Licentiate's Thesis

Using Metadata and Ontologies in Virtual University

Päivi Pöyry

Information Ergonomics Research Group
Software Business and Engineering Institute
Department of Computer Science and Engineering
Helsinki University of Technology
December 9th, 2004

HELSINKI UNIVERSITY OF TEHCHNOLOGY

ABSTRACT OF LICENTIATE THESIS

Department of Computer Science and Engineering

Author
Päivi Pöyry

Date
9.12.2004

Pages
93 (+appendices 29)

Title of thesis
Using Metadata and Ontologies in Virtual University

Language
English

Chair
User interfaces and usability

Chair code
T-121

Supervisor
Professor Marko Nieminen

Instructor
Professor Marko Nieminen

Abstract
In the past few years the field of e-learning has grown significantly due both to the rapid development of the enabling information and communication technologies and to the continuously increasing demands for specific and up-to-date skills in industry. These two factors have lead to the rise of e-learning, where Internet has begun to serve as a channel for delivering education and training to the learners whenever and wherever they needed it. However, the information on the educational resources is dispersed and poorly classified, and the existing web technologies are unable to understand the content and the context of the learning resources. As a possible solution for bringing order to the chaos, a structured and standardized metadata with a domain ontology may be used. A standardized metadata set helps the providers of educational resources to classify their supply efficiently, and a domain ontology enables shared understanding on the subject area by defining the semantics and the relationships of the concepts used.

In this study the main focus was on developing metadata schemas and ontologies suitable for the virtual university taking the users' perspective into account. Thus, this thesis, as being applied constructive research on virtual learning environments, falls in a multidisciplinary field by combining knowledge from education science, computer science, and software engineering with the emphasis on the latter ones. Furthermore, the topics dealt with in this study included e-learning, conceptual modeling, ontology development, metadata specification, web service development and standardization work, to mention but a few. The goal of this thesis is to describe the use and development of an educational metadata model and ontologies in an educational information system supporting an international virtual university.

One of the key findings is that the use of a standardized metadata model together with conceptual modeling of the domain is essential in improving the information systems in the fields of e-learning and virtual universities by unifying the descriptions and classifications of educational resources. However, the use of metadata alone does not solve the problems. A domain ontology integrated into metadata is needed to form a basis for common semantics and shared understanding of the domain concepts and their relationships. Secondly, another key finding of this study is the need to modify general standards to meet the needs of a specific application area. The evaluation of the CUBER metadata proved that its users generally accepted the metadata model, and it also fulfilled the users' information requirements.

Keywords
E-learning, Metadata, Ontology, Conceptual modeling, Virtual university

TEKNILLINEN KORKEAKOULU
Tietotekniikan osasto

LISENSIAATINTUTKIMUKSEN TIIVISTELMÄ

Tekijä Pöyry, Päivi Tuuli Helena	**Päiväys** 9.12.2004 **Sivumäärä** 93 (+liitteet 29)
Työn nimi Using Metadata and Ontologies in Virtual University	**Kieli** Englanti
Professuuri Käyttöliittymät ja käytettävyys	**Professuurin koodi** T-121
Työn valvoja Professori Marko Nieminen	
Työn ohjaaja Professori Marko Nieminen	

Tiivistelmä

Muutamien viime vuosien aikana tapahtunut valtava tietotekniikan kehitys sekä työelämän jatkuvasti kasvaneet osaamisvaatimukset ovat vaikuttaneet e-oppimisalan nopeaan kasvuun. E-oppimisen ja e-opetuksen jakelukanavana toimii internet, mikä on mahdollistanut opiskelun ajasta ja paikasta riippumatta. Kuitenkin opetusresursseja koskeva tieto on ollut internetissä hajallaan ja huonosti järjestettynä. Lisäksi tähänastiset web-teknologiat eivät ole kyenneet tulkitsemaan verkkoresurssien sisältöä eikä niiden kontekstia. Yhtenä mahdollisena ratkaisuna tässä tutkimuksessa esitetään standardoidun metadatan ja ontologian hyödyntämistä osana tietojärjestelmiä. Standardoitu metadata helpottaa verkko-opetusresurssien organisointia, ja ontologia mahdollistaa aihealueen käsitteiden välisten suhteiden ja semantiikan määrittämisen.

Tämän tutkimuksen tarkoituksena oli virtuaaliyliopiston sovellusalueelle sopivan metadatamallin ja ontologian kehittäminen käyttäjien tarpeet huomioiden. Tutkimusote on siten konstruktiivinen, ja luonteeltaan virtuaalisten oppimisympäristöjen tutkiminen on monitieteistä yhdistäen mm. tietotekniikkaa ja kasvatustiedettä. Tutkimuksen aihepiiri sivuaa mm. e-oppimista, käsitemallintamista, ontologian kehittämistä, metadatan määrittämistä, web-palvelun kehittämistä ja standardointityötä. Tämän tutkielman tavoitteena on kuvata virtuaaliyliopiston tietojärjestelmää varten tehtyä metadatamallin ja ontologian kehitystyötä lisäten siihen käyttäjälähtöisen näkökulman.

Tutkimuksen keskeisiä tuloksia on metadatan sekä käsitemallinnuksen käyttämisen merkityksellisyys e-opetusjärjestelmien toiminnan parantamisessa verkko-opetusresurssien kuvauksia ja luokituksia yhtenäistämällä. Metadata yksin ei kuitenkaan riitä ratkaisuksi vaan sen lisäksi tarvitaan ontologia, joka määrittelee sovellusalueen semantiikan ja käsitteiden väliset suhteet. Toinen keskeinen tulos on se, että metadatastandardeja on mahdollista muokata sovellusalueen ja käyttäjien tarpeita vastaaviksi tiettyjen rajoitusten puitteissa. Lisäksi käyttäjien tarpeiden ja vaatimusten huomioiminen tietomallia kehitettäessä on ensiarvoisen tärkeää, jotta tietomallin hyödyntäminen onnistuu luontevasti sovelluksen käyttäjiltä.

Avainsanat
E-oppiminen, Metadata, Ontologia, Käsitemallinnus, Virtuaaliyliopisto

Acknowledgements

I'd like to thank my colleagues in the Information Ergonomics Research Group at the Helsinki University of Technology for their support during this challenging thesis process. Also I want to thank my colleagues during the CUBER project: Heli Kautonen, Kirsi Pelto-Aho, and Juha Puustjärvi. This thesis is dedicated to the sole sunshine in my life, to my little daughter Norma.

In Espoo, December 9th 2004

Päivi Pöyry

Table of Contents

Appendices

1 Introduction

1.1 Motivation

Today people in all professions are faced with increasing demands. Technology has evolved in an ever-increasing speed, and the roles of people in work, society and industry are shifting constantly. Keeping up with the pace requires continuous education. Traditional universities are trying to answer to this need of lifelong learning by building virtual universities, whilst facing competition from the companies providing continuing education in the form of e-learning.

In the past few years the field of e-learning has grown significantly due both to the rapid development of the enabling information and communication technologies and to the continuously increasing demands for specific and up-to-date skills in industry. These two factors have lead to the rise of e-learning, both in the public educational sector and in the private corporate training. Internet has begun to serve as a potential channel for delivering education and training to the learners whenever and wherever they needed it. Not only educational material has been distributed via the Internet but also courses and even degrees can be studied on the net. This has lead to the internationalization of the educational market, since the Internet enables the providers of educational resources to act regardless of the national borders. Furthermore, the traditional universities have faced a challenge from the private educational sector. As new knowledge is needed just on time in industry, the Internet as a fast communication and education channel has proved to offer a competitive advantage when compared to the relatively slow traditional face-to-face education that is more strictly tied to time and place. (Krämer 2000.)

The Internet has become a vast repository of various educational resources, but it has become increasingly difficult for the learners to find the courses or learning materials they are looking for. The information on the educational resources is dispersed and poorly classified, and the existing web technologies are unable to understand the content and the context of the learning resources (Krämer 2000). As a possible solution for bringing order to the chaos, a structured and standardized metadata with a domain ontology may be used, and the technologies of Semantic Web, such as XML and RDF, may prove to be useful. A standardized metadata set helps the providers of educational resources to classify their supply efficiently, and a domain ontology enables shared understanding on the subject area by defining the semantics and the relationships of the concepts used. This not only builds a basis for improved searching and finding of the educational resources from the web, but also enables shared understanding on the educational resources between the educational institutions and the learners seeking for new skills and knowledge.

1.2 The framework of the research

This study was conducted as a part of the CUBER research and development project that was an Information Society Technology (IST) Project in the Fifth Framework Programme of the European Community. The project has started in April 2000 and run for 30 months ending in September 2002. The CUBER project was an international European joint effort project of many distance-teaching universities to provide information infrastructure for an e-learning system. One of the results the CUBER project provides is the e-learning portal and specific metadata set attached to it. (CUBER 2004.) The main focus of this study is the metadata development work carried out within the CUBER project, but this thesis also presents a literature survey about metadata and ontologies, and an additional user centered analysis of the CUBER metadata model.

The name CUBER stands for Personalized Curriculum Builder in the Federated Virtual University of the Europe of Regions. The main goal of the project was to develop an educational information system designed to become a broker system that supports the search for study courses from several European universities. It was expected to facilitate the access of various kinds of learners to a vast collection of higher education courses offered by European course providers, in particular by distance teaching universities. The motivation of the CUBER project lies in the rapidly growing educational demands in the fields of knowledge-intensive industries, such as ICT enterprises as well as the growth of Web-based learning or e-learning. This situation has lead to competition between the European universities that now must strengthen their teaching strategies and expertise profiles in order to succeed in the internationalization of the higher education market. As a part of the CUBER project a metadata schema was to be developed in order facilitate the efficient searching of courses. (CUBER 2004.) The results of this metadata specification work will be presented and further evaluated and analysed in this thesis with respect to the user friendliness.

