScholOnto server would be to enable authors to annotate existing e-print server archive entries in order to give them a presence in the semantic network. One could either mirror the e-print server database in ScholOnto, or perhaps more flexibly, link the two in a more decentralised manner, enabling e-print server users to launch a ScholOnto annotation interface for a given article as an added service.

Our technologies can be seen as a conceptual and technical development from current efforts to develop metadata description schemes for the Web. Metadata in the context of digital libraries refers to ways to encode information about resources in machine-interpretable formats, typically by completing a standard set of descriptive fields. Well known examples include USMARC [38] for library resources, Dublin Core [19] to provide a simple high level scheme for web resources, and IMS [31] for educational resources. Our scheme for representing multiple, possibly conflicting claims about the status of content extends typical metadata schemes which normally focus on encoding uncontroversial content attributes.

From a representational and technical perspective, our approach differs from metadata in that current ontology specification languages support more sophisticated modelling, for example, specifying sufficient and necessary conditions for relations, and providing metalevel modelling support which makes it possible to reason about the ontology itself. OCML in particular also provides powerful inference support making it possible to directly operationalise the ontology and its instantiation as a knowledge base. The W3C's proposed Resource Description Framework [58] employs semantic modelling similar to the knowledge modelling concepts presented here. Exporting OCML to RDF from WebOnto is one route to interoperability with other systems and metadata schemes, but any reasoning capability would be lost without an RDF interpreter/convertor and inference engine (see [17] for further discussion). In addition, should exports to other modelling notations be required, it is simple to translate from OCML to, for instance, Ontolingua [24].

Avoiding the approach of encoding link structure directly into documents leads us, therefore, to a critique of 'ontological HTML-markup' (e.g. [3], [29]). For instance, the (KA)² initiative (Knowledge Annotation for Knowledge Acquisition) [3] aims to support the knowledge acquisition community in building a knowledge base of its own research by populating a shared ontology. The knowledge base is constructed by authors annotating their web pages (e.g. publications; personal and project pages) with tags (analogous to HTML META tags), which can be read by a specialized search engine called Ontobroker [22]. The key architectural difference to our approach is that (KA)² semantic tags are embedded in the physical content, whereas our approach decouples content from claims about its status. This architectural difference reflects the different aims of the two enterprises. The aim of (KA)² is to capture the contents of web pages in a formalism which can be reasoned about by a single

knowledge service (Ontobroker), with a 1-1 mapping between content and the assertion one wishes to make within the ontology. In our scenario we want to represent the *debate* about the *scholarly status* of documents (which may not be online, and even if so, one will not normally have write-access). Moreover, as argued earlier we cannot assume a stable ontology to describe an active research field (see [42] for further details of how this work relates to other knowledge modelling research).

The Open Hypermedia research community seeks to integrate hypermedia functionality into conventional computing environments [47]. One of the key principles is the separation of content from hypertext links, the latter becoming first class objects stored in their own linkbases. ScholOnto is premised on the notion of content-structure separation in order to express competing claims and discourses: multiple relational structures can be layered over a given digital library. Links in ScholOnto are therefore first class objects in that a *claim* is a triple (node-link-node) which can itself be linked to. (In the context of this paper, an interesting application of open hypermedia is the overlaying of active citation links on conventional documents in e-print archives [30].) Finally, this work can be seen as an example of Structural Computing, a generalization of work on open hypermedia, as proposed by [45]. Elsewhere we have considered ScholOnto's relationship to this emerging research field [7].

8. Summary and Future Work

The Web has established itself as a medium for research document dissemination. However, its support for many of the interpretive tasks that scholars perform is weak, despite the fact that semantic hypertext systems (as the Web was originally envisaged by Berners-Lee) afford structural analysis and knowledge-based reasoning. We are developing a representational and technical infrastructure to complement libraries of archived documents with a 'living' semantic network of concepts and discourse. We propose that this will have advantages for tracing the intellectual lineage of a document's ideas, assessing the subsequent impact of those ideas, and engaging in dialogue with other authors through the cumulative construction of a network. The underlying knowledge base is crucial for managing the complexity of the resulting network, as well as opening new possibilities for automatic analysis of a community's collective knowledge. Adopted by a research community, this network could reflect the *evolving understanding and recontextualization of ideas* over time.

