

Applications of semantic web technology to support learning content development

Claus Pahl, Dublin City University, Dublin, Ireland

Edmond Holohan, Dublin City University, Dublin, Ireland

Article originally published
in "Interdisciplinary Journal
of E-Learning and Learning
Objects", V.5 (2009)
<http://ijlko.org/>
Reprinted with permission

ABSTRACT. The Semantic Web is based on ontology technology - a knowledge representation framework - at its core to make meaning explicit and more accessible to automatic processing. We discuss the potential of this technology for the development of content for learning technology systems. We survey seven application types demonstrating different forms of applications of ontologies and the Semantic Web in the development of learning technology systems. Ontology technologies can assist developers, instructors, and learners to organise, personalise, and publish learning content and to discover, generate, and compose learning content. A conceptual content development and deployment architecture allows us to distinguish and locate the different applications and to discuss and assess the potential of the underlying technologies.

KEYWORDS: *Content development and deployment, Knowledge representation, Learning content, Ontologies, Semantic Web*

Introduction

The World Wide Web is an important learning technology platform today. Its accessibility has made it a successful environment in particular for the publication of learning material. Learning resources can be provided in a standardised format that can be accessed at any time from any location. The Web, however, is still evolving. The current evolution of the Web can have an impact on educational technology. This will affect instructors and learners alike. The Semantic Web initiative aims to support explicit semantics and its automated processing (W3C, 2006a). Currently, search and retrieval functionality relies on human interaction and often ad-hoc approaches to selection of documents for a given set of search criteria. Semantic annotations, which can be processed by software applications, will improve the precision of searches. This will enable accurate searches for learning resources. The opportunities that will emerge for educational technology as a result of the Semantic

Web initiative, however, go beyond search and retrieval (Devedžić, 2004a). The overall development and deployment process of educational technology can be affected. Ontology technology - the knowledge representation and inference core of the Semantic Web - promises this wide applicability (Berners-Lee, Hendler, Lassila, 2001). An area such as education, where access to information is central, depends on the representation and organisation of knowledge both for the content but also the metadata level.

Our objective here is a technology assessment based on a survey of different application types, assessing the potential of Semantic Web technologies. We focus in particular on its impact on the development of learning content. We investigate whether ontology technology, if applied suitably, can simplify development for authors and instructors, improve access for learners, and also allow sharing and reuse for all actors involved. Tasks that look simple at first glance, like providing metadata and annotation, have proven to be difficult in practice. Ontology development is an even more challenging problem. We discuss to what extent ontology and Semantic Web technology can actually be utilised to support content-related development aspects of learning technology, including the creation of content, the publication and personalisation of content, the discovery of learning objects, and the generation and composition of complex learning objects. We consider in our investigation Web-based learning technology systems (LTS), which are typically based on software, multimedia, and hypermedia technology, to support learning and training activities. These LTSs facilitate the authoring and delivery of content. Content is represented through usually reusable learning objects that can be processed by LTS to provide learning and training activities for a specific topic.

Ontology technology has already been used in educational technology (Aroyo, Dicheva, Cristea, 2002; Fischer, 2001; IEEE, 2002; Leidig, 2001; Pahl, Holohan, 2004; Sampson, Lytras, Wagner, Diaz, 2004) with different purposes ranging from the definition of a domain-specific terminology to the use of conceptual models and inference in the generation and composition of learning technology content and systems. While some comprehensive accounts exist (Devedžić, 2006; Sampson et al., 2004), we give a systematic and comprehensive technology overview here, surveying a number of different application examples, specifically focusing on learning content development and related activities. We illustrate the benefits

for authors, instructors, and learners, but also the limitations. We take a developer perspective, broadly speaking focusing on content developers and instructional designer roles, to discuss the different aspects important for participants who are involved at an early stage in the development and deployment lifecycle of learning content and learning technology systems. In order to facilitate our investigation, we develop a framework that allows us to distinguish different application types, to characterise the implications for the different actors, and to identify the techniques used in each application type. Although our main investigation focuses on the application of the classification framework to learning content aspects - which we feel is the best understood application of Semantic Web technology to educational technology - we also address the wider context and limitations of the technology in our Discussion section at the end. The next section introduces ontology and Semantic Web technology. The rationale and organisation of our review is presented in the following section. This section identifies seven application categories - ontology development, creation and generation, adaptivity and presentation, packaging and interoperability, organisation and sequencing, metadata and annotation, exchange and sharing - which are discussed subsequently. A critical assessment follows the individual presentations before ending with some conclusions.

Ontology and semantic web technology

The Web creates a space in which content developers, instructors, and learners contribute to and participate in learning processes. Knowledge is a central component in this space. We introduce a knowledge space for learning technology in this section and describe how it can be structured through ontologies and other Semantic Web techniques. This provides us with a classification and comparison framework for the subsequent discussion of ontology-based learning content and its development.

Ontologies are knowledge representation frameworks that describe an area of knowledge by defining the common concepts of that domain and the concepts' properties and relationships (Daconta, Obrst, Smith, 2003; Gruber, 1993; Wilson, 2004). Gruber defines an ontology as a specification of a conceptualisation that is created with the aim of sharing knowledge by committing to this ontology (Gruber, 1993). The knowledge space - defined by Sowa as the

combination of knowledge types, representation formats, and purpose of represented knowledge (Sowa, 2000) - for *learning technology systems* (LTS) comprises several knowledge types relevant to the educational context. Several ontologies inhabit and organise the knowledge space in the educational context: subject content, instruction, user, and system ontologies (see Table 1). Since our focus is on content, essentially, the first three types are relevant; the fourth type is of more importance to the development and deployment of underlying infrastructure, such as the LTS, that allows content to be authored and delivered.

| TYPE | DESCRIPTION |
|-------------|---|
| Content | subject-specific knowledge describes the subject-related aspects of the content |
| Instruction | pedagogic knowledge describes the educational aspects of the content |
| User | user knowledge describes characteristics, preferences, and past learner behaviour |
| System | system-related knowledge describes the LTS implementation aspects of the content |

Table 1.
Educational knowledge types

Another aspect of the knowledge space is its *purpose*, i.e. which functions are supported; see Table 2 in which the functions are listed with increasing semantic level. Ontologies are knowledge representations, but ontologies are often seen as intertwined with logics (Daconta et al., 2003; Sowa, 2000). Ontologies provide the terminological aspects needed in logical reasoning. We can divide the four ontology functions into two different purposes. *Vocabulary* and *terminology* are supported by taxonomies and thesauruses, which mainly address annotation and retrieval needs. *Modelling* and *reasoning* are supported by conceptual models and logical theories, which address requirements arising in the development and composition of educational resources.

Table 2.
Knowledge space functions
and purpose

| FUNCTION | DESCRIPTION |
|------------------|---|
| Taxonomy | terminology definition and classification are the central issues - it supports browsing and retrieval of educational resources |
| Thesaurus | relationships between terms are the central issues - it constrains the use of a vocabulary |
| Conceptual Model | a formal model of a domain - it supports modelling of the subject area and technical aspects which often use more than classification-oriented relationship types |
| Logical theory | reasoning and inference are the central issues - it combines knowledge representation with a logic and, thus, supports reasoning within a knowledge domain |

The knowledge space that we have defined for the educational context is an abstract framework, which needs to be realised through concrete notations and techniques. The *Semantic Web* (Berners-Lee et al., 2001; W3C, 2006a) provides ontology notations and techniques, based on the ontology language OWL (the Web Ontology Language). OWL in turn is based on RDF (the Resource Description Framework) and XML (the eXtensible Markup Language) (see Table 3). Schema languages, such as XML Schema or RDF Schema, are the tools to introduce a vocabulary into an ontological framework. A logic underlying the ontology provides the reasoning facilities. Most applications we review here are based on this stack of Semantic Web techniques.

| TECHNIQUE (semantic level) | DESCRIPTION |
|-------------------------------|--|
| XML | <p>XML and the XML Schema language provide the basic syntactic interoperability for an ontology definition through an approach to define markup languages.</p> <p>Example: The XML expression <code><exercise> What is a <concept> geographical concept </concept> ? Name and define five examples. </exercise></code> marks up this instruction as an exercise with the key concept 'geographical concept' through XML tags.</p> |
| RDF | <p>RDF and RDF Schema introduce semantics. This allows the description of concepts in terms of triples - subject, property, and object. A new concept is defined in terms of its properties in relation to others. RDF Schema allows the definition of classes.</p> <p>Example: <code>Country subClassOf GeographicalConcept</code> is a (syntactically simplified) triple, where each element would have to be expanded to a URI in RDF.</p> |
| OWL | <p>OWL is an ontology language that extends RDF and that combined with a reasoning tool provides logical facilities for reasoning and inference.</p> <p>Example: Countries have a 'name'-property expressed by a string value:</p> <pre data-bbox="353 1030 898 1233"> <rdfs:Class rdf:ID="Country"> <rdfs:subClassOf> <owl:Restriction> <owl:onProperty rdf:resource="#name"/> <owl:allValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/> </owl:Restriction> </rdfs:subClassOf> </rdfs:Class> </pre> <p>The central reasoning concept is subsumption - the subclass relationship between concepts or properties of concepts.</p> |

Table 3.
Semantic web technology
stack

Technology review - rationale and organisation

In the subsequent sections, we discuss applications of ontology and Semantic Web technology for the development of content in learning technology systems - drawing on a literature review and our experience in the development and deployment of these systems. In this section, we develop a framework - which is a central tool in our discussion - that allows us to characterise and describe these application types in a systematic way.