The starting ground for this study is the CUBER system and the different activities related to designing and building it. The goals defined for the CUBER system and for the project directed strongly the metadata specification work that was done during the project. In the project plan of CUBER the issue of user centeredness was not taken into account sufficiently, and this thesis aims to fill this gap by bringing forward the elements of taking users into account in the metadata specification process. Thus this thesis consists of both the results of the CUBER's metadata work, as well as an additional literature survey and a further user centered evaluation of the metadata model, which were not done during the CUBER project but afterwards as a part of this thesis.

The CUBER system was designed to be a search engine or a broker system that enables many kinds of potential students to search for study units from institutions providing higher education. The CUBER system aims at matching the needs of the learners with characteristics of the courses offered by the universities (CUBER 2004.) Technically, the CUBER system consists of three main components: 1) a knowledge base for standardised course descriptions and domain knowledge, 2) a search engine for finding the courses and generating study packages, and 3) an authoring interface for entering and maintaining course metadata (see Figure 1). The knowledge base includes a lexical database, where the standardised vocabularies and classifications are entered. The knowledge base also includes standardised metadata and an ontology, which defines the semantics of the metadata and the vocabularies used in the system.

As a result of the metadata specification work, the information model for the CUBER system was developed. The basic information resources related to the CUBER system are *learning objects*, namely *study programme*, *study package*, *study course* and *study material*. All these learning objects have metadata attached to them. All of these metadata fields and their corresponding data types were defined in a separate *metadata specification*. Some of the metadata fields require structure for the data input, which is dealt with by using an ontology that provides the choice. Essentially, the ontologies give the semantics to the metadata. Consequently, the ontologies provide a shared and common understanding of CUBER's metadata items, which is necessary for a fruitful communication between the learners and the educational organisations with the help of the CUBER system. In order to support this communication well enough, it is important to take into account the users' needs when defining the metadata and ontology.

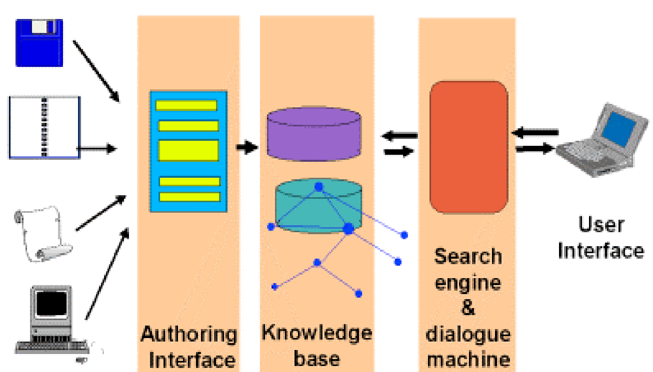


Figure 1. The Architecture of the CUBER system

The goals of this research were very much practical, and the research methods were closely tied to the system development and its practices. Thus the research approach was constructive and oriented towards solving a real life problem instead of solely a

theoretical one. Accordingly, the results of this study are practical solutions to these problems adding pieces of understanding to the existing knowledge. However, the system development work was accompanied by a literature survey in order to ensure that the previous research results and theories were taken into consideration. Moreover, scientific research methods were used throughout the study.
The research topics include the question of requirements for a metadata set in the context of a virtual university, the problematic related to building a domain ontology in this field, and the issues related to metadata and ontology representation in a user-friendly way. As the main results of the metadata related research, the CUBER metadata model, a specific domain ontology, and a conceptual model are presented. Moreover, the requirements related to educational metadata and ontology as well as the results of a user acceptance test are introduced in this thesis.

1.3 The scope of the thesis

This thesis focuses on developing an application profile of the LOM metadata standard for the needs of a virtual university. This thesis also includes a user-centered perspective for the metadata development work, even though only in a supporting role. The thesis is very much based on the work done in the CUBER project, but also additional contributions (extended literature survey and additional evaluation of the model) are included. This thesis, as being applied constructive research on virtual learning environments, falls in the common field of many branches of science. Basically, the nature of this field of study is multidisciplinary combining knowledge from education science, computer science, and software engineering with the emphasis on the latter ones. Furthermore, the topics dealt with in this study included e-learning, conceptual modelling, ontology development, metadata specification, web service development and standardization work, to mention but a few.

Research done in the field of virtual learning environments is however rather dispersed in the sense that each branch of science has its own distinctive perspective, topics, methods, and fora. For example, education science has concentrated on the learning theory, communication, or administrative aspects of the virtual learning environments.

Computer science for its part has focused on the information system perspective, such as algorithms and information retrieval. A good example of the lack of integration is the IQ-Form research project (IQ-Form, 2004) that has been co-operated between several Finnish universities. The computer science departments participating in the project have concentrated on the system development and applying on e.g. Bayesian algorithms, whereas the participating education science departments have focused on researching and applying the current learning theories in the context of supporting e-learning. The diversity of the research done in the field of virtual learning environments is also reflected in the conferences where the results of the research are reported. E-learning and virtual learning environments are included in many kinds of conferences ranging from the educational and social sciences to computer science, human-computer interaction, and even artificial intelligence. In this thesis a multidisciplinary approach was chosen in order to combine knowledge from computer science and educational science, with a user-centered perspective. The results of the study have been reported in conferences ranging from computer science, learning technology, information systems, to human-computer interaction.

The main goal of this thesis is to describe the development and use of metadata and ontologies in an educational information system supporting an international virtual university. Basically, there are three research themes under which the research questions are grouped: 1) Using metadata, 2) Using ontology, and 3) Conceptual modeling and information representation in the context of virtual university. The exact research questions and research methods are presented in the chapter 4 of this thesis.

This thesis consists of work done in the CUBER project’s metadata workpackage, for which the author of the thesis was mainly responsible. Thus most of the results presented are based on the author’s own research work or work done in close collaboration with the other project researchers. However, parts of this thesis are based on other work done in the CUBER project, such as the user study that aimed at defining the user requirements for the metadata model. The other researchers’ work is duly referenced to wherever it is used. In addition to the CUBER project’s work, this thesis presents further work, such as a literature survey, and an additional scenario-based evaluation of the metadata model.

This thesis and the information presented in it are to a great extent based on the following four papers published in international scientific conferences. The author of this thesis had as the first author of all papers the primary responsibility of writing these papers and providing the content related to metadata, virtual learning environments, ontology, and the CUBER system. The other authors have provided limited contribution to the papers. Mr. Puustjärvi and Ms. Pelto-Aho provided advice and some text about the ontology development, whereas Mr. Repokari and Ms. Kautonen contributed to the conceptual model of metadata and the evaluation of the metadata model.

- Pöyry, P., Pelto-Aho, K. & Puustjärvi, J. (2002). *The Role of Metadata in the CUBER system.* The Proceedings of SAICSIT 2002, Port Elizabeth, South Africa. ACM International Conference Proceedings Series. Pp. 172-178.

In this paper the role of using metadata in an educational information system were presented together with the system developed in the CUBER system. The purpose of this paper was to introduce the use of metadata in the context of an emerging virtual university system and the challenges of defining a domain specific metadata model. This paper also presented the use of standardized metadata and an ontology as a possible solution to the emerging problems of the e-learning area.

- Pöyry, P. & Puustjärvi, J. (2003). *CUBER – A personalized Curriculum Builder.* ICALT 2003, Athens, Greece. Pp. 344-345.

This paper presents the use of ontology integrated to metadata as a means to match the learners' profiles (queries) against the courses metadata (course profile). The aim of this paper was also to illustrate the importance of metadata and ontology in supporting the search functionality of learning resources.

- Pöyry, P. , Repokari, L., & Kautonen, H. (2003). *The Conceptual Model for E-Learning Meta-Data Structure.* HCI International 2003, Crete, Greece. Pp. 844-848.

In this paper a Conceptual Model for Metadata information in e-learning based on the research done within the IST R&D project CUBER was introduced. The conceptual model concentrated on the basic concepts or elements of the metadata structure. The first experiences and evaluations of the usefulness of this system and the conceptual model were reported in this paper.

- Pöyry, P. (2003). *Building the Information Infrastructure of a European Virtual University.* SCI 2003, Orlando, Florida, USA.

In this paper, the architecture, functionalities and information structures including metadata and ontologies were presented, and the lessons learned during the process of system development were discussed. This paper drew together the research and development work done in the CUBER project.

This thesis is organised as follows: Section 1 introduces the themes of the thesis and the framework of the study. Section 2 provides a theoretical overview on the fields of e-learning and virtual university. In section 3, the theoretical foundations of using metadata and ontologies are presented. Section 4 introduces the research questions and the methodology used. The research questions will be answered in section 5, where the detailed results of the study are presented. In section 6 the results will be discussed and the thesis will be concluded.

