

Use of the Semantic Web to solve some basic problems in Education

**Increase flexible, distributed lifelong learning,
decrease teacher's workload**

Rob Koper

Open University of the Netherlands
Educational Technology Expertise Center (OTEC)
Valkenburgerweg 167
PO Box 2960
6401 DL Heerlen
The Netherlands
email: rob.koper@ou.nl
www.ou.nl
www.learningnetworks.org

PREPRINT

ABSTRACT

The use of the semantic web in education is explored. Two application areas for use are discussed: a) software agents that support teachers in performing their tasks in flexible online educational settings, and b) software agents that interpret the structure of distributed, self-organized, self-directed learning networks for lifelong learning. The resulting information is used by learners to help persons them perform their tasks in this context more effectively and efficiently. Both of these tasks reuqire a semantic representation of educational entities, specifically the structure of the teaching-learning process, in order to allow for automatic processing.

Self-organised learning networks provide a base for the establishment of a form of education that goes beyond course and curriculum centric models, and envisions a learner-centred and learner controlled model of lifelong learning. In such learning contexts learners have the same possibilities to act that teachers and other staff members have in regular, less learner-centred educational approaches In addition these networks are designed to operate without increasing the workload for learners or staff members. Mechanisms responsible for this efficiency are principles of self-organization and software agents. Both of which are based on semantic web principles that provide support and regulative feedback for both learners and teachers.

Keywords: semantic web, learning design, learning networks, lifelong learning, self-organisation

Introduction

Is the semantic web a tool with realistic educational application? Is it going to change future learning and teaching, and what are the possible directions of this future change?

The basic idea of the semantic web is relatively straightforward: to create a layer on the existing web that enables advanced automatic processing of the web content so that data can be shared and processed by both humans and software (Tim Berners-Lee & Fischetti, 1999; T. Berners-Lee, Hendler, & Lassila, 2001, SW, 2003). Current web pages are structured with (X)HTML tags that provides information about the surface structure of a web page. These tags reveal that every page has a head (e.g. with a title, metadata) and a body with some structured content. The content is structured grammatically in headings, paragraphs, tables, images etcetera. Although some people advocate that this is only a presentation oriented structuring of data, this is only partly true. In essence it provides a semantic structure for the concept of a generic 'page'. A web browser can interpret and process these pages, along with some style sheets, automatically. However, this is only true to the extend that the information is provided in a structured, machine-interpretable way.

For example, a paragraph can be interpreted as a sequence of lines that addresses a common topic. Nothing more, and nothing less. The tagging is generic, doesn't tell anything about *what* content has been structured (a poem, a story, a catalogue, a course), and it doesn't reflect the typical patterns found in different types of documents. This lack of semantic detail is of little consequence if the text is meant for human interpretation only. However, this also implies that the possibilities for automatic processing and manipulation of the web page are restricted to tasks like the ordering and presentation of the paragraph.

Wouldn't it be nice if computers were able to 'understand' web pages so that they can help users to better search for relevant information, make inferences and calculations from the information and combine information in new ways to support knowledge-based tasks such as authoring, planning, navigation, cultural exchange and research. This is the ambitious goal of the semantic web, but it comes at a cost: it requires that more explicit, domain specific meaning ('semantics') be provided by the authors in order to allow for machine-interpretation.

In this article I will explore the use of semantic web technologies in the context of teaching and learning. The usefulness of any technology in any field is dependent on its capacity to address real problems and address practical needs in that field (Mitcham, 1994). Thus, I will make a short inventory of the core technologies in the semantic web, explore some of the current problems and needs in the field of education and will discuss areas where the semantic web technologies can be used to address some of these issues. This exercise cannot be done exhaustively. One way of looking in the crystal ball for future significant developments is to look at current research and technological development (RTD) projects that are working towards the solution of long standing educational problems. So, I will focus on some of our RTD work related to the semantic web, specifically our work on the semantic modelling of educational content and processes, and our work in the realization of self-organized distributed learning networks for lifelong learning. This work is focussed on post-secondary distributed education using Internet technologies.

Some core technologies in the semantic web

The W3C, led by the original creator of the WWW, Tim Berners-Lee defines solutions of the semantic web through use of Resource Description Framework (RDF)-related technologies. However, there are more technologies available to create semantics on the web. I will briefly describe seven of these core technologies in order that the reader can appreciate the multiple tools being used to date to undertake this challenging task:

1. Unified Modelling Language (UML) (Booch, Rumbaugh, & Jacobson, 1999; Fowler, 2000; OMG-UML). UML provides a collection of models and graphs to describe the structural and behavioural semantics of any complex information system. Some of the models provided are:
 - Use case models and scenario's to capture the user requirements and functionality of the system. Scenarios are instances of use cases.
 - Class and object diagrams to specify the semantic information structure of a system. Object diagrams are instances of class diagrams.
 - Activity diagrams to specify workflows.
 - State diagrams to describe the dynamic behaviour of an object in a system.
 - Interaction diagrams (sequence and collaboration diagrams) to model how groups of objects collaborate in some behaviour.
 - Physical diagrams (deployment and component diagrams) to model the implementation structure of a system.
2. XML and XML Schema's (XML, 2003), derived from SGML (ISO 8879). These are tools used to go beyond the fixed, page structure oriented vocabulary that HTML provides. With XML schema's it is possible to structure data and documents according to a personal or community defined vocabularies. These schema vocabularies *can* be of a semantic nature and support a culture of open exchange of data within the communities and tools that understand the vocabulary.
3. RDF and RDF-Schema is the metadata approach from the W3C (RDF, 2003). It doesn't structure the syntax of the data, but defines semantic meaning for data on the web. Multiple semantic perspectives of the same data are possible. The technology is based on lower level technologies: URI's to identify web resources and Namespaces to identify different vocabularies.
4. Topic Maps (ISO/IEC 13250:2000; see Cover, 2003) provide an alternative technology to RDF. Topic maps define arbitrarily complex semantic knowledge structures and allow the exchange of information necessary to collaboratively build and maintain indexes of knowledge. They provide a more general approach than RDF, basically because they are not limited to use in the web environment (and do not use Uniform Resource Identifiers (URI's) as its base).
5. OWL Web Ontology Language According to McGuinness & Van Harmelen (2003), ontology languages provide greater machine interpretability of Web content than that supported by XML, RDF, and RDF-Schema. They do this by providing additional vocabulary along with formal semantics. With OWL it is possible to implement a semantic description of a certain domain by specifying its concepts and the relationships between the concepts. It has three increasingly-expressive sublanguages: OWL Lite, OWL DL, and OWL Full.
6. Latent Semantic Analysis (LSA) (Landauer & Dumais, 1997). The approaches mentioned above require humans to provide the semantic meaning by using a machine interpretable coding scheme. Humans however use natural language to express and interpret meaning. A better approach could be to build programmes that

can understand natural language. This releases humans from typing meta-data and using artificial, labourious coding schemes. One promising approach in this direction is LSA. This technique provides a kind of factor analysis (based on singular value decomposition), that analyses - data (e.g. texts) and orders it on certain underlying dimensions. New text can be interpreted by mapping it on the domain specific underlying dimensions from which it was extracted. This provides for a mechanism to interpret the meaning of words. The development of such language based approaches is a promising field when applied to the semantic web.

7. Software Agents (Axelrod, 1997; Ferber, 1999; Jennings, 1998). One of the basic technologies that can exploit the coded semantics on the web are software agents. The definition of a software agent is rather ill-defined, but most scholars in the field refer to notions such as : a piece of software that can act proactively, is adaptive and (semi-)autonomous and can communicate with other agents and its human creators. In this article I will use the term software agents to refer to all computer programs that can read and process the coded semantics in the data to help humans perform their tasks more efficiently and effectively.

Problems and Needs in Learning and Teaching

One way of looking at problems and needs is by looking at current trends. Howell, Williams, & Lindsay (2003) analysed 32 trends and Merrill (2003) identified current trends in instructional design. Summarized and regrouped on several eLearning domain dimensions (Koper, 2003) we can identify the trends in Table I.

Table I. Summary of problems and needs in education

<i>Dimension</i>	<i>Problems/Needs</i>
I. Changes in Societal Demands	<ol style="list-style-type: none"> 1. Current higher education infrastructure cannot accommodate the growing college-aged population and life-long learning enrollments, making more distance education programs necessary. 2. Knowledge and information are growing exponentially and Lifelong learning is becoming a competitive necessity. 3. Education is becoming more seamless between high school, college, and further studies
II. Changes in Learning Teaching process	<ol style="list-style-type: none"> 4. Instruction is becoming more learner-centred, non-linear, and self-directed. 5. There is an increasing need for new learning and teaching strategies that a) is grounded in new instructional design research and b) exploit the capabilities of technology 6. Learning is most effective when learners are engaged in solving real-world problems; learning environments need to be designed to support this problem-centred approach. 7. Students demand more flexibility; are shopping for courses that meet their schedules and circumstances 8. Higher-education learner profiles, including online, information-age, and adult learners, are changing 9. Academic emphasis is shifting from course-completion to competency 10. The need for faculty development, support, and training is growing 11. Instructors of distance courses can feel isolated
III. Changes in Organization of Educational Institutions	<ol style="list-style-type: none"> 12. There is a shift in organizational structure toward decentralization 13. Higher education outsourcing and partnerships are increasing

	<p>14. Retention rates and length of time taken to completion concern administrators, faculty members, students and tax payers</p> <p>15. The distinction between distance and local education is disappearing</p> <p>16. Faculty members demand reduced workload and increased compensation for distance courses</p> <p>17. Traditional faculty roles are shifting or unbundling</p>
--	---

Stated in more general terms, the longer-term aim of educational change is to (a) increase the effectiveness of education, (b) to increase the flexibility and accessibility of education, (c) increase the attractiveness of education and (d) to decrease the workload for staff (or more in general: to decrease the institutional costs). The relevance of the semantic web for education depends on how much it contributes in the accomplishment of this aim.

My personal expectation is that the semantic web will be of help in two general areas, both related to the fact that it allows for more and better automatic processing of web information:

1. Staff can be helped to perform some of their tasks in flexible, online educational settings more efficiently and less isolated, this includes online course development tasks, learner support tasks, assessment tasks and course management and administration tasks (e.g. setting-up new instances of courses).
2. Persons in different roles (learners, tutors, content providers) can be helped to perform tasks more effectively and efficiently in large, distributed, problem-based, multi-actor, multi-resource learning spaces that are set-up to establish, learner-centred, non-linear, self-directed lifelong learning opportunities.