Objectives and rationale

Our aim is to illustrate and assess the potential of ontology and Semantic Web technologies for the process of learning content development. For this purpose, we present a classification scheme for different application types of ontologies and Semantic Web technologies in the educational context. We have chosen applications that illustrate the identified set of application types in the most suitable way. This contribution is, however, not meant as a comprehensive literature review - the range of application types within the given focus, however, aims to be comprehensive. The aim is rather the discussion of this technology in the context of learning technology.

We present this discussion from the perspective of developers - which covers a range of roles from knowledge engineers, content developers, instruction designers, and software and media designers - all involved at an early stage of an LTS lifecycle. Our classification scheme is based on a development and deployment architecture, which reflects the developer perspective. This architecture is a conceptual architecture that supports the process of learning content development and deployment, thus presenting central aspects from a developer's perspective. Using a process-oriented architectural perspective to classify applications emerges as the most suitable choice. It still allows us to address other perspectives, e.g. those of the different actors involved. The alternative to our conceptual architecture would have been a more functional architecture based on the software components of a learning technology system, such as the IEEE Learning Technology System Architecture LTSA (IEEE, 2003). While the latter would be more geared towards a software developer, our conceptual architecture is more concerned with the content development process rather than LTS component development, thus being more suitable for content developers and instructional designers.

Selection and organisation

The selection of appropriate application types and corresponding concrete examples is based on the following two-step method:

- Firstly, a content-centric conceptual architecture is developed to gain a comprehensive framework in which a complete overview can be given. This architecture is based

on our own experience and a survey of standard literature. The architecture component identification follows general software architecture principles and a use case-based functionality analysis.

- Secondly, for each component or function in the architecture, a literature review has led to the selection of examples that illustrate the potential of the technology, i.e. is reasonably advanced, but also suitable to convey the central principles.

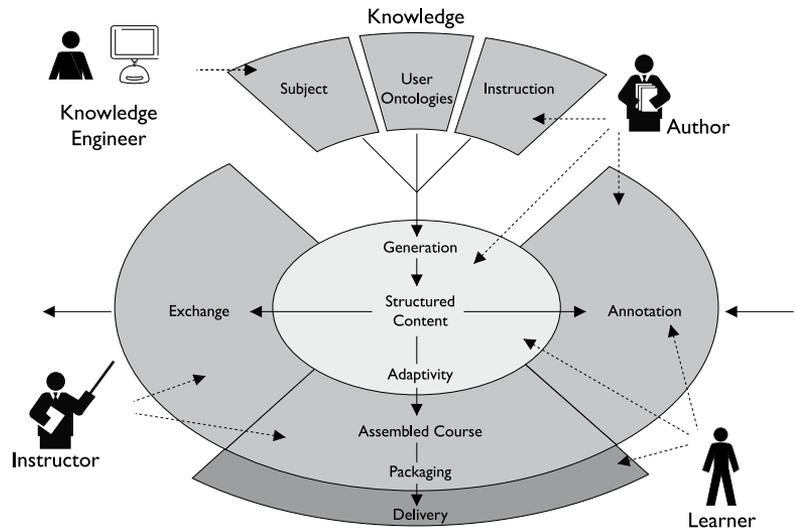
We have used the content and system development lifecycle as the motivation for the classification scheme. We have arranged these application types based on the development process for learning content.

Architecture and content development

The architecture of a learning technology system (LTS) is the description of its structural and abstract behavioural characteristics (including content and content processing activities). It provides, therefore, together with the stages of content development, an ideal basis to discuss the locations of different ontology and Semantic Web applications within a learning technology system.

We can distinguish two central elements of a learning technology system: the content and the learning and content management components. We take a content-centric perspective here - focussing on content development and how content is used by instructors and learners (see Figure 1). We will later on walk through the architecture when illustrating the different applications, starting with the knowledge and ontology part, then addressing generation and structured content, and so on. Learning objects are digitally represented learning content in a learning technology system (IEEE, 2002). Learning objects are rather small and are usually assembled to larger units of study.

Figure 1.
A content centric learning content
development and deployment
architecture



Development

Based on the architecture in Figure 1, we identify seven application types of ontologies and Semantic Web technologies. These application types shall be organised into four contexts, determined by the activity and the actor for whom the application type is most relevant. The development of content and learning technology systems is a participative effort, involving domain knowledge specialists, content authors, instructors, and learners. The perspectives of these four important actors involved shall be addressed in our discussion of knowledge technologies in the educational context.

- Domain knowledge engineers are needed to develop the ontologies for the educational context, such as content, instruction, learner, and LTS-related knowledge, resulting in the first application type: ontology development.
- The content author is the second actor in the overall development process. The focus is on the ontology-aware creation and organisation of individual content units: creation and generation of content and adaptive and personalised presentation of content.
- The development of individual content units is not

complete. These have to be composed, packaged, and assembled as courses by the instructor: packaging of units and interoperability of learning objects and organisation and sequencing of complex learning objects.

- Once content is available, possibly in the form of courses, the material can be used in different forms. Sharing and reuse are aspects relevant for all actors involved: annotation and discovery of learning objects using metadata and exchange and sharing of resources.

We have named the roles typically involved in these activities. The learner, for instance, could also contribute, as indicated in Table 4 and discussed in the remainder of this paper. These seven application types are summarised in Table 4. The “Development Issue” describes the main purpose of the activity, the “Knowledge Aspect” refers to the knowledge types, the “Development Aspect/Actor” categorises application types according to the content artefacts under consideration and the actor typically involved, and the “Semantic Function” refers to the knowledge space functions. The development and knowledge aspect and actors involved and the functions and levels of the knowledge space form a *classification and characterisation scheme* for these application scenarios. The table captures the predominant aspects and is not meant to be exclusive. We illustrate each scenario in a separate section by suitable examples in the remainder of this presentation.

Table 4.
A classification of different forms of application of ontology and Semantic Web technologies for learning technology systems

| Application | Development issue | Knowledge aspect | Development aspect/Actor | Semantic function |
|--------------------------------|---|--|---|---|
| Ontology development | Generate learning ontologies | Content, instruction, user, system | Ontologies (knowledge engineer) | Conc. model, log. theory, taxonomy, thesaurus |
| Creation and generation | Generate content from ontologies | Content | Individual units (author, learner) | Conc. model, log. theory |
| Adaptivity and presentation | Adaptive presentation | User | Individual units (author) | Taxonomy |
| Packaging and interoperability | Interoperability | Metadata (content and instruction) | Individual units (author) | Taxonomy |
| Organisation and sequencing | Educationally sound sequencing and assembly | Instruction | Assembled units (instructor, learner) | Taxonomy, logical theory |
| Metadata and annotation | Abstract description | Metadata (content, instruction and user) | Reusable units (author, learner and instructor) | Taxonomy, thesaurus |
| Exchange and sharing | Sharing and exchanging (technical) | Metadata (content and instruction) | Reusable units (instructor and learner) | Taxonomy, logical theory |

Ontology development

Ontologies, which form the backbone of a knowledge-driven learning content development and deployment approach, often have to be developed for a specific purpose.

The need to provide ontologies for the given educational context was recognised early (Bourdeau, Mizoguchi, 2002). Ontologies for the various aspects - we have classified the relevant knowledge into content, instruction, user, and system - need to be developed and made available, either through the Semantic Web community in general or specifically developed by experts for the educational application context.

Knowledge-based authoring of content and instruction needs to be an ontology-ware process to make explicit the relationship between knowledge and content, but also between knowledge and other components and actors. *Ontology engineering* methods

(Noy, McGuinness, 2001) can be used to develop these content, instruction, user, and system ontologies that can serve as meta-models and input for the authoring process. These methods provide procedures to identify concepts, instances, and relationships and to develop subsumption hierarchies and other richer semantic representations. While general-purpose ontologies, which can be found online, might suffice, for instance, as subject domain ontologies to classify learning objects according to their subject, instruction ontologies require education expertise as input and some activities, such as ontology-based content generation, require rich ontologies based on a variety of relationships. We address the latter aspect in the next section.

Two research activities are important in this context:

- The *richness of ontologies* for description and reasoning within the learning context has been investigated recently by Boyce (2004) and as part of the Diogene Project (Capuano, Gaeta, Micarelli, Sangineto, 2004). We will return to their work in the section on 'Content Organisation and Sequencing', where these ontologies are exploited to support the organisation of larger units of study.
- Another important activity in ontology development is the automated *extraction of knowledge* from existing resources (such as textbooks) in the form of ontologies. This needs techniques different from the classical ontology engineering approaches (Noy, McGuinness, 2001). The work by Buitelaar and Ramaka is an example of how natural language processing techniques are applied to automatically generate ontologies (Buitelaar, Ramaka, 2005).