2 The emergence of e-learning and a virtual university

2.1 E-Learning

E-learning can be defined as information technology enabled and supported form of distance learning, in which the traditional restrictions of classroom learning have disappeared. The main tool of e-learning is a personal computer, and the Internet servers as the principal communication and distribution channel. The learners can participate in online Web-based courses and interact with both the peers and instructors and with the learning materials. The teacher-centeredness of traditional learning does not hold for e-learning, where the learning process has become more and more learner centered. The learning process and the resources may be customized according to the individual needs of the learner. At the same time, the role of the teacher becomes that of a facilitator or of a mentor guiding and supporting the individual process of learning. Typical e-learning environments, such as WebCT and Virtual-U, offer the basic elements for delivering e-learning courses: course content delivery tools, synchronous and asynchronous discussion forums and conferencing systems, possibilities for quizzes and polling, workspaces for sharing resources, white boards, possibilities for evaluation and grading, logbooks, possibilities for submitting assignments etc. (Liu, Chan, Hung, & Lee, 2002.)

The major difference between traditional distance education and e-learning is the enabling technology that helps to overcome the barriers of time, place and distance (Krämer, 2000). E-learning has been referred to as a new form of studying that is free from the constraints of time and place. E-learning has served as a tool of knowledge management and competence development for many knowledge intensive companies. (LearningCircuits, 2002.) However, e-learning is not restricted to business life and corporate training. Instead, e-learning has started to gain increasing interest at the traditional universities that have faced the growing challenge of going on-line with their learning and teaching resources. The rise of e-learning sector has given the universities a chance to start building virtual universities that operate mainly on the Internet as virtual organizations formed by one or several educational institutions.

According to Marquardt and Kearsley (1999) the use of e-learning technologies has many advantages when compared with the traditional training and education.

- E-learning is available when and almost wherever it is needed, and it becomes relatively independent of time and place. The learners can study when they have the time and learning can more easily be integrated into work.
- E-learning is “learner-controlled” in the sense that the learners themselves can decide when and where to study instead of having to adapt to the arrangements of traditional face-to-face education.
- E-learning is cost-effective. With e-learning the travel costs will be reduced when the learners no longer have to travel to different locations in order to participate in training. In addition, the time people will have to spend out of the office reduces, and in many cases fewer teachers are needed for larger audiences.
- E-learning is self-paced, which means that the learners can access and manipulate the learning materials and environments as fast or slowly as they want.
- E-learning can be delivered across wide geographical distances via the Internet. Education is available at the office, on the business trips, and at home. In principle the same courses can be used in several countries, which may be an advantage for multinational organizations that need to educate all their members.
- E-learning enables interactive learning process by offering tools and possibilities for discussions. One-to-one, one-to-many and small-group discussions and interaction is supported in order to facilitate learning in collaboration.
- E-learning enables the companies and educational institutions to increase uniformity of learning content and delivery. Same, consistent learning material can be distributed throughout the organization e.g. via intranet learning environments.

- E-learning can be customized according to the individual learning styles and preferences. There are many learning styles, such as visual, auditive or kinesthetic. The e-learning environments can be used to support these learning styles in order to achieve an optimal learning result.
- E-learning resources can be continuously updated. For example, new information may be easily added to the web sites, and new learning resources can be made available through learning environments for large audiences with little effort.
- With e-learning both push and pull approaches are available. In traditional education the learners may get too much or too little information or the information may be poorly focused. The e-learning technologies enable improved filtering of the learning resources, and the learners get as much information they need on the topics relevant to them. (Marquardt, & Kearsley, 1999.)

2.2 Overview of virtual university

In the recent years the idea of a virtual university has been becoming more and more popular in many countries all over the world (Krämer, 2000). The enormous development in the field of Information and Communication Technologies has enabled the rise of e-Learning and virtual learning environments. As a result, the traditional universities have faced a new challenge emerging from the commercial sector of education. There is a growing need for new kind of learning and teaching as the technology advances rapidly and the skills and competencies required in the working life become more demanding and increasingly dynamic.

Virtual university has been defined as a space where the students are provided with higher education courses with the help of the newest information and communication technology. The degree of utilizing technology in organizing the studies may vary from pure technology-based studies to face-to-face or mixed studies that are supported by learning technologies. The main channel of communication and delivery of teaching is the Internet. (Niemi, 2002; Ryan, Scott, Freeman, & Patel, 2000.) Thus virtual university can be seen as closely related to e-learning that provides learning opportunities via the Internet. The difference between these two concepts is the level of studies offered; virtual university aims at offering higher education studies while e-learning can be used for all educational levels.

A virtual university may be an institution that uses the information and communication technologies for its core activities such as providing learning opportunities, administration, materials development and distribution, delivering teaching and tuition, and providing counseling, advising and examinations. On the other hand, a virtual university may also be a virtual organization created through partnerships between traditional universities and other educational institutes. In addition, the traditional campus universities may be regarded as virtual universities if they offer learning opportunities via the Internet or combine traditional ways of learning with e-learning. (Ryan et al., 2000.)

Virtual university is expected to offer opportunities for life-long learning also for audiences otherwise excluded from university studies. The emerging virtual university can be seen very beneficial especially for the industry, when technology-supported learning can be brought to the workplaces and integrated more closely to work. Moreover, virtual university can enhance organizational learning and bring competitive advantage by continuously developing the skills and knowledge of the employees. (Teare, Davies & Sandelands, 1999.)

Resource-based learning (RBL) is one of the corner stones of learning and teaching in the virtual university. RBL has been defined a student-centered way of learning that exploits various specially designed learning materials, interactive media and technologies. The Internet is able to store and transmit vast amounts of information in different forms and formats. Therefore the Internet is an ideal support for resource-based learning. RBL can be realized as self-study or as interactive group learning both in distance and in the face-to-face mode. (Ryan et al., 2000; Nevgi & Tirri, 2003.) As RBL is seen as a promising form of learning for the web-based virtual university, the problem of managing the learning resources emerges. Effective RBL requires easy access to learning resources, which can be facilitated by classifying and organizing the resources with the help of a standardized metadata set.

The development of the virtual university and virtual learning environments falls into the common fields of many sciences. It requires indeed multidisciplinary research co-operation combining expertise from information technology, computer science as well as from the education science. The research on virtual learning environments is closely related to the research on virtual university. Moreover, the basic research and experiences on distance education, educational systems, emerging e-learning and educational politics are needed as much as knowledge on building information systems. By combining the various branches of science it is possible to offer a deep insight into the research of educational technology and its current applications.

The emerging virtual universities can be seen as some kind of virtual organizations as they do not necessarily have physical organization at all. They typically consist of several universities' co-operative projects that aim at offering distance teaching via the Internet. One example is the emerging Finnish Virtual University (http://www.virtuaaliyliopisto.fi) that serves as a common portal for the various universities' supply. The virtual universities may also be universities that have both traditional campus teaching and distance teaching (www-based teaching) such as Helsinki University in Finland, or universities with only distance teaching such as FernUniversität Hagen in Germany (http://www.fernuni-hagen.de).

Many students and other potential learners are also interested in distance learning because of their possible personal conditions and constraints, such as combining work and studying (Krämer, 2000; Twigg & Oblinger, 1996). Some students may also be interested in courses offered by other universities, and taking a course via Internet enables the student to tailor the contents of the studies according to personal preferences.

On the other hand, some problems may also rise as the virtual universities emerge. If the student takes courses from several (possibly distance teaching) universities and wants to combine them to a single degree, then the question of the recognition of the courses becomes highly relevant for both the student and for the universities. The universities need some kind of a common ground or agreement for the recognition of studies, and the student needs some kind of a affirmation of being able to include other universities' courses to the degree he or she is studying for.

The problematic of recognition is in part political in the nature if there are universities involved from several nationalities in the virtual universities. One solution to this problem is the ECTS standard (European Credit Transfer System) developed by the European Community (ECTS, 2004). However, using this standard alone is not enough to solve the problems, since it is only a measure for the work required by the courses, not a description or classification for the courses' contents. There is a clear need for using standardized metadata with an ontology to give universities a common ground for communication and recognition of studies.

3 Metadata and ontology in virtual learning environments

The field of e-learning is a demanding application environment for the area of computer science developing metadata and ontology, because the semantic classification of the information is affected by many factors, such as nationality, level of pre-existing skills, different educational systems, and socio-economical and educational background of the individuals. Due to all these factors and the different semantics used in the field of education, the efficient retrieval and utilization of the information is troublesome at the moment. From this follows that the question of efficient and easy information retrieval must be successfully dealt with if e-learning is to play a significant role in the educational market, where there are currently lots of courses on-line. But this information is difficult to find due to the unarranged form of the information. There if an evident need for standardized metadata and ontology in order to solve this problem. (Krämer, 2000.)

The field of e-learning is facing a real challenge: how to organize and classify educational information in an effective manner in order to facilitate the information retrieval (Krämer, 2000). Even more challenging may be the problem concerning the consistent use of concepts and terminology in the field of e-learning, where the various national practices. One possible solution for these problems is a conceptual model for e-learning and the use of standardized metadata for organizing the educational information. Domain ontology may be attached to the metadata in order to define the semantics of the concepts and vocabularies used in the application area. However, metadata alone does not solve the problem, but it is essential to organize the metadata in an effective manner. Then, with a highly organized set of standardized metadata the learning resources can be described, organized and classified effectively, which in turn enables effective information searching and retrieval. And in addition, the ontology attached to the metadata will provide ground for common and unambiguous understanding of the domain concepts.

3.1 Metadata and educational metadata

Metadata has often been defined as descriptive and classifying *information about an object*. It describes certain important characteristics of its target in a compact form. Metadata plays a central role in improving searching and categorizing objects within a defined context of use, which in this study is a virtual university. In order to be able to use metadata efficiently across different contexts and systems, the metadata scheme should be standardized. There is a growing interest in using metadata in the field of education and e-learning (Britain & Liber, 1998; Jokela, 2001).