The first expectation is directed at helping staff members to perform their tasks more efficient. This has an effect on the quality of learning. A common finding is that renewal of education (increase flexibility, use of eLearning, learner-centred approaches) leads in most situations to an increase in staff workload. This is one of the (many) reasons why teachers, schools and universities are resistant to change. When the workload of teachers is decreased, more time is available for more and higher quality support activities. It is expected that this will provide a stimulus to implement more fundamental and necessary innovations in the teaching-learning process.

I will now discuss the two issues. The first will be discussed in relationship to our work on formal, semantic representations of course designs, and the second will be related to our work in the development of self-organized, distributed networks for lifelong learning.

Semantic Representation of Learning Designs

An important question related to the educational semantic web, is how to represent a course in a formal, semantic way so that it can be interpreted and manipulated by computers as well as humans. We refer to this process in terms of 'Educational Modelling' (Koper, 2001; Koper & Manderveld, in press). A semantic model is developed using a variety of methods: literature research, expert group discussions, validation sessions, etcetera and the result is described with a formal modelling language, like UML (Booch *et al*, 1999; OMG-UML). The UML class diagrams can be translated to RDF-Schema and/or OWL Web Ontology Language, depending on the richness of the model. (Chang, 1998; Melnik, 2000). XML-Schema's (XSD) and other semantic bindings like Topic Maps can also be generated from the UML models.

When we are able to represent courses in a semantic way, it opens the possibility to solve (parts of) the following problems, all related to teaching tasks:

- The *development of web-based courses* that are flexible, problem-based, non-linear, incorporate multimedia and are adaptive to learner characteristics involves a large number of disciplines, is expensive and extremely time consuming. A semantic framework can help the course developers in the structuring and integration of the development work. It enables authors to reflect on their thinking (for themselves and in teams) and authoring and design support agents and tools can be created to help the developers to do their jobs more effectively and efficiently.
- An explicit notation of courses can *preserve and share knowledge* about effective (prototypical) learning designs. It opens the possibility to build and share catalogues of effective *learning and teaching patterns* that can be communicated very precisely and can be adapted to other contexts, problems and content (see also: Bergin, Eckstein, Manns, Sharp, & Voelter, 2000).
- *Instantiation* of an eLearning course in current Learning Management Systems (LMSs) can be a time-consuming job that has to be repeated for every new run of the course. One has to assign users, create groups, but also has to set-up the communication and collaboration services (e.g. discussion forums, workspaces, etc.) mostly by hand. A representation of a course that includes a specification of the set-up of the services enables the automation of this instantiation process.
- When the representation of the course does include a semantic, higher level description of the interactive processes that occur during the teaching-learning process, software agents can interpret these to *support learners and staff in managing the workflow* of activities in teaching and learning. These agents can also support the filtering of the appropriate resources to be used during the performance of an activity.
- *Adaptation* to individual learner characteristics is highly desirable since no two learners have the same learning pre-requisites, skills, aptitudes or motivations. However such adaptation can only be done realistically when the adaptation is wholly or at least partially automated. Otherwise it becomes too much work for the learner and/or teacher. When the representation includes descriptions of the conditions for adaptation, this process of adaptation can be software supported.
- *Sharing and re-use* of (parts) of courses is one of the major objectives in the field of eLearning, more specifically learning objects (see Littlejohn, 2003). This sharing and re-use is needed to make course development more efficient; however sharing is hard to do when the learning objects are not semantically represented. The learning objects are hard to find, hard to integrate into new contexts and – for new LMSs that receive learning objects from another LMS – hard to interpret and structure in the correct way.
- An explicit semantic representation can serve as a means to create *more advanced and complex, but consistent learning designs* than is possible without such a representation. This is a characteristic of any language with semantic that enables one to write, read, rewrite and share meaning (natural language, musical notation, etc.).
- And last but not least, a semantic representation of courses enables us to perform *research* into more effective and efficient learning designs. This can be done by comparing the experience with (parts of) learning designs structures in the context of real use.

The representation of courses (or more general 'units of learning') we developed is called 'Educational Modelling Language' (EML, 2000). This language has been the input for – and has recently been replaced by – the IMS Learning Design (LD, 2003) specification that was released in February, 2003.

LD uses a semantic conceptual model of the teaching-learning process as its base. This model states that in any instructional design (classes are in italics):

A person is assigned to a *role* in the teaching-learning process, typically a *learner* or a *staff* role. In this role he or she works towards certain *outcomes* by performing more or less structured *learning* and/or *support activities* within an *environment*. The environment consists of the appropriate *learning objects and services* to be used during the performance of the activities. Which role gets which activities at what moment in the process, is determined by the learning design *method* or by a *notification* (a triggering mechanism).

The semantic, conceptual model has been expressed as a series of UML models, from which several bindings were generated automatically. E.g. for the IMS Learning Design specification a XML schema has been derived that keeps the semantics in the tag-names. However other bindings (RDF Schema/OWL, Topic Maps, SGML schema's, relational database schema's) could in principle be generated as well. This implies that the UML model is the dominant part of the specification; it captures the semantic structure and allows other representations to be generated from it.

Course developers can structure and validate learning designs using the XML (or other) bindings, and software agents, that can read and manipulate the learning design, can support them. The resulting syntactically valid courses can be instantiated and run by any LD compliant software system, a so-called LD engine, player or runtime agent. During instantiation, collaborative and communication services can be set-up automatically, including the assignment of the appropriate user rights to groups and persons.