In addition to methodological support, tools are critical for the success of this activity. The Ontology Editor is a tool, described in Bourdeau and Mizoguchi, that enables the collaborative development on ontologies specific to the educational context (Bourdeau, Mizoguchi, 2002). Ontologies are sharable knowledge representation formats, making collaborative ontology development between domain experts and knowledge engineers the ultimate objective. In recent years, stable open source ontology editors, such as Protégé, that are compatible with OWL have become available to support this endeavour and

have increased the chances of ontology technology to become accepted for learning technology.

Content creation and generation

Content authoring comprises the creation of educational content from scratch by an author or instructional designer and the generation of content from resources such as ontologies.

We start with a *generation-oriented* scenario, where stand-alone knowledge such as a subject ontology (see corresponding entry in Table 4) is directly used to generate a content outline (Capuano et al., 2004; Fischer, 2001; Pahl, Holohan, 2004). This is in contrast to knowledgebased organisation or annotation of content, which use ontologies on a meta level. Content does not exist prior to the generation process. Input is solely provided by the subject ontology - see Figure 2 where a simple ontology representing a concept hierarchy is converted into two forms of learning objects.

Figure 2.
Generation of learning content based on ontologies

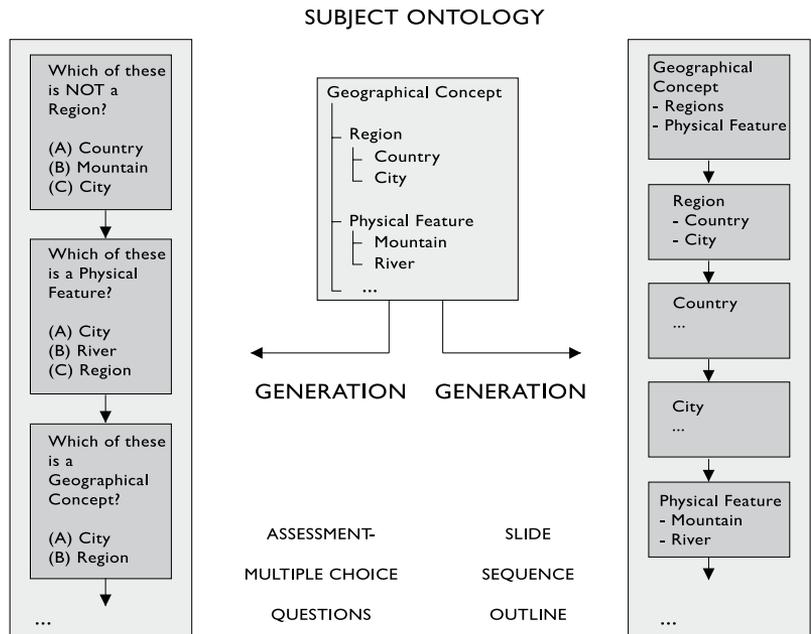


Figure 2 illustrates the application type using the OntAWare authoring tool (Melia, Holohan, McMullen, Pahl, 2005). Ontologies define a conceptual model for learning content components. The latter are assembled to larger content objects. These can be translated into Web representations or other, print-oriented media. Two examples are:

- *Content outlines.* The knowledge represented in the ontology can be converted into learning content. A concept hierarchy usually forms the backbone of such an ontology; this hierarchy also guides the sequentialisation of the concepts and their descriptions (Fischer, 2001), usually based on a depth-first traversal of the hierarchy tree.
- *Assessments.* Multiple choice questions (and answers) can be generated that can be used as input for an assessment tool based on multiple choice questions using related and unrelated concepts from the ontology. The distance between concepts in the hierarchy can be used to generate challenging answers covering closely related false answers (distractors).

The OntAWare aim is to fully automate the first generation step in content development, which can produce reasonably complete outlines if rich ontologies are available as input. A prioritisation of concept relationships (subconcept hierarchies, dependencies, etc.) determines concept ordering, if more than traditional classification-based hierarchical ontologies are used. This can result in more than simplistic outlines. Fischer and Boyce consider extended ontological models based on two knowledge spaces - the concept space, which represents concept-based hierarchies, and a content space, which captures more extensional knowledge in form of definitions or examples (Fischer, 2001; Boyce, 2004). In contrast to OntAWare, another more semi-automatic, interactive approach to ontology-based authoring of content is taken by the AIMS system. Concept-based courseware authoring (Aroyo et al., 2002) is the approach to provide the author with assistance in creating content through domain and instructional models in the form of ontologies and to configure these for delivery. A special feature of AIMS is the support of a generic set of authoring tasks within the system. Based on activities, such as add, delete, and edit, the AIMS system guides an author

through standard interactive authoring dialogues. It suggests solutions along the way and supports the creation of courses, topics, concepts, and tasks, and how they are connected - which has the benefit of allowing the author more influence than an automated solution.

The discussion of the systems and in particular the underlying ontology technology shows that ontology engineering is a discipline that needs to mature. Co-existing and changing ontologies will be the norm; mapping and evolution techniques need to be in place. The OntAWare experience shows that richer ontologies are needed in order to achieve better content quality - these richer ontologies would currently require manual definition.

Metadata and annotation

The abstract description of learning content through metadata is necessary to allow the publication and discovery of these resources. Annotation of fragments can help to link these to underlying knowledge, thus making the knowledge explicit.

Educational resources, ranging from simple text-based material to highly interactive systems, can be provided and accessed using Web technologies. In order to support the discovery of sharable Web resources by potential users, the resources need to be supplemented by suitable abstract descriptions. A prerequisite for this to work is a standardised and agreed upon vocabulary for these annotations. Ontology technology can provide in this context the knowledge support through taxonomy and thesaurus functionalities. The description of learning objects or fragments through annotation and metadata attributes is here the objective. Both metadata and annotation provide meta-level information, but at different levels of granularity.

Metadata

The Learning Object Metadata standard LOM (IEEE, 2002), although not ontology-based, provides a basic metadata framework for the faceted description and classification of learning objects. The learning object notion comprises a variety of educational technology applications. LOM defines the attributes required to fully describe a learning object. It classifies attributes into nine categories addressing, for example, general, technical, educational,

and lifecycle aspects. The provider of the learning object describes the object in terms of content and infrastructure properties. A potential user - learner or instructor - then uses a related query language (or a general-purpose Web search engine) to formulate requirements in terms of the abstract properties described. It relies on the provider to describe a learning resource adequately, and this task will be facilitated if the learning object in question is already explicitly linked to ontologies and this can be exploited. These attributes and their values form a vocabulary, which could be captured in terms of an ontology to enhance the description and discovery functionality.

Silva Munoz and Palazzo Moreira de Olivera (2004) propose using automatic metadata generation in their AdaptWeb system, based on domain and content knowledge ontologies, to support the storage of learning resources. The aim is personalised delivery based on the knowledge represented in both the learner model and the content metadata. The content knowledge ontology captures rules that allow the correct assembly of content units. This knowledge is linked to domain ontologies in the form of concept hierarchies. Complex learning objects that suit the profile of individual learners are then assembled based on the learning requirements. These are expressed in the learner model, which is also an ontology, in terms of domain concept and are guided by the meta-level rules to obtain educationally sound compositions. The learner's knowledge is expressed in terms of hasKnowledge links into the domain ontology. This allows creation of a learning trajectory based on knowledge about learner and content. This is an example where the learner can directly access the ontology to express learning objectives.

Annotation

Annotation is another form of attaching information to an existing resource. The Tangram system (Jovanović, Gašević, Devedžić, 2006) is a system that allows ontology-based, fine-granular annotation of learning content. The predominant form of representing learning content is a text document. These documents are often based on an inherent (but implicit) structure. In content documents we find definitions of new concepts, their illustration, examples, exercises, and so on. XML and ontologies provide ideal syntactical and semantical notations to make this

structure explicit through fine-granular annotation. This has a number of advantages:

- Firstly, an explicit structure supports the instructional designer in the learning design process. It gives guidance, allowing the designer to construct content from small individual building blocks.
- Secondly, an explicit structure makes the document accessible to others. The document can be searched for particular content items; for instance, a learner can search for exercises on a particular concept.
- Thirdly, the fine-granular organisation of documents into small, classified units allows the flexible storage and assembly of these units. Adaptive delivery is an example of this approach where personalised content can be assembled from these small units. This aspect shall be discussed in another section.

Jovanović et al. attempt to automate this annotation process using ontologies (Jovanović et al., 2006). Ontologies can play two roles here: firstly, to realise links from resource fragments into a knowledge space and, secondly, to determine the annotation itself by allowing text fragments to be classified based on concepts from an ontology occurring in the text. The second case is pursued in Tangram. The first case is a good example of direct learner involvement. Ontologies can be part of explanations of a subject domain or a knowledge representation for learners who work on exercises. For example, given a set of terms for concepts in the domain, the learner may have to construct a concept map with them, introducing relations as necessary.