Metadata has also been defined as *data about data* by Wason (2001). In this sense metadata describes a data set and the format of this data. In addition, metadata can be described by a set of *meta-metadata*. Meta-metadata is descriptive information on the metadata record itself. (Wason, 2001.)

According to Jokela (2001), metadata can be categorized in many ways, and for example, the following three main categories can be used; structural, control, and descriptive metadata (Jokela, 2001). *Structural metadata* is used to describe the structural characteristics of the object, such as the format of the object, but it does not itself contain any information on the content of the object. *Control metadata* is created and used for controlling the flow of content in the information system in question. *Descriptive metadata* can be further divided in two sub-categories: *contextual metadata* and content-based *semantic metadata.* With contextual metadata we mean the conditions and the environment in which the metadata is created, e.g. the equipment needed to produce the actual object. Semantic metadata refers to the semantic characteristics of the object, i.e. the semantic metadata explains the meaning of the object (Jokela, 2001). Using semantic metadata requires commonly agreed semantic interpretations among all the users of metadata. Semantic metadata is very much domain specific, which means that the nature of semantic metadata is highly dependent on the concepts and semantic structures of the specific field (Jokela, 2001), e.g. higher education. Semantic metadata is also closely related to the concept of *ontology*, which will be discussed later in this thesis.

Educational metadata is needed for improving the retrieval of learning objects, for supporting the management of collections of learning objects, and for supporting the decision process of the learners looking for educational resources. LOM seems to be the most powerful and most widely used metadata standard for educational information systems. (Lamminaho, 2000; Holzinger, Kleinberger & Müller, 2001) More generally, educational metadata can be used by educational institutes and professionals as well as by learners in order to describe e.g. the content, structures, and relationships of the learning objects and to search for educational objects (Lamminaho, 2000; Stojanovic, Staab & Studer, 2001).

Educational metadata may describe any class of educational objects, such as study courses. The pedagogical features of the course, the contents, special target groups, and the technical requirements of the study course can be described with the help of a metadata schema (Lamminaho, 2000). More generally, educational metadata can be used to describe e.g. the content, structures, and relationships of the learning objects (Stojanovic, Staab & Studer, 2001). Educational metadata can be utilized by educational and pedagogical professionals, by the institutions offering education, and by the students searching for education. Well-designed and sufficient metadata aid the decision making process of the students and help the educational institutions to provide suitable information about their educational supply (Lamminaho, 2000). Educational metadata is very much semantic metadata, but a thorough metadata schema must include also at least structural metadata in order to be able to describe the learning objects efficiently.

The idea of using standardized metadata schemas is being able to develop universally applicable tools dealing with the metadata descriptions of the learning objects. In order to create metadata records containing the resource descriptions specific tools are needed for creating the metadata according to the standards. (Kassanke, El-Saddik & Steinacker, 2001.) Metadata is also useful when guiding non-experienced users through a large collection of learning resources (Strijker, 2001). Moreover, metadata is seen as value-added information that is used to arrange, describe, track or otherwise enhance the access to the object content. At the moment metadata becoming increasingly important when digital government and e-commerce are emerging. Metadata enables increased accessibility, expanded use of objects, multi-versioning, and system improvement. The granularity of metadata, which refers to the level of details in the description, is an important question when developing a metadata set. (Gilliland-Swetland, 2000.) This is the case in CUBER, where the system helps with the help of metadata the users to find the best matches for their needs.

3.2 Educational metadata standardization activities

According to Duval (2001) there have been many kinds of activities going on in the field of metadata that are specialized in education and pedagogy. These activities and projects are located all over the world and using existing standards offers possibilities for enhancing interoperability and reuse of data between different information systems, but the standards developed for different projects may not suffice because they may lack the information fields necessary for some specific information system. Standards are needed in educational context especially due to the diversity of the domain. In this context the standards offer a basic information infrastructure into which local solutions can be integrated. All the major educational metadata initiatives are linked together through different levels of collaboration. (Duval, 2001.)

Numerous metadata standardization activities were investigated as a part of this study. In the following chapters four educational metadata standardization activities are presented. The Dublin Core metadata element set was originally intended to facilitate discovery of electronic resources, but at the moment there are on-going activities that aim to add educational elements to DC, too. (Dublin Core, 2004; Lamminaho, 2000.) IEEE Learning Objects Metadata (LOM) Working Group has been developing a metadata standard for the specific purposes of education, especially Web-based education. (LOM, 2004; Lamminaho, 2000.) . IMS (Instructional Management Systems Project) has been a major contributor to the LOM standard developed by IEEE's LTCS working group. (IMS, 2004; Lamminaho, 2000.) ARIADNE metadata developed as a part of ARIADNE project is compatible with the IEEE's LOM standard. (VanDurm, Duval, Verhoeven, Cardinaels & Olivié, 2001; ARIADNE, 2004; Lamminaho, 2000.)

All the above-mentioned metadata initiatives were thoroughly investigated when defining the metadata schema in this study, but none of them were found sufficient in expression for the purposes of the CUBER system. However, the LOM standard was chosen in early phase of the study to serve as the backbone to the metadata to be developed, and it formed the framework of this study. The fundamental requirement for the metadata to be defined was to be interoperable and compatible with the LOM standard. The challenge of this study was to fulfill the partly contradictory requirements: to be conforming to a standard and to meet the specific needs of the services of CUBER.

3.2.1 Dublin Core

The Dublin Core Metadata Initiative (DCMI) is an open forum for developing *"interoperable online metadata standards that support a broad range of purposes and business models"*. DCMI's metadata development activities include workgroups, workshops, conferences, standards liaison, and educational activities in order to promote the acceptance of the metadata standards and practices of Dublin Core. DCMI aims to facilitate finding resources from the Internet by developing metadata standards that support the discovery of resources, defining frameworks for the interoperation between different metadata specifications, and by facilitating the development of community- or discipline -specific metadata specifications. (DCMI, 2004.)

The Dublin Core metadata element set is intended to facilitate discovery of electronic resources, especially from the World Wide Web. Dublin Core is a widely known metadata standard that has been developed since 1995 by a series of workshops. The Dublin Core standard contains 15 metadata elements that describe the content, the intellectual property rights and the instantiation of the object. Even though Dublin Core does not contain any educational metadata elements, it has been used as a basis for many educational metadata projects. At the moment there are on-going activities that aim to add educational elements to DC, too. (DCMI, 2004; Lamminaho, 2000.)

The Dublin Core metadata is a standard for *"cross-domain information resource description"*. The terms information resource refers to anything that has identity. Dublin Core metadata can be assigned to any kinds of resources. The Dublin Core metadata elements are: Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights. (DCMI, 2004.) These elements are very generic and they can be used for describing many kinds of information resources. These elements become meaningful when attached to a course or any other learning object. However, they do not include any domain specific or process related elements that are needed for example in the field of virtual university where the actual contexts and activities related to learning and teaching have to be described.

3.2.2 Learning Object Metadata (LOM)

IEEE Learning Objects Metadata (LOM) Working Group has been developing a metadata standard for the specific purposes of education, especially Web-based education. LOM describes the educational learning objects with a hierarchical metadata structure that is grouped into nine top-level categories for describing the pedagogical characteristics of the learning objects. Within these categories more detailed descriptions are provided further down in the hierarchy. LOM aims to define the minimal set of metadata elements needed in managing and locating the learning objects on the web. This minimal set of metadata elements may be locally extended, which will be later explained in this thesis. (LOM, 2004; Lamminaho, 2000.)

The LOM standard originates from the ARIADNE and IMS projects, and also build on the metadata work of the Dublin Core Metadata Initiative. Consequently, the LOM standard has been mapped with the Dublin Core standard. The LOM standard specifies Learning Object Metadata (LOM) with the help of a conceptual data schema that defines the structure of a metadata instance for a learning object. LOM defines a learning object as an *"entity -digital or non-digital- that may be used for learning, education or training"*. In LOM a metadata instance describes the relevant characteristics of the learning object to which it attached. These characteristics have been grouped into nine categories that are: general, life cycle, meta-metadata, educational, technical, educational, rights, relation, annotation, and classification. (LOM, 2004.)

The intended context of use for LOM may be, for example, technology supported learning including computer-based or -aided training and instruction systems, interactive and collaborative learning environments, and distance learning systems. The learning objects for their part may be, for example, multimedia or instructional content, learning objectives, instructional software and software tools, as well as persons, organizations, or events related to technology-supported learning. LOM has tried to focus on a minimal set of attributes that are required for the learning objects to be managed. The standard includes the possibility to extend the metadata elements and types locally, and the elements can be obligatory or optional. The most relevant attributes for describing the learning objects are: type of object, author, owner, terms of distribution, and format. LOM also includes pedagogical attributes such as: teaching or interaction style, grade level, mastery level, and prerequisites. (LOM, 2004.)

LOM aims to enable the learners and teachers to find, exchange and re-use easily the learning objects with the help of technology supported learning systems. LOM also aims to offer the educational institutions a standard that would be easy to accept in wide use. One of the LOM's goals has also been offering the possibility to automatically and dynamically create individual learning entities composed of various learning objects. Moreover, LOM aims to be a simple standard that can be extended in order to serve the needs of different application domains. (LOM, 2004.) The metadata work reported in this thesis is based on the extendibility of the LOM standard.