To summarise, the work reported makes the following conceptual and technical contributions:

- an approach to the distributed construction of a knowledge base, by a community working in a changing domain, holding potentially conflicting views about the structure of their domain;
- a representational scheme which seeks to balance simplicity with expressiveness in providing a language for researchers to describe and debate their work, and in so doing create authorial perspectives which can be computationally detected and critiqued by peers;
- an approach to tackling the problems that have obstructed previous efforts to represent argumentation in hypertext systems;
- the adaptation of a knowledge-based web server, Java ontology management tools and underlying knowledge modelling language to implement the digital library server.

We are currently refining the ontology as we model a variety of literatures to evaluate ScholOnto's usability and digital library services. We are encouraged by the small scale modelling and literature analyses that we have undertaken ourselves, but the broader scope and acceptability of the infrastructure is as yet unproven. The suite of tools illustrated requires further development and integration, coupled with better user interfaces to access, navigate and visualize the semantic network. We welcome approaches from communities interested in testing early releases of the ScholOnto server as a means of managing their own research knowledge.

Acknowledgements. This research was partially funded by the UK EPSRC (GR/M78151), and the CEC-funded *Enrich* (P29015) and *PatMan* (P4017) projects. *Enrich* is part of the ESPRIT programme on IT for Learning and Training in Industry; *PatMan* is part of the Healthcare Sector of the Telematics Application Programme.

References

1. Allan, J. Automatic Hypertext Link Typing. In *Proc. Hypertext'96*, 1996, ACM Press: NY, pp. 42-52
2. Benjamins, V.R. and Fensel, D. Special Issue on Problem Solving Methods (Eds.). *International Journal of Human-Computer Studies*, 49, 4, 1998
3. Benjamins, V.R., Fensel, D., Decker, S. and Asuncion, G.P. (KA)²: Building Ontologies for the Internet: A Mid-term Report. *International Journal of Human Computer Studies*, 51, 3, 1999, pp. 687-712
4. Bieber, M., Vitali, F., Ashman, H., Balasubramanian, V. and Oinas-Kukkonen, H. Fourth Generation Hypermedia: Some Missing Links for the World Wide Web. *International Journal of Human-Computer Studies (WWW Usability Special Issue)*, (Eds.) S. Buckingham Shum & C. McKnight, 47, 1, 1997, pp. 31-65
<<http://www.hbuk.co.uk/ap/ijhcs/webusability/>>