Ontology technologies can make the inherent knowledge structure of content explicit. Taxonomy and thesaurus functionalities are used here to support the markup, in particular the definition of tags. Two ontologies describing two types of knowledge are important:

- Instructional knowledge is needed to give structure to educational documents. Education-specific markup languages, such as EML (Koper, 2001) and its successor, the IMS Learning Design LD (IMS, 2003), can form a notational system for content development and representation. These can provide

the primary structure. An ontology acts as a taxonomy, introducing a vocabulary through tags. This knowledge forms part of a development technique, which is essential for the instructor as an instructional design tool.

- Subject-specific knowledge can be used to support the educational structuring. It adds another dimension of access to the document in the form of subject-related query and retrieval functions. It is in particular suitable for learners searching for topic-specific units. Ideally, the subject-specific knowledge is based on a common, accepted ontology for the topic domain. The development of these ontologies has begun for various domains; examples are software engineering and genetics (Boyce, 2004). In addition to an introduction of a vocabulary with concept classifications (a taxonomy), domain ontologies that are used as subject ontologies often comprise thesaurus functionality as well, supporting, for instance, synonyms in searches.

There is, however, a limitation connected to markup as a structuring tool. Firstly, as already noted, sufficiently rich and accepted ontologies are only slowly emerging. Also, current standards like LOM are not sufficient as they do not cover all learning object aspects in order to support automation. Secondly and specific to the annotation context, the technique is essentially limited to textual resources. Multimedia, however, is an essential element of learning content.

Content adaptivity and presentation

Learning content needs to meet the expectations and requirements of the learners. Adapting content to individuals and groups of learners before the content is presented to the user is of major importance.

In general, separating content from its final presentation (or published appearance that is delivered to the end user) adds some flexibility that is widely used in the Web environment. Adapting content to the needs or preferences of the user requires the matching of learner knowledge with knowledge represented in content. The two traditional forms of adaptivity that have been most often implemented are layout and navigation adaptivity. Both forms of adaptive presentation of content are useful for

the educational context. The variety of forms and the degree of adaptivity of delivery can be enhanced through the use of ontology-based content organisation (De Bra, Aroyo, Cristea, 2004).

- In the simplest case, layout aspects can be separated from the content and added in an additional processing step. XML-based content documents separate structure of content from its presentation. The XML Stylesheet Transformation Language XSLT allows the transformation of XML-based input into a variety of output or publication formats. We can use this technique to create a richly laid out version for local use or a more reduced version for distance or mobile learning.
- Before creating different output formats in the translation, we could let the learner decide on the assembly of individual fragments or objects into larger learning objects. We could, for instance, generate a summary version that includes concept definitions but no exercises. In this advanced scenario, content itself, not only its presentation, can be tailored towards the needs of users. The standard technique here is to adapt the navigation between content units. XML is an abstract data and document-structuring format that allows machine processing, which enables flexible storage and retrieval of XML-documents. It allows content to be assembled from small units into learning objects that suit the needs of individuals or groups based on concepts and their dependencies organised in an ontology (Aroyo et al., 2002; De Bra, Brusilovsky, Houben, 1999).

Similar to creation and generation (see section on “Content creation and generation”), this application is based on a transformation step, but the purposes are different (content delivery vs. content creation and composition). While user and content knowledge is the central input, technology-related systems knowledge (see Table 1) can also help us to choose the most appropriate delivery technology.

The AIMS system supports adaptivity in an LTS (Aroyo et al., 2002). Central to this system are the knowledge bases. At the core is a domain ontology that captures the central concepts of the course and allows structuring the course along these concepts. A user model captures the learner profile, which provides the information

on which the adaptivity is based. A third pillar of this system is the instruction model, which captures suitable, educationally sound approaches for learning within a domain. The typical application of adaptivity is to adapt the navigation infrastructure between content units based on the learners' preferences and knowledge of the course subject. Concepts that have already been learned can be excluded from presentation. Depending on the concepts that a learner has already learned, the concepts that he or she intends to learn, and dependencies between concepts represented in the ontology, a personalised navigation path through the concepts to be learned can be generated. Overlay models, such as the one presented in Figure 3 to illustrate the principles, abstract and represent concepts dealt with in learning objects in a subject ontology, which in turn can be used as a reference point for user models to express the past learning activities of learners and the current state of their knowledge.

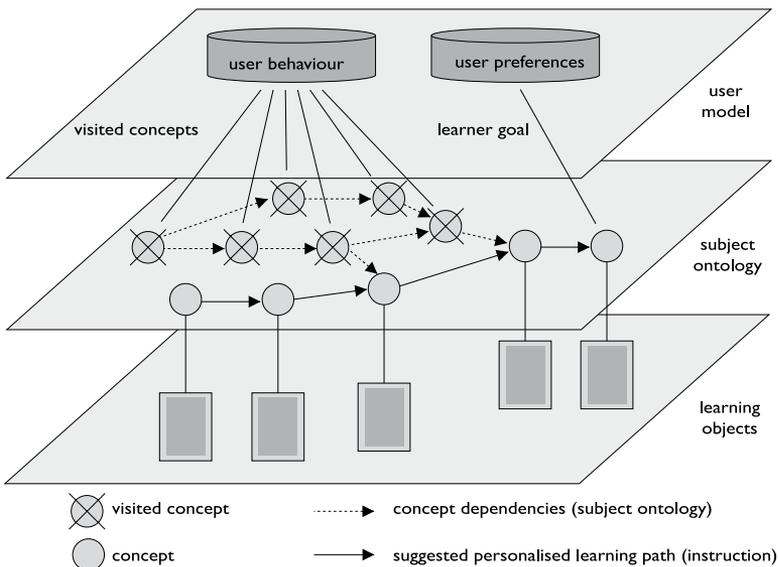


Figure 3.
Adaptive content navigation

The OntAWARE system (Melia et al., 2005) uses two criteria to determine the instructional component, i.e. the personalised navigation paths to the concepts the learner wants to learn: knowledge (determined through pre- and post-tests) and behaviour

(the concepts that the learner has visited). This is an example of an adaptive hypermedia system where the knowledge bases are needed beyond the content creation stage until the final delivery. Learners access and navigate through ontologies here directly. A possibility is to use ontologies to support learners in exploring their knowledge of subject domains - as in open learner modelling - and to access relevant content via ontology-based exploration and navigation to improve their knowledge.

Personal Reader (Dolog et al., 2003) and LAOS/MOT (Cristea, 2004) are other adaptive systems that also use Semantic Web technologies to support the adaptation of content. Personal Reader is based on an RDF/RDFS-based representation of knowledge. LAOS is an adaptive hypermedia framework of abstract models that has been formulated in terms of an XML Schema. MOT is an authoring system based on LAOS. These two systems are indicative examples of research prototypes that illustrate current activities in this active area of research. Systems like OntAWare, Personal Reader, and LOAS/MOT are research prototypes that exhibit a major shortcoming of these systems. These generic, subject-independent systems are not sufficiently tried out in practice. This has been achieved so far only by subject-specific systems such as the ActiveMath environment (Libbrecht et al., 2001). ActiveMath represents mathematics knowledge as conceptual content in an ontology. It captures structures, dependencies and pedagogical information at metadata level. This is used to automatically generate interactive content according to the learner's goals, competence and preferences. It is also one of the few systems that tackle interactivity.

Content organisation and sequencing

Individual content units need to be assembled into courses or other units of study. The organisation and sequencing of units in an educationally sound way is the central activity. Knowledge captured in the form of ontologies can support this task. Standards exist that allow the sharable implementation of these sequencing definitions through navigation infrastructures.

Knowledge about a collection of learning content units can be used to organise individual units into a larger learning object by sequencing the units based on inherent dependencies that are

derived from the knowledge. In addition to basic taxonomy and thesaurus functionalities, conceptual modelling and logical theory functions play an important role. While this is traditionally the task of the instructor, with increasing automation and reasoning and interface support, learners could equally well compose content units from repositories according to their individual needs.

Knowledge-based content organisation

Content units (either learning objects or fragments) can address different learning aspects. Some might define and explain concepts; others might provide examples of a concept. Knowledge about these units - derived from internal structuring through annotation or metadata attributes - can help an author, instructor, or learner in assembling educationally sound courses. For instance, ontology rules can be formulated that capture the idea of soundness and integrity of these assemblies (Henze et al., 2004).

```
FORALL D, E, example(D,E) <-
definition(D) AND example(E) AND
    EXISTS C1 (D[dc:subject->C1]) AND
    FORALL C2 (D[dc:subject->C2] -> E[dc:subject->C2]).
```

This rule links a content unit D defining a concept and an example E, requiring that E relates to the concept defined in D. The second line requires D to be a definition and E an example. Through the third line it is verified that D is about a concept - the term `dc:subject` refers to a metadata attribute. The last line requires that example E is indeed about the concept referred to in D. A reasoning tool - supporting the logical theory aspect of the ontology framework - would carry out these inferences and would establish the soundness of an assembly. While this is ontology-assisted organisation of content, the organisation can be automated to a higher degree through sequencing.