Although the LD specification has been released quite recently, several first runtime agents have been developed. Important initiatives are Edubox (Perot & OUNL, 2003), Reload (JISC/BOLTON, 2003), the LAMS system (Dalziel, 2003) and the Open Source LD engine (Vogten & Martens, 2003).

It is expected that the semantic model underlying LD, as expressed in UML, is a critical component for the realisation of the Educational Semantic Web, because it provides a tested, generic and (within the IMS community) accepted semantic notation. Whether this model is implemented in XML, RDF-Schema, OWL, Topic Maps etc. depends which tools and technologies are used at any moment in time.

Currently we are working to establish an approach to analyse LD coded courses. We look at the coded learning designs of 31 courses (between 20-200 hours of study each) that have been put into real practice. Some questions we have are:

1. What is the frequency of use of any of the LD elements and of constructs (two or more elements)?
2. Can we find patterns in the tagging within actual deployed LD courses?
3. Which LD constructs are repeatedly used in practice and what is the typical use and typical experience with the construct in practice?
4. Can we set-up a system (including software agents) that can evaluate the effectiveness of learning design patterns in practice?
5. Can we formalize learning design patterns in such a way that they can serve as building blocks in new designs?
6. Is it possible to derive the higher-level role of a semantic object type by using techniques like Latent Semantic Analysis? For example: can we derive a classification of learning activities (explicitly coded in LD), based on an LSA interpretation of the learning activity descriptions and titles? In this case the XML structuring is used to identify certain objects and LSA (or other) to classify the objects at a higher level of abstraction.

7. Can we provide feedback to course developers by comparing a new developed course with the patterns derived from existing (effective) comparable courses. This would enable us to identify weak spots in course designs and suggest improvements.

The advantage of such a research approach is that we can learn from direct experience and share that experience among practitioners. It opens the route to what we refer to as 'inductive learning design' models and approaches. In this approach we aim to inductively abstract patterns from real practice. Most current instructional design approaches evaluate the effectiveness more indirectly, because an explicit semantic representation of the design itself is not available. One thing that has can be developed in future, based on inductive and deductive analysis, are ontology's for *specific* pedagogical approaches. This allows for more specific guidance of agents in helping to create specific units of learning according to the principles of a certain pedagogical ontology. A long-term goal is the possibility of building software agents that can construct some simple units of learning on their own. However, there is still a long road to go before any of these ambitions will be realised in the actual practice of education.

Beyond course centric models: Self-organized Learning Networks for Lifelong Learning

In 2003 we started a new five year RTD programme, called: "Learning Networks: connecting people, organizations, autonomous agents and learning resources to establish the emergence of effective lifelong learning" (Koper & Sloep, 2003). In this programme we will develop self-organized, distributed learning networks for lifelong learning, based on self-organization theory (e.g. Hadeli *et al*, 2003; Maturana & Varela, 1992; Varela, Thompson, & Rosch, 1991) and software agent approaches. The software agents are expected to perform tasks that instantiate principles of the semantic web - specifically LD, Web ontology's and inductive analysis techniques like LSA.

As discussed above, LD is able to represent any learning design model. However, most users in eLearning and distance learning interpret this in terms of modelling courses based on some underlying assumptions. These assumptions are:

- A curriculum consisting of one or more courses that is explicitly designed by teachers, institutes and/or other parties in society.
- The course is developed by a teacher and/or other expert developers.
- The developed course is put into practice by enrolling students and assigning teachers/tutors.
- A student takes a course and the support is provided by the teacher/tutor.
- Assessment is a responsibility of the teacher or a (super-) institutional entity.

However, given the current need for lifelong learning scenarios, the demand for more flexible, self-directed informal and formal learning opportunities and the need for more efficient teaching scenario's, this model is quite restricted and labour intensive. In lifelong learning the roles are not so fixed as implied above: students can be (co-) producers of course materials, can perform assessments (e.g. in peer and self assessment), and can support other students, just like teachers and experts can both teach and learn at the same time in a certain field of expertise. We want to examine a form of education delivery that goes beyond course and curriculum centric models, and envisions a learner-centred and learner controlled model of lifelong learning where learners have the same possibilities to act that teachers and other staff members have in regular, less learner-centred educational approaches, but without increasing the workload for learners and staff members. Mechanisms responsible for this efficiency are principles of self-organization and software agents, based on semantic web principles that provide support and feedback for persons in performing their learning and support tasks in the learning and teaching process.

Self-organization allows the creation of an efficient system with a minimum of planning and control overhead while maintaining maximum flexibility to adapt to learners' needs thereby reducing the current overhead costs in maintenance, planning, control and quality issues. The essence of self-organisation is stated Heylighen and Gershenson (2003) who wrote:

"A self-organizing system not only regulates or adapts its behaviour, it creates its own organization. In that respect it differs fundamentally from our present systems, which are created by their designer. We define organization as structure with function. Structure means that the components of a system are arranged in a particular order. It requires both connections, that integrate the parts into a whole, and separations that differentiate subsystems, so as to avoid interference. Function means that this structure fulfils a purpose."