3.2.3 IMS Global Learning Consortium's Meta-data Specifications

IMS (Instructional Management Systems Project) is a global consortium of several educational institutions, commercial entities, government agencies, and developers in the area of distance education. Among the aims of IMS is to develop and to promote open specifications for facilitating online distributed learning activities such as locating and using educational content, tracking learner progress, reporting learner performance, and exchanging student records between administrative systems. IMS is a major contributor to the LOM standard developed by IEEE's LTCS working group, and IMS has introduced the use of XML for representing the metadata. IMS also uses the LOM standard as its basis for its current metadata work. IMS has also contributed to LOM by introducing best practice guides for metadata developers and implementers. (IMS, 2004; Lamminaho, 2000.)

IMS has defined and published its own metadata specification that is also very closely related to the LOM standard. IMS has been a major contribitor to LOM, but in addition it has published extensions to the standard. According to IMS, the LOM standard may not be able to capture all metadata elements or dimensions needed to describe learning objects and their use. However, LOM allows creating extensions for the metadata elements and structures. Moreover, the IMS Learning Resource Meta-Data XML Binding Specification defines ways for treating all user-defined, proprietary extensions in a uniform manner. (IMS, 2004.)

3.2.4 ARIADNE's Work in Educational Metadata

The ARIADNE Educational Metadata Recommendation is based on work performed by several European and international organisations since 1995. The metadata work of ARIADNE started with two European ARIADNE Projects, and after 2000 the ARIADNE Foundation has continued this work. ARIADNE has contributed significantly to the educational metadata standardisation work, as it introduced an early version of its metadata specification in 1998 to the IEEE LTSC Workgroup developing the Learning Object Metadata (LOM). Together with a similar metadata specification submitted by the IMS Project, the early ARIADNE metadata specification served as the starting point for the emerging LOM standard. The current ARIADNE Educational Metadata Recommendation is an application profile of the LOM standard. This means that the recommendation is fully compatible with the LOM standard, and instantiates LOM for the ARIADNE community. It has taken into account the specific needs and requirements in the fields of European higher education and continuing professional training. It is designed to support cultural and linguistic diversity and sharing and reusing of knowledge resources between the educational institutions. (ARIADNE, 2004.)

The ARIADNE metadata recommendation includes several metadata elements, and they can grouped according to the following categorisation:

1. general information on the resource itself
2. semantics of the resource
3. pedagogical attributes
4. technical characteristics
5. conditions for use

6. meta-metadata.

The work of ARIADNE has been supported by the European Commission in the Information Society Technologies (IST) Programme. The primary goal of ARIADNE was to enable and enhance sharing and reusing of electronic pedagogical material by universities as well as other organisations providing education and training. ARIADNE has developed a Knowledge Pool System (KPS), which is a Europe-wide distributed repository for pedagogical documents, with associated indexation and query tools. As a central part of the KPS is the metadata ARIADNE has been defining in collaboration with LOM and the IMS project. The ARIADNE metadata is strictly focused on describing the educational documents, not the processes or outcomes related to education. (ARIADNE, 2004.)

ARIADNE metadata is fully compatible with the IEEE's LOM standard (VanDurm, Duval, Verhoeven, Cardinaels & Olivié, 2001). The ARIADNE project has also used LOM as the basis of its current metadata work. ARIADNE has added some more educational metadata element to the original LOM schema, such as educational objective, pre-requisite, and pedagogical classification. (ARIADNE, 2004; Lamminaho, 2000)

3.3 Creating an application profile of a metadata standard for a specific domain

Dublin Core Metadata Initiative (DCMI) and IEEE's Learning Object Metadata Standard (LOM) have formulated shared principles for the creation of new metadata sets (Duval et al., 2002). According to Duval et al., metadata should be modular in order to enable different kinds of combinations. *Modularity* enables the re-usability and flexibility of a metadata set. In addition, metadata sets should allow for *extensibility*, i.e. it should be possible to add additional elements to the base schema for the local and domain specific needs without compromising the interoperability provided by the base schema. As a solution *application profiles* may be created in order to construct domain specific metadata sets that have selected elements from one or more metadata schemas and combined them into a one coherent metadata schema. This expresses the principles of modularity and extensibility required from a good metadata set. By using application profiles it is possible to create metadata sets tailored to the requirements of the particular application/system. (Duval et al., 2002.)

The use of application profiles enables the communities of practice to concentrate on defining a community-specific, detailed and complex metadata schema whilst preserving the interoperability provided by the conformity to the standardized base schema. According to Duval et al., metadata schemas may for example define rules of *cardinality*, i.e. define which elements should be mandatory, optional or conditional. Also the *value space restrictions* may be even more restricted in the application profile than in the standardized base scheme in order to meet the application specific requirements. Another means to increase the precision of the description and to minimize the misunderstandings is using pre-defined *controlled vocabularies*, where commonly agreed semantics are necessary. (Duval et al., 2002.)

According to the above principles, the metadata schema produced in the framework of this study and the CUBER project has been designed to be an application profile of LOM standard. The goal of the metadata work was to fill the gaps identified in the LOM standard by finding out the users' information needs related to finding education from the CUBER system and to provide metadata elements capable of describing these information needs.

The need for application profiles has been recognized in the research and development projects in the field of e-learning. During the UNIVERSAL project the metadata model of LOM was examined and evaluated. The metadata elements Educational objective and Method of instruction were found important but these elements were not available in LOM. Moreover, some LOM elements such as 'semantic density' and 'interactivity type and level' were found irrelevant among the users of the metadata system. Detailed examination of LOM revealed that in many places the vocabularies were not rich enough to be able to describe sufficiently the educational activities and some learning resource types. (Simon, 2001)

According to Simon (2001) a sufficient metadata model for describing the educational activities must include means to describe the roles for teaching contributors, more learning resource types, educational objectives, instructional design, and communication media used in education. At the moment LOM lacks the elements for describing any pedagogical or instructional models that are needed both by learners and teachers (Heinrich & Anders, 2003). Moreover, when trying to move towards increased semantic interoperability, semantic modeling of the learning domain is needed. More precisely, a domain ontology is needed as a basis of the information systems. (Simon, 2001.) In this study and in the CUBER metadata all these issues have been addressed in some form, e.g. through metadata extension elements or alternative vocabularies.

3.4 The need for an ontology

In order to transfer data seamlessly and transparently in the virtual university, there has to be a standard way for both people and computers to communicate all necessary knowledge, with both people and computer systems (Stojanovic, Staab & Studer, 2001). The objective of our research was to enable this communication by providing a conceptual model and relevant *ontology* for the purpose of building an information exchange infrastructure for a virtual university.

Furthermore, the need for common semantics and vocabularies has become evident during many multi-cultural research and development projects. The concepts and vocabularies used in the different countries varied significantly. The idea of using an ontology is to give semantics and a limited amount of possible interpretations for the vocabularies used. This is of utmost importance in order to avoid misinterpretations and misuse of the metadata. The ontology not only defines the central concepts of the metadata but also presents the relations between the concepts. In general, ontologies provide a shared and common understanding of a domain (Noy & McGuinness, 2001). In this sense the metadata model in this study incorporates the ontology.

In philosophy, ontology is a theory about the nature of existence, and ontology as a discipline studies such theories (Berners-Lee, Hendler & Lassila, 2001). However, in computer science an ontology means *a set of formally specified metadata structures consisting of commonly agreed concepts that bear a limited sense of meaning with them* (Jokela & Turpeinen, 2000). The most commonly cited definition in the literature is that of Gruber (1993), according to which an ontology is "*an explicit specification of a domain conceptualization*". However, this definition is argued to be insufficient and vague, to tell more about the state of affairs than the actual conceptualization (Guarino, 1997). In this view, the different ontologies for metadata fields are thought to be just *dimensions* of one big ontology, rather than numerous separate ontologies.

Ontologies are *formal, explicit, and shared* specifications of some conceptualizations. Formal means that the ontology should be machine readable, and explicit refers to having defined the types of concepts and the constraints on their use are explicitly defined. Shared refers to the fact that an ontology must reach a consensus. *A meta-ontology* is used to define the structure of the ontology itself. (Fensel, 2001.)

The goal of ontology building is to simplify reality while retaining fidelity to it (Weinstein, 1998). An ontology can also contain inference rules and logic (Hendler, 2001). It enables both people and computers to understand things, because all terms have been explicitly defined and assumptions clearly written down. In this research, the focus is not on ontology theory. Rather, the viewpoint is the one of applying ontologies and conceptual modeling in real-life problems. This way, it was hoped to be able to acquire valuable information on the usefulness and relevance of the information model.

Ontologies together with metadata enhance efficient access to information by offering possibilities to organize and categorize the content of the information system in question. In this context an ontology is defined as a means to formalize and to specify a common terminology for a defined area of interest. (Turpeinen, 2000.)

A *domain ontology* is used in order to model the semantics of the information content in a specified area. An ontology can incorporate several dimensions according to which the content can be organized. The dimensions can be, for example, the subject, author or the publisher of the target object. The dimensions of an ontology are often predefined by standards such as MARC that is designed for libraries and Dublin Core that is developed for various network resources. (Turpeinen, 2000.)