Ontology-based sequencing of content

In ontologies, the standard organisational form is a hierarchy, categorising concepts into classes and subclasses. In more elaborate ontologies a variety of relationships between concepts might be

represented. The knowledge represented can comprise subject and education-related aspects (Leidig, 2001; Rius et al., 2008):

- Subject-related knowledge is often based on a semantic concept network. A subject-related ontology can be richer than a vocabulary or concept hierarchy. Often it forms a conceptual model describing a full domain.
- Educational knowledge often involves relationships that express dependencies, e.g. *is-BasedOn*. This knowledge can also comprise a vocabulary to classify educational units, such as definition, example, or exercise.

In this composition-oriented scenario of content organisation, knowledge is explicit and separated from the representation of content in the form of learning objects (Garlatti, Iksal, 2003). Concepts, which are addressed within the content units, and their relationships can be used to organise and sequence these units. Different types of relationships have to be dealt with in the process of arranging the content units in a suitable sequence (Fischer, 2001). A possibility is to prioritise relationships in this process, for example, to consider the subconcept relationship as the most important one. This sequencing algorithm uses the ontological reasoning facilities of an underlying logic to determine the ordering dependencies.

As an example, we consider an enhanced database ontology, which can support learning object generation or composition (Boyce, 2004). Central concepts are defined: *relation*, *record*, *table*, *database object*, etc. Concepts are related through a central subconcept or *is_a* relationship:

| | | |
|----------|-------------|-----------------|
| Table | <i>is_a</i> | database object |
| relation | <i>is_a</i> | database object |
| record | <i>is_a</i> | database object |

Concepts are in addition to *is_a* related through *isPartOf* and *is Basis For* relationships:

| | | |
|--------|------------------|----------|
| table | <i>isBasedOn</i> | relation |
| record | <i>isPartOf</i> | table |

The subconcept hierarchy is the backbone of the sequentialisation but cannot, as we can see here, resolve all dependency problems. Using a prioritisation approach - here `isBasedOn` as the secondary and `isPartOf` as the tertiary relationship - we could obtain the following order of concepts:

database object, relation, record, table

The constraints expressed in the ontology, however, might not lead to a unique solution. This example shows that ontological modelling for educational domains might require a rich set of relationship constructs. Boyce (2004) has used relationship types such as `isBasedOn` and `isPartOf` in addition to the standard subconcept relationship `is_a`. The Diogene project (Capuano et al., 2004) came to the same conclusion, using `HasPart`, `Requires`, and `SuggestedOrder` as necessary additional relationship types in rich ontologies.

This exploration demonstrates a major problem. Current approaches rely on formal notations, which would require a corresponding background of the content developer. It also shows that full automation is currently an ambitious aim and a highly interactive development process is more likely. In terms of the underlying ontology languages, there is only a slowly emerging trend towards a common set of relationship types for content and instruction modelling.

Sequencing and navigation implementation

Sequenced content units are usually connected through navigation links when they are presented to the learner.

- The SCORM Sequencing and Navigation standard SN (Advanced Distributed Learning, 2004) defines a notation for representing the intended learning behaviour in a consistent way.
- IMS Learning Design (LD) (IMS, 2003) is another standard addressing the development of composite learning activities.

SCORM SN, for example, defines how content can be sequenced through a set of learner-initiated or system-initiated navigation events. This complements the more abstract ontology-based sequencing we just described with implementation aspects. The branching and flow of content units is described by a predefined set of activities, typically defined at design time. SCORM SN also defines how a SCORM conformant learning management system interprets the sequencing rules expressed by a content developer with the navigation events and their effects on the run-time environment. IMS Learning Design (LD) (IMS, 2003) is another standard addressing the development of composite learning activities. IMS LD is widely supported by authoring and runtime environments, such as the Reload authoring tool. SCORM SN models the branching and flow of learning activities in terms of an activity tree, based on the results of a learner's interactions with learning objects and an authored sequencing strategy. An activity tree is a conceptual structure of learning activities managed by the learning management system for each learner. In SCORM, a learning activity may reference content objects that are delivered to the learner. SCORM SN describes how navigation events can be triggered and processed, resulting in the identification of learning activities for delivery. Each learning activity identified for delivery will have an associated content object.

Content packaging and interoperability

Interoperability and reuse of educational resources is only possible if these content resources are provided (i.e. packaged) in a widely accepted and supported format.

The IMS Content Packaging (CP) standard aims at providing interoperability of Internet-based learning content with content creation tools, learning management systems, and run-time environments. This standard is supported by an XML-based format as part of the SCORM Content Aggregation Model CAM (Advanced Distributed Learning, 2004). A content package consists of two files:

- An XML file describes the content organisation and resources. The organisation part describes the different views or organisational paths through the content. These

organisations are defined in the form of hierarchies.

- The resources themselves are described in XML format. Different resource types are supported. For instance, one resource type is “webcontent”, which encompasses HTML and other media that can be handled by a Web browser including standard plug-ins.

These two are joined together to form a transportable interchange file. This acts as a reusable unit in the form of a logical directory of resources and their abstract description. This resembles metadata, but instead of discovery and retrieval, the aim here is to provide interoperability descriptions. XML technologies provide an education-specific notation to assemble content units (resources) and organise them through hierarchical, educationally sound content packages. While this is similar to content composition, the focus is on interoperability, not primarily on learning design aspects of composition and assembly.

Current standards such as the SCORM suite are often lowest common denominators, making the implementation of ontology-based adaptive and interactive content difficult to achieve.

Content exchange and sharing

Sharing and exchanging refers here to a technical context, where content objects remain under control of the creator or owner, in contrast to annotation, packaging, and assembly, where resources are assumed to be fully (physically) available to potential users.

Ideally, the reuse and sharing of educational resources across organisations should reduce costs and improve quality. Annotation and metadata are the first steps towards reuse; they allow resources to be described by providers and discovered by potential users. A common problem with this scenario is that many institutions are reluctant to give up control over their learning resources. So-called peer-to-peer architectures provide a solution here. These enable institutions to participate in a sharing network without losing control over their resources. A suitable implementation can make the architecture invisible for the user. We contrast this here with a less flexible centralised repository architecture solution, which on the other hand

provides a more advanced description and retrieval approach. The Edutella project (Nejdl et al., 2002) suggests an RDF-based notation to describe sharable learning resources in a peer-to-peer architecture. RDF is suitable since bindings to most educational standards (LOM, SCORM, etc) are available. It provides enhanced taxonomy and logical theory functions. An RDF-based query language plays a central role for the discovery of reusable and sharable resources. Edutella shows how annotation in a heterogeneous environment works. At the core is a knowledge base that captures abstract descriptions of learning resources; here are examples for three text books:

```
<lib:Book about="http://www.xyz.com/se.html">
  <dc:title>Software Engineering</dc:title>
</lib:Book>
<lib:Book about="http://www.xyz.com/ai.html">
  <dc:title>Artificial Intelligence</dc:title></
lib:Book>
</lib:AIBook><lib:Book about="http://www.xyz.com/ai.html">
  <dc:title>Prolog</dc:title>
</lib:AIBook>
```

The execution of query *Return all resources that are a book having the title "Artificial Intelligence" or that are an AI book* in Edutella - here formulated in the logical notation Datalog -

```
aibook(X) :-title(X, 'Artificial Intelligence'),
type(X, Book).
aibook(X) :- type(X, AI-Book).
?- aibook(X).
```

would return a pointer to suitable resources retrieved from the knowledge base:

```
<lib:Book about="http://www.xyz.com/ai.html">
  <dc:title>Artificial Intelligence</dc:title>
</lib:AIBook>
<lib:Book about="http://www.xyz.com/ai.html">
  <dc:title>Prolog</dc:title>
</lib:AIBook>
```

Edutella provides a query language implementation for a distributed and possibly heterogeneous architecture. The Edutella aim is to provide an infrastructure that makes this nature of the underlying network transparent to the user. The execution of the query is an example where the reasoning aspects of the knowledge space - the logical theory - comes into play.

Sharing often involves, in contrast to the previous peer-to-peer architecture, dedicated and centralised learning object repositories. The use of ontologies for the description and retrieval of learning objects is here a natural idea. Tan and Goh introduce a suite of tools that support these activities (Tan, Goh, 2004). Their approach is based on general-purpose ontologies, such as SUMO and Cyc, for the description and classification of learning objects. Domain-specific ontologies can be used to refine the search for learning objects. While the architectural setting is different, and to some extent less flexible than Edutella, the underlying knowledge-based reasoning to support the query execution and result determination, as in the RDF example above, is in principle the same. However, Tan and Goh have, in comparison to Edutella, added layered general-purpose and domainspecific ontologies that increase the accuracy of searches through a more faceted classification of educational resources.