It is expected that the application of self-organization principles will help empower learners to move beyond passive consumption of e-learning content towards active production (Fischer & Ostwald, 2002). This shift of control aims to help relieve the burden on providers to predict needs, costs, expected use and income, and tilts the balance of responsibility for learning processes towards the learners themselves (see Tattersall *et al*, 2003).

It is recognised that, in putting the learner centre-stage, care must be taken that the shifting of control does not lead to an overburdening or abandonment of learners. Instead, support and guidance must be given to learners in taking up their new responsibilities. Here lies the opportunity for educational providers - to create the best conditions for self-organizing learning networks to flourish. Part of these conditions will be the provision of software agents that provide support in areas like navigation through the network.

Figure 1 provides a high-level use case of a learning network (A UML use case specifies the different functions that different actors can perform in a learning network).

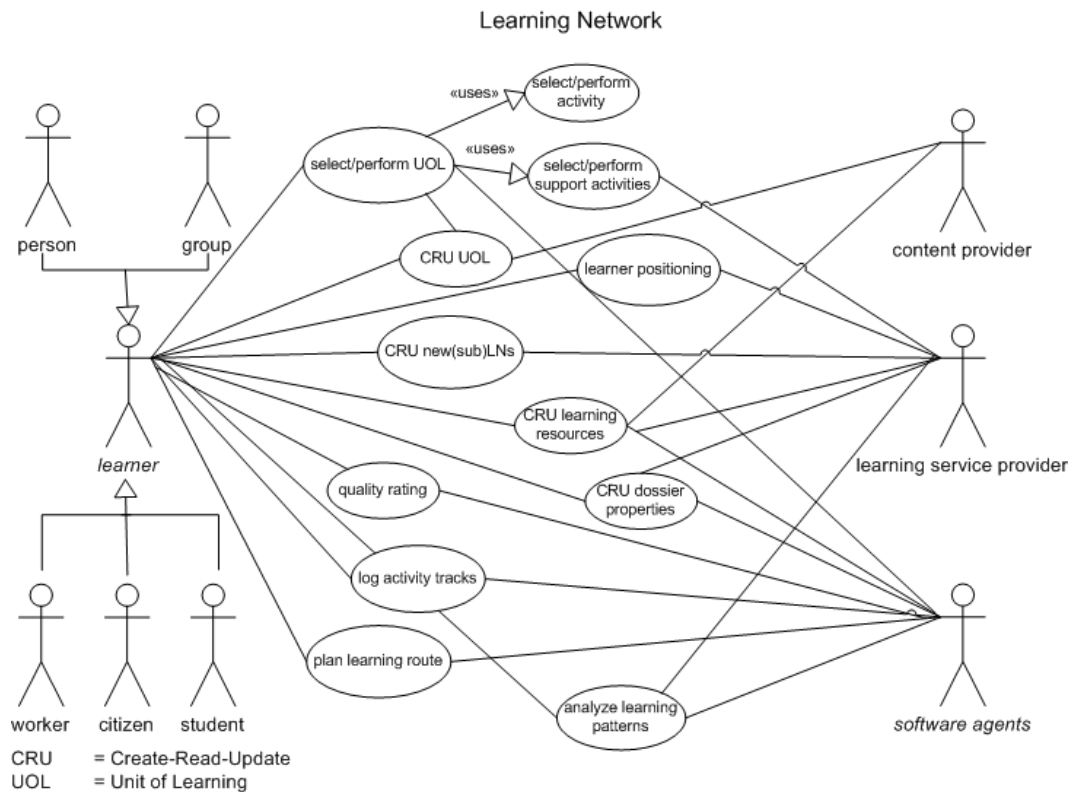


Figure 1: a general use case for learning networks

The use case diagram (figure 1.) specifies several *actors* in a learning network: learners, providers and autonomous software agents. A learner can be an individual person or a group of persons. A further specialization of learners can be given in terms of workers, citizen and students (in regular educational institutes). Different kinds of providers may be distinguished, such as content providers responsible for the provision of learning content (e.g. experts, publishers, libraries) and learning service providers can be distinguished, responsible for tutoring, mentoring, assessment and other learning support functions (e.g. schools, universities, training institutes). Software agents can perform a variety of activities in collaboration with the human actors: sometimes they take over some human activity but in most cases they will support the humans in performing their activities.

The figure specifies several *use cases*, i.e. the activities performed by the actors, represented with oval boxes. A learning network is defined in a certain knowledge or application domain (e.g. psychological diagnosis or eLearning) and consists of a heterogeneous community of humans with a variety of backgrounds and offerings in this field. You can enter the network to learn something new, to increase your level of competence in a certain area or to offer something for others to learn or use. The core concept of the learning network is that it consists of a collection of nodes, each representing a unit of learning (UOL nodes). Every node contains some study tasks, knowledge resources, collaborative services and learning support facilities organized around some learning objective and some prerequisites. Learners can create their own UOL nodes, can use nodes created by others, can collaborate with others to create nodes and can evaluate and rate the quality of UOL nodes. Providers of high-quality materials and courses can do the same. Someone who wants to learn something (a learner) can search for his or her own learning path (a sequence of nodes), explore node after node or can use a predefined route defined by someone else. This route can be analysed on the basis of previous successful path of others or can be pre-planned by e.g. an educational institution. Some UOL nodes can serve as assessment nodes, resulting in a certificate or

diploma that reflects the acquired competencies in the learning network. In a learning network, the user will find several software agents that can support him by performing certain tasks, like the creation of new UOL nodes, selection of adequate learning path, etcetera.