A domain ontology defines the concepts and the relationships between them in the area of interest in question. Conceptual models can be derived from these concepts and their relationships. In the conceptual model, the information objects of the content area are classified under different dimensions. *Dimensions* are used to describe the aspects of the domain of which the content is organized. In some cases the dimensions may be independent of each other, but usually they are closely interrelated, which complicates the process of conceptual modeling. *Concept* has been defined as a category, mental entity, or a generalized idea. *Categorizing* means grouping and dividing the concepts or information objects into categories according to similarities and differences perceived. Conceptual models can be used in this process for information classification and organization, which in turn improves information retrieval. (Turpeinen, 2000.)

There are some ontology languages that are used to describe and formally the define the ontologies. OIL and DAML+OIL are the currently most known ontology languages for the semantic web, whereas XML and RDF are also technologies used in ontology development too. However, their role is to serve as a basis for using ontology languages in the layered architecture of the semantic web. (Fensel, van Harmelen & Horrocks, 2003; Mohan & Brooks, 2003.) XML as a meta-language describing other languages or resources can be used as an exchange format for relational database systems. In addition, XML facilitates the generation and management of metadata and it is capable of storing metadata. XML is used with DTDs (Document Type Definitions) that define the allowed elements. (Bradley, 2000.) However, XML does not have a semantic model and DTD is not an appropriate semantic language. Thus RDF and ontologies are needed for describing the learning objects in the web.

(Stojanovic, Staab & Studer, 2001) A more simple solution is, for example, UML (Unified Modeling Language) that is used as a tool for defining semantic interrelations when building conceptual models besides for defining and modeling information systems. (Turpeinen, 2000) In this study UML was used for conceptual modeling and representing the ontology and its dimensions. Additionally, XML with a DTD was defined for information exchange between different systems CUBER anticipated to co-operate with.

3.5 Conceptual modeling of a domain

Generally, building an information system starts with an analysis of the domain, its features and conceptual aspects. The observations will then be turned into a formal (or semi-formal) model that includes the actors, variations of information and the relations between the actors and the information. According to Sowa (2000) the model must include the functions that represent the real world functions in order to make the system usage possible. However, the reality is always more complex and more diverse than the models representing it. Thus the model can only be an approximation of the reality, a focused and interpreted view based on the intended use of the system. (Sowa, 2000.) In order to further define the metadata requirements, the world of higher education was studied and models representing it were created. This included surveys, interviews, and brainstorming sessions within the multinational CUBER project consortium. As the result, detailed requirements were defined for the metadata set.

Sowa defines *conceptual analysis* as a task of investigating and analyzing the concepts that are expressed in some natural language in order to make explicit the otherwise implicit relationships between these concepts. Conceptual modeling is a central task for e.g. database modeling, knowledge engineering and systems analysis. Ontology development can be seen as a part of the process of conceptual modeling and as an integral part of application development. Sowa also suggests that all necessary information of the application domain should be gathered together into a conceptual schema that serves as an integrator between the system's database, applications, and the user interface. (Sowa, 2000.)

According to Männistö (2000) a conceptual model is needed to offer the semantic terms that are needed in the process of modeling an application. Two components are recognized as important for conceptual modeling: *static and dynamic ontology* (Männistö, 2000). A domain ontology can incorporate a set of conceptual models that may also be called metadata structures. They represent the content domain, and they consist of the domain concepts and their relationships. Thus conceptual modeling and ontology development can be seen as a prerequisite for using semantic metadata. (Jokela, 2001.)

3.6 The Semantic Web as a platform for e-learning

Semantic web is defined as an extension of the current Internet in which the documents are annotated with metadata. This metadata is used to define the semantics of the documents in a machine-understandable way, which improves significantly the services provided by the web. Using ontologies in semantic web is needed in order to be able to cope with the various web resources. The ontologies provide a shared understanding of a specified domain, and this understanding is possible to communicate both between people and computer systems. The basic technologies for the semantic web include XML, RDF and ontologies. (Hyvönen, 2001; Davies, Fensel & van Harmelen, 2003; Mohan & Brooks, 2003.)

The semantic web seems as a promising technological basis for implementing e-learning. Especially using ontologies is a characteristic that enables using semantically enriched learning objects. This is a prerequisite for creating customized, on-demand on-line learning services. This encourages learner centered learning as the learning materials and other resources are available for the learners according to their needs. (Stojanovic, Staab & Studer, 2001.) In this research influences were taken from the semantic web research, e.g. the use of ontologies whilst some other elements of semantic web were abandoned.

3.7 The need for user centered metadata model for the virtual university

The need for developing user-centered metadata for the virtual university can be recognised both from existing research results and the practice. The current metadata models and standards offer either only general metadata elements or educational metadata that does not meet the needs of the users. For example, the description of the learning process and the teaching activities of the virtual universities are to a great extent missing from the existing models. However, the LOM standard offers a good basis for developing an application profile, i.e. a metadata schema that is based on LOM but has been extended in order to meet the specific user needs and requirements of a particular community or domain. The development of such application profile of the LOM metadata schema has to be user centered so that the usability, understandability, and expression of the metadata model can be ensured.

This thesis builds on the metadata development work done in the CUBER project that resulted in an application profile of the LOM standard for the domain of virtual university. This application profile was developed on the basis of a user study and in close collaboration with the intended users of the metadata and the CUBER system. In addition to the CUBER's metadata work, this thesis aims at introducing user centeredness for the development of metadata model.

4 Research questions, methodology and implementation of the study

The research approach of this study is mainly constructive. The goal of *constructive research* approach is to construct reality on the basis of existing knowledge and at the same time to decide what this new reality to be constructed should be like (Järvinen & Järvinen, 1996).
The goals of constructive research include developing and building something new in a concrete manner, and forming a better understanding of the phenomena explored. In other words, constructive research aims to create functioning artifacts that serve some predefined ends, and to at the same time create new and apply existing information and knowledge

needed in producing these artifacts. The results of this kind of a research include constructs, models, prototypes, and methods. Basically, this research approach consists of two phases: building and evaluating. The first phase refers to constructing the artifact, and the latter one refers to assessing how well the artifact meets its requirements and expectations. (Järvinen & Järvinen, 1996.) Moreover, constructive research includes a contribution to the existing knowledge or theory, and during the research process there should occur learning based on the experiences. (Metodix, 2004.) In this study, both of the building and evaluating phases were realised. First, as a part of the CUBER system development, the CUBER information model incorporating both metadata and ontology was specified and implemented, and second, the end users evaluated the acceptability of the model. Furthermore, both existing information was examined and new understanding was developed during this research.

Due to the constructive nature of the research approach, the research questions in this thesis are relatively practical, even though the existing knowledge from literature has been investigated and taken into account. The research questions were defined in order to support producing practical solutions to the real-life problems of the study. These research questions including some rather general research questions served as guidelines for the study and the answers and results contribute and are closely related to the goals of the research project's metadata specific part, i.e. developing metadata and ontology for an innovative educational information broker system for the purposes of virtual universities. The research questions that are grouped under three rubrics are the following:

1. Using metadata in the context of virtual university:
    a. What kinds of user requirements are there for the metadata?
    b. Are the current and emerging educational metadata standards sufficient for virtual learning environments?
    c. What kind of metadata is needed in virtual university?
    d. How should metadata be structured in a virtual university application?

2. Using ontology in the context of virtual university:
    a. What kind of ontology is needed in virtual university?
    b. What kinds of requirements are there for ontology?
    c. How should ontology be structured in a virtual university application?
3. Conceptual modeling and information representation in the context of virtual university:
    a. How can metadata be modeled and represented?
    b. How can an ontology be modeled and represented?
    c. Is the information model developed in this study understood and accepted by its intended end users?

The methodology used in this research is based on the constructive tradition and it emphasizes the importance of producing functioning solutions to real life problems, such as functional information systems. In this study the multidisciplinary approach is reflected also in the methods used in this research: methods from several disciplines (e.g. education science, software engineering, computer science) could be used in parallel. The research methodology was a combination from these disciplines' methods, including different methods in different phases of the study, e.g. semi-structured interviewing and creating SQL definitions.

In the first phase of the study the requirements for the metadata were defined. In this phase surveys, interviews and brainstorming sessions were used. In the second phase the different metadata standardization activities were investigated by conducting a literature survey, and comparative content analysis was carried out for the metadata schemas. In the third phase, the CUBER metadata schema and ontology was defined in an iterative manner including several cycles of development, reviewing, and refining by the project partner organisations. This was the constructive phase of this study, targeted at developing a significant part of the CUBER system. Thereafter, the CUBER system was implemented and integrated, and the user interfaces were built. These are not considered as central phases for this thesis, but anyhow significant as implementing the results of the study. As the final phase of this study the metadata and conceptual model were evaluated by the intended end users of the system, i.e. the representatives of the educational organisations involved in the CUBER project. This evaluation was performed with the help of a web-questionnaire and a thematic interview, the results of which were analyzed qualitatively. However, it has to be noticed that only one of the intended user groups (i.e. the providers of the learning objects, not the learners) participated in the evaluation, which naturally affects the results. The research methods related to each research question are presented in Table 1.