Problems in this area result from the need to generate adequate responses, which need to be user-specific and context-informed (Wang, 2008). On a different level, rights-related problems and the often-found unwillingness of content developers to make their content available is a further obstacle.

Discussion

Technology benefits

The Semantic Web is essentially a stack of techniques with XML providing syntactic interoperability at the bottom and knowledge representation and reasoning in the form of OWL ontologies at the top. Gruber (1993) emphasises two central aspects of ontologies - the representational perspective resulting in the defining and specification of a conceptualisation and the communicative perspective addressing the aim of a group of agents committing to an ontology, thus sharing the knowledge captured in this ontology. Ontologies can

benefit content development activities in these two ways.

- *Representing knowledge.* Ontologies make knowledge, on which learning content is based, explicit - supporting the development of content. This relates to the following application types: ontology development, creation and generation, adaptivity and presentation, packaging and interoperability, and organisation and sequencing.
- *Communicating knowledge.* Knowledge can be communicated and shared - supporting the reuse of content. This relates to the following application types: ontology development, metadata and annotation, and exchange and sharing.

These benefits could result in more flexibility in the creation and management of content and, as a consequence, in reduced costs. In this section, we summarise and discuss the current and future potential, but also the limitations of Semantic Web technologies for learning content support. We also broaden our view here and look at learning technology trends beyond content in both pedagogical and infrastructural directions.

We can conclude from our investigation that Semantic Web technologies excel in learning technology applications where the full potential of ontologies is used (see Table 4): creation and generation of content, organisation and sequencing of content, and metadata for and reuse of content. The most convincing applications essentially exploit the full semantic richness of ontology-based knowledge representation and reasoning. This, however, does not imply that the lower Semantic Web stack layers are without benefit - hence, we included some applications of this type in our discussion. XML technologies are enablers providing the required syntactic interoperability, which is often a major step ahead and additionally an enabler for further improvements. The higher layers add flexibility, quality improvement, and cost reduction.

Technology limitations

The technical potential, which we have illustrated through a number of research prototypes, is often limited in actually making an impact in practice by a number of factors, including human factors and technical factors such as tool support and availability of ontologies. We indicate in this overview whether it affects

the representation or communication of knowledge and content dimensions.

- Human factors relate to the attitude and the abilities of the content developers and instructional designers, clearly visible in relation to the representation of knowledge. Acceptance of the Semantic Web technologies is still a central limitation - even the necessity of providing XML-based metadata is not generally recognised. Even if this importance were recognized, it is a challenging and costly task that requires skills and expertise. A certain degree of familiarity with knowledge engineering technology and ability to use these is always required. Content modelling in an ontological framework is the minimal skills requirement.
- Technical factors that hinder the full exploitation of the potential of ontologies include the limited tool support - which affects both the representational and communicational dimensions of ontology support. Generic tools for ontology processing such as the Jena engine are available; their adaptation towards learning contexts is, however, still in its infancy. A current difficulty is the support of advanced semantic functions (semantic models, logical theories) in the form of tools accessible to the non-expert based on the higher-level semantic technologies RDF and OWL. Protégé is an ontology editor that, for example, supports OWL as the ontology format. Other interesting tools include the Reload authoring framework, which, although not ontology-based, connects to learning design and composition issues we have discussed earlier on.
- The application types we have discussed often assume the existence of ontologies for particular domains - the general expectation is that these will be developed once the Semantic Web becomes a reality. However, often ontologies that can be shared and communicated (or used to share and communicate content) do not exist and will have to be created or extracted from existing resources, or they do not satisfy the needs of ontologies for the educational context, which might encompass, for instance, richer structures than just classification hierarchies.

Areas that indicate current limitations or additional potential, such as adaptivity, are discussed in the next subsection. We also discuss the learner perspective and formats of learning that require more support through ontology technology.

Technology - trends and current developments

We have discussed application types in this investigation that have been realised and whose benefits have been demonstrated. Semantic Web technologies should, however, also provide input for some currently unexplored aspects of learning technology.

- Adaptivity is a broader technique in which semantic knowledge could potentially support more than the classical content and navigation adaptivity that we have looked at here. Ontology-linked user modelling would allow user characteristics such as background to be exploited, for example, by using metaphors or examples that match the background of the learner.
- Reasoning about the soundness of content compositions (sequencing) shall also be mentioned here, as it employs the most advanced form of reasoning of all the application types presented. Ideally, this support can be automated in the future.

Considering that the Web has become one of the predominant platform for e-learning, all its aspects should be addressed. We have focused on Semantic Web activities. The Semantic Web is only one of the directions in which the World-Wide Web is developing. While the Semantic Web ultimately focuses on the human end user by providing automation through explicit, machine-processable knowledge, the Web Services architecture (W3C, 2006b) aims to enable software-to-software uses of the Web. The principle of software services, provided at certain locations on the Web that can be used by other software applications, is the basis. Service-oriented learning technology systems are the application of this principle in the educational context (Devedžić, 2004b). The consequence is a blurring of the distinction between content and management systems. Content will be provided through services, as will standard functions of a learning technology system such as user management or evaluation support. Technically, this is a simplification. From the perspective of instructors and learners,

seeing content as dynamic objects is required. One of the consequences is that the composition of content (sequencing - the educational perspective) and composition of functions (software assembly - the technical perspective) will become the same. Learning- and training task-oriented composition is the objective. Ontologies and Semantic Web technologies can also play a central role here to support composition.

Learning approaches and the learner

We have looked at learning content within the traditional paradigm of instructor- or technology mediated learning as knowledge transfer. The broader educational context also currently involves paradigm changes towards more active learning approaches. Our focus in this investigation has been on content support within the current predominant paradigm in order to provide a comprehensive account. Semantic Web technologies, however, have the potential to support a broader range of educational aspects.

- We have already mentioned collaboration between learners as an integral part of a successful learning experience. While the relationship between knowledge and content is reasonably clear, the area of collaboration between learners (a central element of learning) and its support by ontologies is a less researched context.
- Our investigation has also focused on knowledge transfer as the paradigm of learner-content interaction. Active learning and skills training approaches are educational paradigms that might equally well benefit from, for instance, ontologically represented procedural knowledge. These are, however, still in their infancy and stable theories and methodologies in the context of technology-supported learning and training and best practice expertise are required before Semantic Web applications can be investigated.
- Problem-based and experimental learning, which adds a realistic setting and proximity to real-world situations, can also benefit from knowledge bases to support their activities.

We have focussed on a developer perspective, addressing roles usually associated with the early stages of the LTS and content lifecycle. With improved tool support and increased levels of automation, the role of learners can become more active and allow for stronger participation in the development activities we have discussed. The learner-centric composition of learning objects is clearly an option for learners that will see increasing demand, but the creation and annotation of content is also becoming a realistic scenario. This scenario, however, needs a deeper investigation than possible here.

Conclusions

Knowledge is of major importance for the development and deployment of learning technology. Content, learning objects, and learning technology system components are different notions of parts of a learning technology system. Knowledge is central to their structure, metadata, presentation, creation, and composition. Different types of knowledge - content, instruction, and learner-related and meta-level knowledge about content and technical aspects - can be captured in ontologies. Knowledge in general and ontologies in particular can support learning technology systems in various ways. If ontologies are made widely available and are supported through Semantic Web technologies, then various perspectives on content in the development and deployment process can be supported:

- *Individual content units.* Explicit knowledge in form of ontologies can be used to create and structure content and to make these content units adaptive.
- *Content assemblies.* Subject knowledge within content units, but also external, instructional knowledge can be used to organise content into educationally sound units of study and to package the content units in order to obtain interoperable content objects.
- *Reusable content.* Meta-level knowledge is required to allow content for reuse to be discovered, reused, and shared within different environments by groups of learners.

The process- and architecture-based classification framework

that we have developed has helped us in structuring and exploring the different applications of ontology technology. It has, as an instrument for this specific purpose, successfully facilitated the classification and structuring of existing work in the area into a consistent development-oriented framework. It allowed us to categorise all projects and systems we have encountered in our investigation, and it has also demonstrated its adequacy through the fact that each application type is supported by widely discussed or accepted techniques and systems. Since the framework links development stage-related application types to specific ontology techniques, it provides also a support tool for technology selection for ontology-aware content development. Overall, the Semantic Web technologies allow further evolution and gradual improvements of traditional, non-Web based approaches (Diessel, et al., 1994; Van Marcke, 1990), but will not revolutionise learning technology. The main benefit of ontology technology is the support it can provide for an instructor or instructional designer. While this approach does not directly impact the quality of content or the learning experience, improving content creation, configuration, and management support and improving the reusability and exchange of tested resources will impact the quality positively and can be a contributor to cost reduction. Ultimately, the learner will benefit from these through extended availability and access to content, lower education costs, and possibly better quality.