A key notion in the learning network is that it supports learners performing all types of use cases including the ones that traditionally are only available for content and learning service providers. There are no central control actors; the control is expected to emerge under favourable conditions (local feedback, pattern detection) and in a democratic way. This is another way of saying that a person can take all the other roles in a learning network.

A similar argument holds true for quality control: there is no central quality control foreseen in learning networks. It is expected that the network will uphold a variety of different qualities, but that the feedback mechanisms (like ratings and paths) will assure that on the average a satisfactory quality level will be maintained. Thus factors like development costs, frequency of use, incentives, price, and satisfaction may be dynamically balanced. Again this is expected to be an emergent behaviour that will only occur at a certain scale of interactions within the network.

In order to establish learning networks, we identify some core themes and RTD questions (table II).

Table II. Overview of Themes and Questions in Learning Networks

<i>Theme</i>	<i>Questions to solve</i>
MAKE AND USE UNITS OF LEARNING IN LEARNING NETWORKS	<ul style="list-style-type: none"> • How can actors and/or software agents create, share, (re)use and rate units of learning and related artifacts in learning networks? • What is an optimal structure and size of a unit of learning, given IMS LD, and the self-organization characteristics of learning networks? • Can we analyse re-usable patterns in units of learning? • Can we create software agents that create, update and/or use units of learning or help persons in doing so? • What principles and facilities allow the sharing of artifacts/resources in learning networks? • What principles govern the lifetime of units of learning, how and when are they fading out and ultimately deleted, also given principles of partial trust and the time-limited availability of artifacts/resources?
LEARNER POSITIONING IN LEARNING NETWORKS	<ul style="list-style-type: none"> • How can we – at any time - determine the position of a learner in a learning network, given the prior knowledge, situational circumstances and needs of the learner? • Can we create an abstract representation of a learner's position independent of its role in the learning network itself? • Can we create software agents that help the positioning process? • How does a learner know which units of learning match its entry and exit requirements? • How to communicate the position of a learner in a reliable, certified way to external parties (e.g. teachers, employers, other actors).
NAVIGATION IN LEARNING NETWORKS	<ul style="list-style-type: none"> • How do actors know how to proceed in a learning network and can they learn from previous success of others? • How can we store, share and use tracks and track patterns in learning networks to facilitate (personalized) learning plans? • Is it possible to create software agents that help actors to navigate, create optimal routes, etc.?
LEARNING NETWORKS INTEGRATED	<ul style="list-style-type: none"> • How can we create a distributed network of actors/agents and artifacts that optimises the emergence of effective, efficient and attractive lifelong learning in its participants and in the network as a whole? • What principles, theories, models, methods, rules and technologies

	govern such a network? What are its benefits and its restrictions? • What is the critical size in terms of actors/agents, interactions and artifacts for emergence of lifelong learning to become possible?
--	--

As can be seen in the questions, a critical aspect of Learning Networks is to represent units of learning nodes and other artifacts in the network in a semantic way so that software agents can operate on it as well as humans. In this way the principles of the semantic web are expected to be a key issue in the realisation of learning networks for lifelong learners.

In Conclusion

In this article I explored the use of semantic web technology in the educational field. The core ambition of the semantic web is to allow software agents to interpret the meaning of web content, in order to support users in performing their tasks. In order to be able to interpret the meaning of learning objects and services, several semantic modelling and coding techniques are available, like UML, XML schemas, RDF (-Schema), Topic Maps, OWL Web Ontology Language and techniques like Latent Semantic Analysis.

I started the exploration of use of the semantic web in educational by looking at some of the basic problems and needs in education that could be addressed by semantic web technology, at least in principle. Two area's of interest where identified: a) software agents that interpret the semantic structure of units of learning to decrease teacher workload and b) software agents that interpret the structure of distributed, self-organized, self-directed learning networks for lifelong learning to help persons to perform their tasks in this context. Examples of these tasks are: finding appropriate units of learning, creating and adapting units of learning, creating and adapting learning resources, navigating through the network (creating effective, efficient and sensible learning routes), access the current position in the network and provide help with support tasks (e.g. providing feedback on performance; organizing and replying email).

Several areas of work in these two domains of use where specified. To conclude this exploration I want to make some specific additional comments.

Explicit tagging of metadata versus inductive approaches

One of the points addressed in the introduction is the difference between approaches where humans are forced to expend considerable extra effort in adding machine interpretable code to express the meaning of learning artifacts. Specifically: XML tagging of content, RDF and OWL Web Ontology Language are intended to be specified by humans. This is only feasible when there is a clear advantage in terms of cost effectiveness. This is not quite clear at the moment: are the advantages of software manipulation so great that it justifies the increase of effort in the development of ontology's and extra tagging? In practice we can see a lot of practitioners showing resistance when asked to add structured metadata. Why can't machines do this job for us? Because it is meant for machines to understand ...

Perhaps in the long run language interpretation approaches like LSA are a more convenient and more efficient way to proceed. However, these approaches are still very much in their infancy and far from any practical use today. It is also not clear yet, whether these techniques will be used to create the currently specified semantic metadata (RDF, OWL) from inductive analysis of text, or that they represent the semantics more directly in multidimensional schemes (e.g. LSA dimensions are stored in large matrices).