5. Bradshaw, S., Hammond, K. and Scheinkman, A. Guiding People to Information: Providing an Interface to a Digital Library Using Reference as a Basis for Indexing. In *Proc. IUI 2000: International Conference on Intelligent User Interfaces*, New Orleans, Louisiana, USA - Jan. 9-12, 2000, ACM Press: NY
6. Buckingham Shum, S. Negotiating the Construction and Reconstruction of Organisational Memories. *Journal of Universal Computer Science (Special Issue on Information Technology for Knowledge Management)*, 3, 8, 1997, pp. 899-928 <http://www.iicm.edu/jucs_3_8/>
7. Buckingham Shum, S., Domingue, J. and Motta, E. Scholarly Discourse as Computable Structure. In *Proc. SC2: Second International Workshop on Structural Computing*, San Antonio, Texas. June 3, 2000. Lecture Notes in Computer Science: Springer-Verlag: Berlin
8. Buckingham Shum, S. and Hammond, N. Argumentation-Based Design Rationale: What Use at What Cost? *International Journal of Human-Computer Studies*, 40, 4, 1994, pp. 603-652
9. Buckingham Shum, S., MacLean, A., Bellotti, V. and Hammond, N. Graphical Argumentation and Design Cognition. *Human-Computer Interaction*, 12, 3, 1997, pp. 267-300. PrePrint: <<http://kmi.open.ac.uk/kmi-abstracts/kmi-tr-25-abstract.html>>
10. Buckingham Shum, S. and Sumner, T. New Scenarios in Scholarly Publishing and Debate. In *The Knowledge Web: Learning and Collaborating on the Net*, Eisenstadt, M. and Vincent, T., (Ed.), Kogan Page: London, 1998, pp. 135-151 <<http://kmi.open.ac.uk/knowledgeweb>>
11. Bunge, M.A. *Treatise on Basic Philosophy, Vol. 3. Ontology I: The Furniture of the World*. Reidel: Boston, 1977
12. Special Issue on the Dexter Hypertext Reference Model: *Communications of the ACM*, 37, 2, 1994, pp. 26-86 <<http://www.acm.org/pubs/contents/journals/cacm/1994-37/#2>>
13. Carter, L.M. Arguments in Hypertext: A Rhetorical Approach. In *Proc. Hypertext 2000*, San Antonio, TX, 2000, ACM: New York
14. Chen, C. and Czerwinski, M. From Latent Semantics to Spatial Hypermedia: An Integrated Approach. In *Proceedings of Hypertext'98: 9th ACM Conference on Hypertext*, Pittsburgh, 1998, ACM Press: New York
15. *Digital Document Discourse Environment*. Knowledge Media Institute, Open University <d3e.open.ac.uk>
16. De Bra, P., Houben, G.-J. and Wu, H. AHAM: a Dexter-Based Reference Model for Adaptive Hypermedia. In *Proc. Hypertext'99: 10th ACM Conference on Hypertext*, Darmstadt (21-25 Feb.), 1999, ACM Press: New York, pp. 147-156 <http://www.acm.org/pubs/citations/proceedings/hypertext/317601/p147-de_bra/>
17. Decker, S., Brickley, D., Saarela, J. and Angele, J. A Query and Inference Service for RDF. In *QL'98: The W3C Query Languages Workshop*, Boston (Dec. 3-4, 1998), 1998 <<http://www.w3.org/TandS/QL/QL98/pp/queryservice.html>>
18. Domingue, J. Tadzebao and WebOnto: Discussing, Browsing, and Editing Ontologies on the Web. In *Proc. KAW'98: 11th Banff Knowledge Acquisition Workshop*, Banff, Canada, 1998, Dept. Computer Science, University of Calgary, CA. PrePrint: <<http://kmi.open.ac.uk/~john/banff98-paper/domingue.html>>
19. Dublin Core Metadata Initiative <<http://purl.oclc.org/dc/>>
20. Embley, D.W., Campbell, D.M., Smith, R.D. and Liddle, S.W. Ontology-Based Extraction and Structuring of Information from Data-Rich Unstructured Documents. In *Proc. CIKM'98: Conference on Information and Knowledge Management*, Bethesda, MD, USA, 1998, ACM Press: New York, pp. 52-59
21. *ENRICH Project: Enriching Representations of Work to Support Organisational Learning*. Knowledge Media Institute, Open University, UK <<http://kmi.open.ac.uk/projects/enrich/>>
22. Fensel, D., Decker, S., Erdmann, M. and Studer, R. Ontobroker: The Very High Idea. In *Proc. FLAIRS-98: 11th Annual Florida Artificial Intelligence Research Symposium*, Florida, 1998
23. Grønbaek, K., Bouvin, N.O. and Sloth, L. Designing Dexter-Based Hypermedia Services for the World Wide Web. In *Proceedings of Hypertext'97: 8th ACM Conference on Hypertext*, Southampton, 1997, ACM Press: New York, pp. 146-156 <<http://journals.ecs.soton.ac.uk/~lac/ht97/pdfs/gronbak.pdf>>
24. Gruber, T.R. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5, 2, 1993, pp. 199-220
25. Gruber, T.R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies*, 43, 5/6, 1995, pp. 907-928
26. Grudin, J. Groupware and Social Dynamics: Eight Challenges for Developers. *Communications of the ACM*, 37, 1, 1994, pp. 92-105