Table 1. Summary of the research questions and methods

The research questions	The research methods used
1. Using metadata in the context of virtual university:	
a. What kinds of user requirements are there for the metadata?	User requirement study including a questionnaire, interviews and brainstorming
b. What kind of metadata is needed in virtual university?	Literature survey, iterative brainstorming and constructive research (developing the metadata schema)
c. Are the current and emerging educational metadata standards sufficient for virtual learning environments?	Literature survey and comparative content analysis
d. How should metadata be structured in a virtual university application?	Analysis of the existing standards, iterative constructive research (metadata schema development)
2. Using ontology in the context of virtual university:	
a. What kind of ontology is needed in virtual university?	Literature survey, further analysis of the metadata requirements, brainstorming sessions
b. What kinds of requirements are there for ontology?	Literature survey, further analysis of the metadata requirements, brainstorming sessions
c. How should ontology be structured in a virtual university application?	Literature survey, iterative constructive research (developing the ontology)
3. Conceptual modeling and information representation in the context of virtual university:	
a. What is a feasible way to model and represent metadata?	Literature survey, iterative constructive research (metadata development, use of UML diagrams and SQL definitions)
b. What is a feasible way to model and represent an ontology?	Literature survey, iterative constructive research (ontology development, use of UML diagrams and SQL definitions)
c. Is the information model developed in this study understood and accepted by its intended end users?	Questionnaire and thematic interviews with the end users, qualitative content analysis, scenario-based evaluation/ analysis

5 Results of the study

The main result of this study is the metadata schema incorporating a domain ontology for European higher education and virtual university context. The results reported in this thesis were accomplished as a part of the CUBER project, especially a workpackage focused on the metadata development work.

As an initial phase of the metadata development work, a user requirements study was conducted. Based on these requirements, the CUBER metadata was developed as an application profile of the LOM standard, which means that the standard was modified according to the special needs of the application domain. In addition, a domain ontology was created in order to enable seamless communication of educational information both for humans and computers. The metadata and ontology were developed with the help of an iterative collaboration process, in which the end users were involved to considerable extent. Next, a conceptual model and a meta-ontology were created to illustrate the relationships and structures of the concepts central to the field of virtual university and e-learning. The implementation of the metadata and ontology of CUBER as a relational database was supported by creating a database definition by using standard SQL. Finally, the information model of CUBER was evaluated by the end users of the CUBER system. The following chapters present the results of the study, and the Appendices 1-3 including the metadata schema, example of the conceptual model in UML, and example of the SQL definition are added in this thesis in order to illustrate the results in a concrete manner.

5.1 The Metadata Requirements in CUBER

In order to meet the needs of the users a system must provide certain functions. A *user need* is defined as a task or a goal the user wants to perform or accomplish with the help of a product (Kotonya & Sommerville, 1998). A *requirement* can be defined as a function, constraint or other property that the system must have in order to satisfy the user's needs (Faulk, 1997). The requirements of a product to be developed usually include technical requirements related to the system and the user requirements concerned with the users' tasks, needs and expectations (Kauppinen, 2002). In order to capture the requirements from the users' point of view, a user-centered analysis of requirements may be used. (Rumbaugh, 1994) It is recommended that information be gathered on the current practices and behaviour of the end-users in order to gain insight into future needs, because the problems and needs of today are likely to continue existing (Patnaik & Beeker, 1999). User requirements are used to create the system concept and specification; they also are used as evaluation criteria in the later phases of the product development process (Mäkelä & Battarbee, 1999). Requirements engineering presents a systematic way for defining, managing and testing the requirements of a system (Kauppinen, Kujala, Aaltio & Lehtola, 2002).

At the beginning of the CUBER project, a user needs survey was carried out in order to gather information about both the general requirements related to the user interface of the CUBER system and the more detailed information needs related to the metadata descriptions of the learning objects to be stored in the CUBER system. This survey was focused on the second potential end user group of CUBER, namely the students searching for learning objects in the Internet. The metadata related needs and requirements of the other end user group, the course providers, were investigated separately. Altogether 100 persons (most of them German) participated in this survey. (Wiendieck, Preisendanz & Filla, 2000.) As this survey was not actually a part of the metadata study presented in thesis, only the results of the survey will be summarised concisely. However, the results of the survey were used as an input for the metadata development work carried out in the CUBER project, which is the subject of this thesis.

The goal of the students searching for learning objects in the Internet is to find easily the best matches according to their specific criteria. These criteria are (Wiendieck et al., 2000):

- Learning objectives
- Stage of education
- Short description
- Summary of the content of the learning object

- Learning method
- Prerequisites to enrol
- Information about examinations
- Workload required
- Tutor or teacher
- Difficulty level
- Title (of the learning object)
- Addition with the topic of learning objects
- Authors of the learning objects
- Starting date and enrolment date
- Availability of the learning object
- Fees related to studying
- Support to prepare a study plan
- Positions and relationships of the learning objects
- Contact persons
- Information on literature and other learning material
- Links to more detailed information
- Institution offering the learning objects
- Duration of the learning object
- Degrees related to the learning objects

The above criteria can be seen as requirements for the metadata in the sense that all this information should be able express with the metadata elements. In addition to these rather direct metadata requirements, the survey highlighted the need to describe different learning objects in different ways. For example, descriptions of the study courses and study programmes should be distinguished from each other by offering different metadata elements for the learning objects according to their level of abstraction. (Wiendieck et al., 2000.) This differentiation was implemented in the CUBER metadata work by offering different metadata elements for each four aggregation levels (see chapter 5.4 for details).

At the beginning of the CUBER metadata development the metadata requirements of the second end user group, i.e. the course providers of CUBER were investigated. In this study the term user requirement refers to the features, structures, elements, and contents of the metadata to be used in virtual university for *describing* the learning objects. Because the users were from several European countries, a questionnaire for the requirements study was decided to be delivered. The users in this study are the representatives of university personnel who were expected to produce the metadata descriptions of the learning resources offered by their organizations. The questionnaire consisted of open-ended questions that were further clarified with more narrow sub-questions. (Lamminaho, 2000.) The goal of this metadata user requirements study was to analyse the widely varying terms and concepts used in European higher education. This was necessary so that a consensus on the meaning of these terms and concepts could be reached. Moreover, the structure and representation of the metadata had to be analysed.

The questionnaire was delivered by email to the participants. The gathered data was analysed with qualitative content analysis, and the responses from different countries and organizations were compared with the help of tables. (Lamminaho, 2000.) As the result of the study, requirements and recommendations for the metadata development work were given. One of the main findings was that the requirements and needs related to metadata did not completely match with the LOM standard; thus the standard needed to be extended. Secondly, the study revealed the differences between the partners with regard to the use of educational concepts and terms; a common agreement had to be reached on the central concepts in order to be able to co-operate efficiently. Language, culture and educational practices had to be taken seriously into account when developing the metadata set in CUBER.

The concepts and terms included in the questionnaire were chosen in advance, and the respondents were asked to define the meaning of these terms according to their understanding. The concepts and terms included e.g. the following:

- Web-based course
- Course registration procedures
- The procedure of entering the course pages
- Study process
- Course completion
- Course material
- Guidance related to studying
- Discipline
- Target audience
- Educational level.

In order to gather information on the course provider organisations' needs and requirements related to metadata, a questionnaire with open-ended questions was developed. It was distributed via email to the CUBER project partner organisations, and five organisations out of eight answered the questionnaire. (Lamminaho, 2000.)

The questionnaire included questions of the following issues (Lamminaho, 2000):

- What is the meaning of the term "web-based course" in the participating countries?
- How the studying process of a course can be described?
- How do the students pass the courses?
- What kinds of learning materials are attached to the courses?
- How is Information Technology perceived and understood as a subject area of studies?
- What kinds of granularity levels are there in use for the units of study?
- What kinds of prerequisites are there for the courses?
- Are the courses organised by only one or several educational institutions?
- What are the prices of a course in the participating countries?

The main results can be summarised as follows (Lamminaho, 2000):

- The concept "web-based course" is understood in many ways, varying from static self-study material to interactive virtual learning environment with supporting activities for the learners.

- In order to participate in a course the students need to register and to enrol to the course. For this purpose they need an ID and a password.
- The studying process can be characterised by several factors, such as teaching method, interactivity level, methods for examination and assessment, and course completion. The courses may be taught in traditional face-to-face classrooms, in the Internet, or by mixing these methods. Consequently, the exams can be taken in the Internet in some cases. Moreover, there are several - even alternative - ways to complete the course, such as projects work or an examination. Several assignments or attendance to classes may be required.
- Degrees or certificates may be given for completed courses or other studying units but these vary significantly. The credits given in each university are not comparable. ECTS may bring a solution.
- It is important to be able to indicate the number of students to be admitted to the course. Also the duration of the course can vary from days to years depending on the organisation.
- The learning materials related to studying have to be defined in conjunction with the learning objects. Most used materials were books, CD-ROMs, audio/videotapes, PDF-documents on the Web, and videoconferencing activities. Links from courses to the material are useful.
- Teachers, tutors or other personnel can provide guidance for the students in various ways.
- The prerequisites of a course can be related both to the skills or previous studies of the student or to the technical prerequisites, all of which are dependent on the course in question.
- The price of the studying varies from free of charge to several thousands of euros. The fees should indicate the total price of the course or other study unit.
- The courses or other study units may be organised by one or several course provider institutions.

- In addition to using the term “course”, there is a need to use terms that are either smaller or wider than it. Courses may be compiled of smaller modules, and on the other hand they may build up wider entities, such as study packages, programmes or degrees.
- The disciplines describe the field of science or research the course is related to. These categorisations vary widely in the participating countries.

This study yielded only general view on the requirements related to the metadata. For this reason, the requirements study became a continuous process between the project partners in CUBER. Consequently, the metadata development became an iterative process in which metadata specification version development and requirements or feedback elicitation followed each other. In latter phases when most of the requirements had been implemented, the partners still participated in the metadata development by providing valuable comments and improvement items. This way, the metadata development of CUBER can be seen as a participatory design process (Carroll, Chin, Rosson & Neale, 2000), due to the fact that the users of the CUBER system and metadata were involved throughout the development process.