Overlapping open standards

One of the problems in realizing the semantic web in education is that there are different overlapping and non-harmonized standards from different organizations available. This is due, in part, to the immaturity of the field, but also because different techniques for addressing the same problem are sometimes needed. The context and experience of both the problem and the problem solvers are often better tuned and experienced within a particular domain. For most of us, it is not completely clear how these standards relate to each other and whether some of them can be made obsolete or are better suited for a certain job. This is not only true for standards from different organizations, for example Topic Maps from ISO/IEC versus RDF based approaches from W3C, but this is also true for standards that have been developed within the same standards organization.

Examples are the unclear relationship between CSS and XSLT and the relationship between RDF-Schema and OWL, all from W3C. This is even true for relationship of specifications within the same standard. UML from OMG consists of several models that overlap to a certain extent in their modelling abilities (e.g. UML Sequence and Collaboration diagrams) and the unclear relationship of activity diagrams to all other types of diagrams have been heavily debated.

It is not expected that this situation will change for some time. A way of dealing with this problem is not to ignore the standards, but to use the more expressive ones (like UML) and generate – when needed – the ones that can be considered subsets. For example, the UML/RDF-schema note from W3C states explicitly that: 'The RDF-Schema model itself is equivalent to a subset of the class model in UML' (Chang, 1998, p. 2). The same is true for IMS LD, being a semantic model (or ontology) of the teaching-learning process expressed in UML and can derive bindings in XML-schema or OWL. The expressive power of the semantic model of LD allows the encoding of LD units of learning in packaging and implementation schemas like SCORM or IMS Content Packaging. These two open standards provide less semantics: without human interpretation and the provision of a manual it is not possible to interpret the pedagogical structure of – for instance – a SCORM coded course. This hinders conversions from SCORM content to other specifications. Benneker, Hermans, Kresin, Piet, & Verhooren (2003) created for example a mechanism to generate IMS Question and Test Interoperability (QTI) code from EML (predecessor of LD). The other way around is not possible. This is one example that illustrates the power of semantic models, it is possible to convert to less rich schemas at any time.

Lack of user-friendly tools

One of the basic problems at the moment – and expected to continue in the near future – is the lack of user-friendly tools that conform to the different interoperability standards. This is specifically true in the educational field, partly because most vendors are not willing to invest much in this rather small nice market when compared to the general ICT business market. In Europe most educational institutes are also financed with scarce public means. Hard choices have to be made at the government level, in terms of investing in eLearning tools, new buildings or better teacher salaries and job conditions. Because of this situation I advocate the following principles: piggy back where possible on general business software. Avoid using specific eLearning applications where (configurations of) generic tools are possible. Try to influence the (further) development of generic applications by suggesting requirements from the educational field to the developers. However, for certain educational applications – mostly where it comes to the heart of teaching and learning – this will not be appropriate: we find that generic tools are simply not suitable. In this case, I think that we should try to convince our governments and other financing organizations (at least in Europe and other places where education is primarily a public responsibility), to invest in the development of eLearning tools that fully comply to open eLearning standards, are easy available (e.g. open source) and user-friendly. This will open-up new possibility for learning in society: the possibility to upgrade the competency levels of the workforce and the provision of lifelong learning mechanisms to support flexible deployment.

Implementation of change within education

Last, but not least, there is a large number of issues related to the implementation of the semantic web in educational practice. More specifically, because the education system is famous for its resistance to change. The changing role of the teacher in the educational system has to be considered carefully. It is too often ignored. Teachers must learn to tutor in the context of new pedagogical models, and to use online tools – like the agents discussed - so that they can perform the core of their job more effectively, efficiently and with more appreciation. Given attention to 'teacher workload' as advocated above is one step in this direction. However, it is not only the technology, but also the organizational conditions that are a critical part in implementation. One danger is that the organisational constraints are so heavy, that new online approaches loose during implementation some of its critical functionalities and will not work or offer little added value. An example is the use of eLearning applications in residential universities where the technology is just an add-on to existing work, leading to increased workload and costs without increase in the effectiveness of education.

One of the instruments to support further implementation is to establish appropriate communities of practice. An example is the UNFOLD Project (UNFOLD, 2003), that has been funded by the European Commission under Framework 6, to support the development and operation of communities of practice around IMS Learning Design.

References

- Axelrod, R. (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, NJ: Princeton Press.
- Benneker, F., Hermans, H., Kresin, F. J., Piet, M., & Verhooren, M. (2003). *Lab sessie Rechten Online [lab session Law Online]*. Utrecht: Digital University.
- Bergin, J., Eckstein, J., Manns, M., Sharp, H., & Voelter, M. (2000). *The Pedagogical Pattern Project*. Retrieved June 12, 2003, from <http://www.pedagogicalpatterns.org>
- Berners-Lee, T., & Fischetti, M. (1999). *Weaving the Web : the original design and ultimate destiny of the World Wide Web by its inventor (1st ed.)*. [San Francisco]: HarperSanFrancisco.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*(May 17).
- Booch, G., Rumbaugh, J., & Jacobson, I. (1999). *The Unified Modeling Language User Guide*. Reading: Addison-Wesley.
- Chang, W. W. (1998). *A discussion of the Relationship Between RDF-Schema and UML*. Retrieved October 13, 2003, from <http://www.w3c.org/tr/note-rdf-uml>
- Cover, R. (2003). *(XML) Topic Maps*. Retrieved October 14, 2003, from <http://xml.coverpages.org/topicMaps.html>
- Dalziel, J. (2003). *Discussion Paper for Learning Activities and Meta-data*. Retrieved October 14, 2003, from http://mdlet.jtc1sc36.org/doc/SC36_WG4_N0043.pdf
- EML (2000). *Educational Modelling Language*. Retrieved October 14, 2003, from <http://eml.ou.nl>
- Ferber, J. (1999). *Multi-agent Systems*. Reading: Addison-Wesley.
- Fischer, G., & Ostwald, J. (2002). *Transcending the Information Given: Designing learning Environments for Informed Participation*. Paper presented at the Proceedings of ICCE 2002 International Conference on Computers in Education, Auckland, New Zealand.
- Fowler, M. (2000). *UML distilled (second edition ed.)*. Upper Saddle River, NJ: Addison-Wesley.
- Hadeli, P. V., Zamfirescu, C. B., Van Brussel, H., Saint Germain, B., Holvoet, T., & Stegmans, E. (2003). *Self-Organising in Multi-agent Coordination and Control*