27. Halasz, F.G. Reflections on Notecards: Seven Issues for the Next Generation of Hypermedia Systems. *Communications of the ACM*, 31, 1988, pp. 836-852
28. Halasz, F.G. and Schwartz, M. The Dexter Hypertext Reference Model. *Communications of the ACM*, 37, 2, 1994, pp. 30-39
29. Heflin, J., Hendler, J. and Luke, S. Reading Between the Lines: Using SHOE to Discover Implicit Knowledge from the Web. In *AAAI-98 Workshop on Artificial Intelligence and Information Integration*, Stanford, CA, 1998 <<http://www.cs.umd.edu/projects/plus/SHOE/shoe-aaai98.ps>>
30. Hitchcock, S., Carr, L., Jiao, Z., Hall, W. and Harnad, S. Developing Services for Open E-Print Archives: Globalisation, Integration and the Impact of Links. In *Proc. ACM Digital Libraries Conference*, San Antonio, Texas, 2000, ACM Press: New York
31. The Instructional Management System (IMS) Project. <www.imsproject.org>
32. *Journal of Interactive Media in Education*. Knowledge Media Institute, Open University <www-jime.open.ac.uk>
33. Special Issue on Concept Mapping: *Journal of Research and Science Teaching*, 27, 10, 1990
34. Kolb, D. *Socrates in the Labyrinth: Hypertext, Argument, Philosophy (A Hypertext)*. Eastgate Systems: Watertown, 1994
35. Kolb, D. Scholarly Hypertext: Self-Represented Complexity. In *Proceedings of The Eighth ACM Conference on Hypertext*, Southampton, 1997, pp. 29-37
36. Lawrence, S., Giles, L.C. and Bollacker, K. Digital Libraries and Autonomous Citation Indexing. *IEEE Computer*, 32, 6, 1999, pp. 67-71 <<http://www.neci.nj.nec.com/~lawrence/papers/aci-computer98/>>
37. Lee, J. SIBYL: A Qualitative Decision Management System. In *Artificial intelligence at MIT: Expanding Frontiers*, Winston, P. and Shellard, S., (Ed.), MIT Press: Cambridge, Massachusetts, 1990, pp. 105-133
38. USMARC Format for Bibliographic Data, Including Guidelines for Content Designation. Network Development and MARC Standards Office, Library of Congress, Cataloging Distribution Service, Washington DC, 1994
39. Marshall, C.C. and Rogers, R.A. Two Years before the Mist: Experiences with Aquanet. In *Proceedings of the Fourth ACM Conference on Hypertext*, 1992, pp. 53-62
40. Motta, E. An Overview of the OCML Modelling Language. In *Proceedings KEMML'98: 8th Workshop on Knowledge Engineering Methods & Languages*, Karlsruhe, Germany, Jan. 21-22, 1998, 1998
41. Motta, E. *Reusable Components for Knowledge Modelling*. IOS Press: Amsterdam, NL, 1999
42. Motta, E., Buckingham Shum, S. and Domingue, J. Ontology-Driven Document Enrichment: Principles and Case Studies. In *Proc. KAW'99: 12th Banff Knowledge Acquisition Workshop*, Banff, Canada, 1999, Dept. Computer Science, University of Calgary, CA <<http://kmi.open.ac.uk/projects/scholonto/>>
43. Newell, A. The Knowledge Level. *Artificial Intelligence*, 18, 1, 1982, pp. 87-127
44. Novak, J.D. Concept Maps and Vee Diagrams: Two Metacognitive Tools to Facilitate Meaningful Learning. *Instructional Science*, 19, 1990, pp. 1-25
45. Nürnberg, P.J., Leggett, J.J. and Schneider, E.R. As We Should Have Thought. In *Proceedings of Hypertext'97: 8th ACM Conference on Hypertext*, Southampton, 1997, pp. 96-101 <<http://journals.ecs.soton.ac.uk/~lac/ht97/pdfs/nuern.pdf>>
46. *Open Archives Initiative* <<http://vole.lanl.gov/ups/>>
47. *Open Hypermedia Systems Working Group* <<http://www.ohswg.org/>>
48. Riva, A. and Ramoni, M. LispWeb: a Specialized HTTP Server for Distributed AI Applications. *Computer Networks and ISDN Systems (Proc. WWW96 Conference)*, 28, 7-11, 1996, pp. 953-961 <<http://kmi.open.ac.uk/~marco/papers/www96/www96.html>>
49. Schuler, W. and Smith, J. Author's Argumentation Assistant (AAA): A Hypertext-Based Authoring Tool for Argumentative Texts. In *Proc. ECHT'90: European Conference on Hypertext: Argumentation, Design & Knowledge Acquisition*, Cambridge, 1990, Cambridge University Press, pp. 137-151
50. Selvin, A. Supporting Collaborative Analysis and Design with Hypertext Functionality. *Journal of Digital Information*, 1, 4, 1999 <<http://jodi.ecs.soton.ac.uk/Articles/v01/i04/Selvin/>>
51. Selvin, A.M. and Buckingham Shum, S.J. Rapid Knowledge Construction: A Case Study in Corporate Contingency Planning Using Collaborative Hypermedia. In *Proc. KMaC 2000: Knowledge Management Beyond the Hype*, Birmingham, UK (16-19 July 2000), 2000, Operations Research Society <<http://kmi.open.ac.uk/kmi-abstracts/kmi-tr-92-abstract.html>>