We have seen in the application types we discussed that the full spectrum of ontology technology from simple taxonomies to logical reasoning can be beneficial to educational technology. It supports a variety of tasks of both learners and instructors ranging from a basic (e.g. creation) to an advanced level (e.g. adaptivity). A number of aspects have emerged from our discussion of these applications:

- The relationship between knowledge and content becomes explicit when ontologies are used to create, organise, and describe learning content. Explicit knowledge provides an easy and flexible access to content and structures creation and presentation. Knowledge management is a central aim in various areas and environments for individuals and organisations and it is also relevant for educational

technology.

- An important theme connected to the use of ontologies is sharing and reuse. Ontologies are enablers of the reuse of learning resources in different contexts and environments and their use by a community of learners. Ontologies make educational resources more accessible through a standardised and accepted representation of knowledge that can be communicated among the participants involved.

Ontology and the Semantic Web provide a vision and technologies to realise this vision. Nonetheless, the ongoing process of research and development causes problems in relation to standards and tool support. Core elements such as OWL are standardised and adequately supported by editors and processing engines. Standards in the learning technology area such as SCORM, however, support ontology technology only to a very limited extent. Another problem is the current lack of domain ontologies for individual subjects. While markup and metadata are well understood, the exploration of the reasoning capabilities of ontology technology for learning technology has only begun. This leaves ontology-based reasoning to support activities, such as composition and the extraction and generation of ontologies that satisfy the educational needs, on the agenda for researchers. For developers aiming to embark on the development of knowledge-aware learning content and learning technology systems, this means that description and query support are feasible to date or in the near future. In conclusion, the potential of ontology and Semantic Web technologies for learning technology has only been exploited successfully to some extent, leaving some of the promises and objectives still to be accomplished.

References

Advanced Distributed Learning ADL (2004), *SCORM Sharable Content Object Reference Model*, <http://www.adlnet.gov/Scorm/index.aspx>

Aroyo Lora, Dicheva Darina, Cristea Alexandra (2002), *Ontological support for web courseware authoring*. Proceedings of the 6th International Conference on Intelligent Tutoring Systems ITS'2002, France and San Sebastian, Spain, 2-7 June 2002, Berlin, Springer-Verlag, pp. 270-280

Berners-Lee Tim, Hendler James, Lassila Ora (2001), *The semantic web*, "Scientific American", V. 284, n. 5, pp. 34-43

Bourdeau Jaciqueline, Mizoguchi Riichiro (2002), *Collaborative ontological engineering of instructional design knowledge for an ITS authoring environment*. Proceedings of the 6th International Conference on Intelligent Tutoring Systems ITS'2002, France and San Sebastian, Spain, 2-7 June 2002, Berlin, Springer-Verlag, pp. 399-409

Boyce Sinéad (2004), *Investigation of ontology relationships used for the modelling and generation of educational content*, M.Sc. Dissertation, School of Computing, Dublin City University

Buitelaar Paul, Ramaka Srikanth (2005), *Unsupervised ontology-based semantic tagging for knowledge markup*. Proceedings of the Workshop on Learning in Web Search at the International Conference on Machine Learning, 7th August 2005, Bonn, Germany

Capuano Nicola, Gaeta Matteo, Micarelli Alessandro, Sangineto Enver (2004), *Diogene: A semantic web-based automatic brokering system*, "AIS SIGSEMIS Bulletin", V. 1, n. 3, pp. 65-67

Cristea Alexandra I. (2004), *What can the semantic web do for adaptive educational hypermedia?* "Educational Technology & Society", V. 7, n. 4, pp. 40-58

Daconta Michael, Obrst, Leo, Smith, Kevin (2003), *The semantic web - a guide to the future of XML, web services, and knowledge management*, Indianapolis, USA, Wiley & Sons

All URLs checked
June 2009

De Bra Paul, Aroyo Lora, Cristea Alexandra (2004), *Adaptive web-based educational hypermedia*, in Levene Mark, Poulouvassilis Alexandra (Eds.), *Web dynamics, adaptive to change in content, size, topology and use*, Berlin, Springer, pp. 387-410

De Bra Paul, Brusilovsky Peter, Houben Geert-Jan (1999), *Adaptive hypermedia: From systems to frameworks*, "ACM Computing Surveys", V. 31, n. 4

Devedžić Vladan (2006), *Semantic web and education*, Berlin, Springer-Verlag

Devedžić Vladan (2004a), *Education and the semantic web*, "International Journal of Artificial Intelligence in Education IJAIED", V. 14, pp. 165-191

Devedžić Vladan (2004b), *Web intelligence and artificial intelligence in education*, "Educational Technology & Society", V. 7, n. 4, pp. 29-39

Diessel Thomas, Lehmann Axel, Vassileva Julita (1994), *Individualised course generation: A marriage between CAL and ICAL*, "Computers & Education", V. 22, n. 1-2, pp. 57-64

Dolog Peter, Gavrioloaie Rita, Nejd Wolfgang, Brase Jan (2003), *Integrating adaptive hypermedia techniques and open RDF-based environments*. Proceedings of the World-Wide Web Conference WWW'2003, 20-24 May 2003, Budapest, Hungary, pp. 88-98

Fischer Stephan (2001), *Course and exercise sequencing using metadata in adaptive hypermedia learning systems*, "Journal on Educational Resources in Computing JERIC", V. 1, n. 1

Garlatti Serge, Iksal Sébastien (2003), *Declarative specifications for adaptive hypermedia based on a semantic web approach*. Proceedings of the 9th International Conference User Modeling 2003, 22-26 June 2003, Johnstown, PA, USA, pp. 81-85

Gruber Thomas R. (1993), *A translation approach to portable ontologies*,

"Knowledge Acquisition", V. 5, n. 2, pp. 199-220

Henze Nicola, Dolog Peter, Nejd Wolfgang (2004), *Reasoning and ontologies for personalised e-learning in the semantic web*, "Educational Technology & Society", V. 7, n. 4, pp. 82-97

IEEE Learning Technology Standards Committee (LTSC) (2003), *IEEE P1484.1/D11 Draft standard for learning technology systems architecture (LTSA)*, IEEE Computer Society

IEEE Learning Technology Standards Committee (LTSC) (2002), *IEEE P1484.12/D4.0 draft standard for learning object metadata (LOM)*, IEEE Computer Society

IMS Global Learning Consortium (2003), *IMS learning design (LD) standard*, <http://www.imsglobal.org/learningdesign/>

Jovanović Jelena, Gašević Dragan, Devedžić Vladan (2006), *Automating semantic annotation to enable learning content adaptation*. Proceedings of the 4th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems AH 2006, 21-23 June 2006, Dublin, Ireland

Koper Rob (2001), *Modeling units of study from a pedagogical perspective: the pedagogical meta-model behind EML*, Open University of the Netherlands, <http://dspace.ou.nl/handle/1820/36>

Leidig Torsten (2001), *L3 - Towards an open learning environment*, "Journal on Educational Resources in Computing", V. 1, n. 1

Libbrecht Paul, Melis Erica, Ullrich Carsten (2001), *The ActiveMath learning environment: System description*. Proceedings of the Calculemus Workshop at IJCAR, 21-22 June 2001, Siena, Italy, pp. 173-177

Melia Mark, Holohan Edmond, McMullen Declan, Pahl Claus (2005), *Ontology-based adaptive content navigation*. Proceedings International Conference on Methods and Technologies for Learning ICMTL 2005, 11 March 2005, Palazzo Steri, Palermo, pp. 435-440

Nejdl Wolfgang, Wolf Boris, Qu Changtao, Decker Stefan, Sintek Michael, Naeve Ambjörn, Nilsson Mikael, Palmér Matthias, Risch Tore (2002), *EDUTELLA: A P2PNetworking infrastructure based on RDF*. Proceedings of the World-Wide Web Conference WWW2002, 7-11 May 2002, Honolulu, Hawaii, USA, pp. 604-615

Noy Natalya, McGuinness Deborah (2001), *Ontology development 101: A guide to creating your first ontology*, Technical Report KSL-01-05, Knowledge Systems Laboratory, Stanford University

Pahl Claus, Holohan Edmond (2004), *Ontology technology for the development and deployment of learning technology systems*. Proceedings of the International Conference on Educational Hyper- and Multimedia Edmedia'04, 21-26 June, Lugano, Switzerland, pp. 2077-2084

Rius Àngels, Sicilia Miguel-Angel, García-Barriocanal Elena García (2008), *An ontology to automate learning scenarios? An approach to its knowledge domain*, "Interdisciplinary Journal of E-Learning and Learning Objects", V. 4, pp. 151-165, <http://ijello.org/Volume4/IJELLOv4p151-165Rius.pdf>

Sampson Demetrios G., Lytras Miltiadis D., Wagner Gerd, Diaz Paloma (Eds) (2004), *Special issue on ontologies and the semantic web for e-learning*, "Educational Technology & Society", V. 7, n. 4

Silva Munoz Lydia, Palazzo Moreira de Oliveira José (2004), *Adaptive web-based courseware development using metadata standards and ontologies*. Proceedings of the International Conference on Advanced Information Systems Engineering CAiSE 2004, 7-11 June 2004, Riga, Latvia, pp. 414-428