- Using Stigmergy*. Paper presented at the The First Workshop on Self-Organising Engineering Applications (ESOA 2003), Melbourne, Australia.
- Howell, S. L., Williams, P. B., & Lindsay, N. (2003). *Thirty-two Trends Affecting Distance Education: An Informed Foundation for Strategic Planning*. Retrieved October 12, 2003, from <http://www.westga.edu/~distance/ojdla/fall63/howell63.html>
- Jennings, N. R. (1998). A Roadmap of Agent Research and Development. *Autonomous Agents and Multi-agent Systems*, 1(1), 7-38.
- JISC/BOLTON (2003). Reload [authoring environment, packager & player]. Bolton: JISC.
- Koper, E. J. R. (2001). Van verandering naar vernieuwing [from change to renewal]. In P. Schramade (Ed.), *Handboek Effectief Opleiden* (Vol. 26, pp. 45- 86 (47.10-41.01- 47.10-41.42)). Den Haag: Elsevier.
- Koper, E. J. R. (2003). Learning technologies: an integrated domain model. In W. Jochems & J. Van Merriënboer & E. J. R. Koper (Eds.), *Integrated eLearning* (pp. 64-79). London: RoutledgeFalmer.
- Koper, E. J. R., & Sloep, P. (2003). *Learning Networks: connecting people, organizations, autonomous agents and learning resources to establish the emergence of effective lifelong learning* (OTEC RTD Programme Plan 2003-2008). Heerlen: Open University of the Netherlands.
- Koper, E. J. R., & Manderveld, J. M. (in press). Educational modelling language: modelling reusable, interoperable, rich and personalised units of learning. *British Journal of Educational Technology*, xx-xx.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*(104), 211-240.
- LD (2003). *IMS Learning Design. Information Model, Best Practice and Implementation Guide, Binding document, Schemas*. Retrieved October 3, 2003, from <http://www.imsglobal.org/learningdesign/index.cfm>
- Littlejohn, A. e. (2003). *Reusing Online Resources: A Sustainable Approach to eLearning*. London: Kogan Page.
- Maturana, H., & Varela, F. J. (1992). *The Tree of Knowledge: The Biological Roots of Human Understanding* (Rev.Ed. ed.). Boston: Shambhala/New Science Press.
- McGuinness, D. L., & Van Harmelen, F. (2003). *OWL Web Ontology Language Overview (W3C Candidate Recommendation 18 August 2003)*. Retrieved October 14, 2003, from <http://www.w3.org/TR/owl-features/>
- Melnik, S. (2000). *Representing UML in RDF*. Retrieved October 14, 2003, from <http://www-db.stanford.edu/~melnik/rdf/uml/>
- Merrill, M. D. (2003). *First Principles of Instruction*. Retrieved October 3, 2003, from <http://www1.moe.edu.sg/itopia/download/abstracts/Applying%20First Principles of Instruction to Technology-Based Education.pdf>
- Mitcham, C. (1994). *Thinking through technology: the path between engineering and philosophy*. Chicago: The University of Chicago Press.
- OMG-UML (2003). *UML Specification, version 1.4*. Retrieved October 14, 2003, from <http://www.omg.org/technology/documents/formal/uml.htm>
- Perot, & OUNL. (2003). Edubox (Version 3.06) [EML Player].
- RDF (2003). *Resource Description Format (RDF)*. Retrieved October 14, 2003, from <http://www.w3c.org/RDF/>
- SW. (2003). *Activity Statement Semantic Web*. Retrieved October 14, 2003, from <http://www.w3.org/2001/sw/Activity>
- Tattersall, C., Manderveld, J., Van den Berg, B., Van Es, R., Janssen, J., Waterink, W., Bolman, C., & Koper, E. J. R. (2003). *Road Mapping (ROMA)* (OTEC/LTD Project Plan). Heerlen: Open University of the Netherlands.
- UNFOLD (2003). *Understanding New Frameworks of Learning Design (IST-2002-1_507835)* (Project Plan).
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: MIT Press.
- Vogten, H., & Martens, H. (2003). Open Source Learning Design Engine (software). Heerlen: Open University of the Netherlands.

XML (2003). *Extensible Markup Language (XML)*. Retrieved October 14, 2003, from <http://www.w3c.org/XML/>