52. Shipman, F.M. and Marshall, C.C. Formality Considered Harmful: Experiences, Emerging Themes, and Directions on the Use of Formal Representations in Interactive Systems. *Computer Supported Cooperative Work*, 8, 4, 1999, pp. 333-352. PrePrint: <<http://bush.cs.tamu.edu:80/~shipman/cscw.pdf>>
 53. Shipman, F.M. and McCall, R. Supporting Knowledge-Base Evolution with Incremental Formalization. In *Proc. ACM CHI'94: Human Factors in Computing Systems*, Boston, Mass., 1994, ACM Press: New York, pp. 285-291
 54. Smolensky, P., Fox, B., King, R. and Lewis, C. Computer-Aided Reasoned Discourse Or, How to Argue with a Computer. In *Cognitive Science and its Application to Human-Computer Interaction*, Guindon, R., (Ed.), Ablex: Norwood, New Jersey, 1988, pp. 109-162
 55. Sumner, T. and Buckingham Shum, S. From Documents to Discourse: Shifting Conceptions of Scholarly Publishing. In *Proc. CHI 98: Human Factors in Computing Systems*, Los Angeles, CA, 1998, ACM Press: NY, pp. 95-102. PrePrint: <<http://d3e.open.ac.uk/general/d3e-chi98/>>
 56. Toulmin, S. *The Uses of Argument*. Cambridge University Press: Cambridge, 1958
 57. Trigg, R. and Weiser, M. TEXTNET: A Network-Based Approach to Text Handling. *ACM Transactions on Office Information Systems*, 4, 1, 1983
 58. W3C Resource Description Framework (RDF). World Wide Web Consortium, 1997 <<http://www.w3.org/Metadata/RDF/Overview.html>>
 59. WebConf: *A Guide to Conferencing Software for the Web* <<http://thinkofit.com/webconf/>>
 60. *WebOnto Project*. Knowledge Media Institute, Open University, UK <<http://kmi.open.ac.uk/projects/webonto/>>
 61. Weinstein, P. and Alloway, G. Seed Ontologies: Growing Digital Libraries as Distributed, Intelligent Systems. In *Proceedings of the Second ACM Digital Library Conference*, Philadelphia, PA, USA (July), 1997, ACM Press
 62. Whitehead, J. and Wiggins, M. WEBDAV: IETF Standard for Collaborative Authoring on the Web. *IEEE Internet Computing*, September/October, 1998, pp. 34-40. PrePrint: <http://www.ics.uci.edu/~ejw/authoring/intro/webdav_intro.pdf>
-