Sowa John F. (2000), *Knowledge representation – Logical, philosophical, and computational foundations*, Pacific Grove, CA, USA, Brooks/Cole Publishing

Tan Marvin, Goh Angela (2004), *The use of ontologies in web-based learning*. Proceedings of the Workshop on Applications of Semantic Web Technologies for E-Learning SW-EL'2004, 23-26 August 2004, Eindhoven,

The Netherlands, pp. 75-80

Van Marcke Kris (1990), *A generic tutoring environment*. Proceedings of the 9th European Conference on Artificial Intelligence (ECAI), 6-10 August 1990, Stockholm, Sweden, pp. 655-660

W3C - the World Wide Web Consortium (2006a), *The semantic web initiative*, <http://www.w3.org/2001/sw>

W3C - the World Wide Web Consortium (2006b), *The web services architecture*, <http://www.w3.org/TR/ws-arch>

Wang Shouhong (2008), *Ontology of learning objects repository for pedagogical knowledge sharing*, "Interdisciplinary Journal of E-Learning and Learning Objects", V. 4, pp. 1-12, <http://ijello.org/Volume4/IJKLOv4p001-012Wang200.pdf>

Wilson Ruth (2004), *The role of ontologies in teaching and learning*, JISC Joint Information Systems Committee - Technology and Standards Watch Report TSW0402, <http://www.jisc.ac.uk/media/documents/techwatch/acf1lac.pdf>

Sintesi

Qual è il potenziale delle tecnologie del web semantico e delle ontologie per lo sviluppo del contenuto nei sistemi tecnologici di apprendimento (LTS) dal punto di vista degli sviluppatori?

Quale l'impatto reale di tali tecnologie, che promettono di assicurare un supporto valido alla semantica esplicita e al suo trattamento automatico, sull'area educativa, in cui sono evidentemente centrali l'accesso all'informazione, la rappresentazione e l'organizzazione del contenuto?

Alle due domande si tenta di rispondere attraverso una rassegna sistematica della letteratura scientifica del settore e una puntuale disamina di sette tipi di applicazioni, focalizzati appunto sullo sviluppo del contenuto di apprendimento; per scoprire in particolare se le tecnologie delle ontologie sono in grado di semplificare lo sviluppo del contenuto per gli autori e i formatori/instructors, migliorare l'accesso per gli allievi e consentire la condivisione e il riuso.

Il primo passo è la messa a punto di un modello di classificazione e confronto per il contenuto di apprendimento fondato sulle ontologie e il suo sviluppo.

Lo spazio di conoscenza, inteso come “la combinazione di tipi di conoscenza, formati di rappresentazione e scopo della conoscenza rappresentata” secondo la definizione di Sowa (2000), all’interno dei LTS comprende innanzitutto quattro tipi di conoscenza rilevanti per il contesto educativo: tre concernenti specificamente il contenuto - contenuto, formazione/instruction, utente - e uno relativo al vero e proprio sistema tecnologico di apprendimento, cruciale quindi per uno sviluppo e un utilizzo efficaci della sottostante infrastruttura.

Dato che sono le ontologie a fornire gli aspetti terminologici indispensabili in un ragionamento logico, esiste una strettissima relazione fra ontologie e logica, che si rispecchia fedelmente nelle quattro funzioni essenziali delle ontologie. In ordine crescente di livello semantico, tassonomie e tesauri supportano il vocabolario e le ontologie, modelli concettuali e teorie logiche supportano la modellazione e il ragionamento.

Lo spazio di conoscenza è destinato però a rimanere uno schema astratto se non ricorre a notazioni e a tecniche concrete necessarie alla sua pratica realizzazione, a un livello semantico di crescente complessità: XML, che garantisce l’interoperabilità sintattica elementare; RDF, che introduce la semantica; da ultimo OWL, che permette il ragionamento e la deduzione.

Infine, l’architettura concettuale impiegata nella selezione e nella classificazione dei tipi di applicazioni è coerentemente orientata al processo di sviluppo del contenuto, piuttosto che allo sviluppo della componente tecnologica (LTS). In essa vi sono due elementi chiave: da un parte il contenuto e dall’altra le componenti di apprendimento e di gestione del contenuto.

Sono quindi esaminati nel dettaglio i sette tipi di applicazioni di tecnologie di ontologie e di web semantico - sviluppo dell’ontologia, creazione e generazione, adattività e presentazione, packaging e interoperabilità, organizzazione e sequencing, metadati e annotazione, scambio e condivisione - ciascuno con le sue specifiche funzioni semantiche. Essi sono organizzati in quattro contesti, determinati dall’attività e dall’attore principale, per il quale l’applicazione è più rilevante: l’ingegnere del dominio di conoscenza, l’autore del contenuto, il formatore/instructor e l’allievo.

Quanto ai vantaggi, le ontologie possono contribuire allo sviluppo del contenuto sia rappresentando la conoscenza, dal momento che la rendono esplicita, sia comunicando la conoscenza; grazie alla conoscenza così comunicata e condivisa, infatti, s’incoraggia e si facilita ampiamente il riuso del contenuto. I benefici risultano evidenti in una maggiore flessibilità nella creazione e nella gestione del contenuto e in una significativa riduzione dei costi.

Le tecnologie del web semantico eccellono nelle applicazioni delle tecnologie dell’apprendimento in cui si dispiega appieno il potenziale delle ontologie: creazione

e generazione di contenuto, sua organizzazione e sequencing, metadati per il suo reimpiego. Le applicazioni migliori mettono così a profitto la ricchezza semantica della rappresentazione della conoscenza.

I limiti maggiori al completo sfruttamento del potenziale tecnologico sono costituiti dai fattori umani, e cioè dagli atteggiamenti e dalle abilità degli sviluppatori di contenuto e dei progettisti formativi/instructional designers, e dai fattori tecnici, come l'esigua disponibilità sia di supporto strumentale che di ontologie. Difatti spesso le applicazioni presuppongono l'esistenza di ontologie specifiche per particolari domini che nella realtà non esistono e dovrebbero pertanto essere create sulla base delle risorse esistenti.

Fra le tendenze attuali e gli sviluppi che appaiono oggi più promettenti, le tecnologie del web semantico aprono scenari davvero interessanti in due ambiti decisivi, l'adattività e il ragionamento.

Poiché il web è diventato una delle piattaforme fondamentali per l'e-learning, tutte le sue mutazioni incidono profondamente sulla configurazione degli LTS. In primo luogo, il venir meno della distinzione fra il contenuto e il sistema di gestione, una delle linee evolutive prevalenti del web, fa del contenuto un oggetto dinamico per i formatori come per gli allievi. In secondo luogo, si assiste specularmente alla progressiva fusione fra la prospettiva educativa e quella tecnica; in altre parole, la composizione del contenuto (sequencing) tende a divenire tutt'uno con la composizione di funzioni (assemblaggio dei software). Ed è proprio nella composizione che le ontologie e le tecnologie del web semantico potranno giocare un ruolo chiave. Infine, queste tecnologie paiono capaci di favorire in misura considerevole gli approcci educativi oggi predominanti, in special modo l'apprendimento attivo, l'addestramento delle abilità/skills training, l'apprendimento basato sui problemi e sulla esperienza diretta.

Per tracciare un primo bilancio, grazie alle ontologie e alle tecnologie del web semantico si delineano molteplici prospettive nel processo di sviluppo e di spiegamento del contenuto: sono cioè ugualmente supportati unità di contenuto indipendenti, assemblaggi di contenuto, contenuto riutilizzabile. Le tecnologie del web semantico assicurano un progresso graduale negli approcci tradizionali non basati sul web, ma non rivoluzioneranno le tecnologie di apprendimento.

La tecnologia delle ontologie, anziché influenzare direttamente il contenuto e l'apprendimento in quanto tali, offre un valido aiuto al formatore e al progettista formativo, migliorando la creazione del contenuto, la configurazione e il supporto di gestione, l'utilizzabilità e lo scambio di risorse: i benefici maggiori si otterranno in termini di maggiore qualità e in una diminuzione notevole dei costi. Gli allievi trarranno inoltre vantaggio dalla maggiore disponibilità di contenuti di apprendimento e da un accesso ad essi più facile e diffuso. Dallo studio delle applicazioni disponibili, emerge che sono essenzialmente le ontologie a permettere la condivisione e

il riuso delle risorse di apprendimento in una varietà di contesti e ambienti; le ontologie, infatti, le rendono largamente accessibili tramite una rappresentazione di conoscenze standardizzata che può essere comunicata agevolmente. Tuttavia, il potenziale delle tecnologie delle ontologie e del web semantico all'interno dei LTS è stato finora sfruttato con successo soltanto fino ad un certo punto, e si sono al contrario evidenziati barriere, difficoltà e problemi. I principali riguardano gli standard e il supporto strumentale, la scarsità di ontologie esistenti per specifici domini e, da ultimo, la limitata esplorazione delle capacità di ragionamento della tecnologia delle ontologie.