In addition to the requirements study the goals of the CUBER system were analysed. There is an abundance of courses available in the Internet, but the information on the courses is still very unstructured and spread out. This was one of the problems the CUBER system aimed to solve by using a well structured and standardized metadata combined with a powerful search engine and knowledge base that contains also a lexical database for defining the central terms in the field. The purpose of using metadata in the CUBER system was to facilitate the learners’ searching, comparing and selecting of study elements offered by the European universities and other (online) educational institutions. The goals of the CUBER system have had direct consequences for the metadata. It was possible to deduce most of the metadata elements needed by exploring the goals and aims of CUBER. The preceding goals were considered in depth when specifying the metadata. Thus the metadata serves the CUBER system by making many of its services and functionalities possible. Metadata is very meaningful for both the learners and the course providers because it enables matching the learners' needs to the course supply of the universities.

The original goals of the CUBER system were, in short:

- Enabling the comparison of learning resources from different providers and to find the best matches according to learners' individual educational goals
- Taking learner's previous skills and knowledge and qualification objectives into account
- Taking various target groups and the learner's specific constraints into account
- Providing information on how the courses are integrated together
- Making it possible to generate a complete curriculum plan according to individual preferences
- Providing information on whether the studies lead to some kind of degree or certificate and
- Providing information whether there is a possibility for inter-university recognition and international approval of the degrees or certificates gained in the virtual university
- Facilitating the search functionalities according to learners' personal preferences and commonly understood keywords
- Enabling the learner to create an own user profile and to save it for later searches in the CUBER system.

Even though these goals were defined before the user requirements analysis, it is possible to identify a strong user centered approach, but during the CUBER project it was not fully recognised. By analysing the CUBER system it was possible to identify three different **stakeholder or user groups** that have different tasks and requirements related to the system and the metadata.

- CUBER system developers: This stakeholder group is responsible for the system development and maintenance.
- Learners: This user group actually uses the CUBER system for searching the learning objects.

- Course providers: This user group uses the CUBER system and the metadata for describing the educational supply of the educational organisations.

The learner and course provider groups can be seen as closely related in the sense that they actually use the CUBER system and the metadata model, but at the opposite ends of a continuum. The course providers use metadata and the system for describing the learning objects, and the learners use the system and metadata for searching the learning objects and for deciding whether the learning objects match their needs and preferences. The CUBER system developers for their part need the metadata schema for building the database. They are not interested in the contents or semantics of the metadata; instead they are interested in the structures and datatypes of the metadata schema. All these three user or stakeholder groups have been taken into account in the metadata development according to the retrospective user centered analysis of the metadata model. Even though the user approach was “hidden” in the goals of the CUBER project, the different viewpoints of these groups have been realised in the following way in the metadata schema (see Appendix 1):

- System developers: The metadata schema is structured in a consistent way that supports system development. All metadata elements have been defined by the same data elements: name, explanation, size, order, value space, data type, example, aggregation levels, and mandatory. By providing this information in the metadata schema the system developers were able to build the CUBER system.
- Learners and course providers: The metadata elements and especially the CUBER extensions have been chosen on the basis of the user groups’ need to communicate the learning object descriptions through the CUBER system and its metadata model. The metadata elements contain also explanations that facilitate the understanding of the meaning of the elements.

The usability of educational metadata is an important issue due to the fact that the systems perceived practicality in users’ view is the main factor affecting the adoption of a system by the users. Thus the metadata elements selected to the schema and the system must meet the requirements of the users. Currently the most common problems are the ambiguous field names, irrelevant metadata fields, and using too many fields in the system. (Monthienvichienchai, Sasse & Wheeldon, 2001.) In the CUBER system the usability of metadata was enhanced by carrying out a user requirements study before designing the metadata, by gathering continuously feedback from the users, conducting several usability tests for the metadata interface, and by studying the acceptability of the model when the final prototype was completed.

The CUBER system focused on educational resources supply, the study elements or learning objects. It was not designed to hold any student administration or organizational administration data. Instead, the system was merely concentrated on information considered relevant for the learner to decide on choosing a study element offered on the web. For this reason the metadata schema does not include the administrative metadata elements.

5.2 Overview of the CUBER Metadata model

As a result of the metadata definition work, the CUBER metadata model was created. The metadata model of CUBER consists of nine categories that contain metadata elements. The metadata categories do not carry any information. Instead, they function only as rubrics under which the related metadata elements are gathered. The metadata elements, for their part, can carry information as values. They can also include sub-elements that carry more detailed information about the study elements described. In this latter case, the metadata elements containing sub-elements cannot themselves have values directly. The sub-elements carry the values in question, and the metadata elements themselves can have their values only indirectly through their sub-elements. (LOM, 2004.)

The metadata schema is organized hierarchically in the form of a tree as can be seen in the Figure 1. The metadata categories are placed in the top of the hierarchy, and the metadata elements are right below them. One category can have several metadata elements under it. The metadata sub-elements are subordinated to the metadata elements, and one metadata element can have one or several sub-elements. (LOM, 2004.)

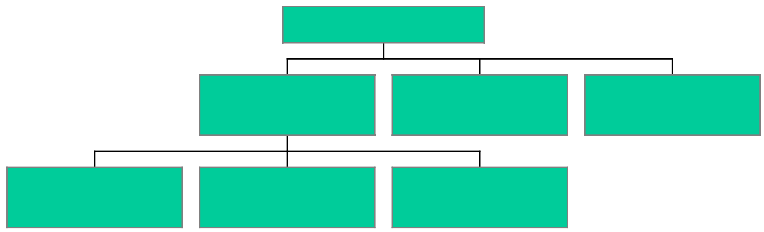


Figure 1. The structure of CUBER metadata

There are nine categories in LOM and they all are used in CUBER metadata. The metadata categories of LOM and CUBER are presented in Table 2. The complete CUBER metadata schema is attached to this thesis as Appendix 1.

Table 2. Categories of CUBER metadata

1. General	The general information that describes the study elements of CUBER as a whole.
2. Lifecycle	The features that describe the history and current state of the study element. Information about the contributors.
3. Meta-metadata	Information on this metadata record itself.
4. Technical	The technical characteristics and requirements for use of the study element.
5. Educational	The educational and pedagogical characteristics of the study element.
6. Rights	The intellectual property rights and the conditions of use for the study element.
7. Relation	Defines the relationships between given study element and other targeted study elements.
8. Annotation	Provides comments on the educational use of the study element. Information on the commentators and the comments themselves.
9. Classification	Describes where the study element falls within a particular classification system.

5.3 The CUBER Metadata Extensions

There are three kinds of *metadata elements* in CUBER metadata. 1) metadata elements adopted from LOM, 2) LOM metadata elements with altered vocabulary, and 3)

CUBER extension metadata elements. All *metadata categories* defined in the LOM standards are used in CUBER, but some of the metadata elements defined in the standard were considered to be of little or no relevance for the purposes of CUBER, and they were not used. They were not removed nor transformed due to the requirements of compatibility defined in LOM. Instead, these LOM elements are to be left empty in CUBER metadata when exporting or importing data from other systems compliant to the LOM standard.

The *CUBER extensions* to the LOM metadata were added as complementary elements to the base metadata schema in order to be able to describe all the characteristics of the CUBER study elements and to be able to provide the functions of the CUBER system. The extensions include also additions to the vocabularies of LOM. According to LOM (2004; Duval et al., 2002), the vocabularies can be extended, although it is strongly recommended to use the vocabularies defined in LOM. However, the metadata elements and their meanings and attributes must not be changed. There are three basic ways to adapt the metadata schema to meet the particular local needs: 1) using Classification systems in Category 9 Classification, 2) creating extensions to the metadata schema that do not override the original LOM elements, and 3) changing the vocabulary used in the LOM elements. (LOM, 2004.) All these methods were used in the metadata specification of CUBER. In this study the focus is on the metadata elements level.

In the CUBER extension elements information about the institutions that are providing the education was included. Elements for the important dates related to the studying and information on the possible recurrence of the study elements were also added to the schema. In order to facilitate internal identification of the study elements within the CUBER system, a “CUBER Identifier” element was introduced. The most significant difference from the standard was the decision not to use the aggregation levels defined in LOM. Instead, the CUBER metadata defined its own aggregations based on the metadata requirements analysis. This decision will be dealt with more elaborately later in this thesis. The metadata element “date” defined in LOM was not used in CUBER because of the need to express both the begin and end dates of the study elements such as study courses, which would not have been possible with the LOM standard. Moreover, the CUBER metadata introduced its own metadata elements for the providers and institutions of the study elements for the aggregation levels course, package and, programme in order to be able to keep the information on the persons and the institutions separate and independent.

The greatest number of changes was made within metadata category “Educational“, because of the special need of CUBER to describe the educational characteristics of the study elements. First of all, the vocabulary of metadata element “Educational.Context” was changed to meet CUBER’s requirements. The teaching activities, examination, ECTS-credits, study guidance, enrolment, and dedication to studying, pre-requisites, and related official degrees could be described with the additional metadata elements. In addition, the cost of the study element in Euro, and the possibilities for the student to obtain financial aid could be announced. The metadata category “Relation” was also extended by adding one element with which the users of the system could describe the dependencies among the study elements and the conditions under which they could be combined into a larger entity. This description would be given as free-text input, instead of using pre-defined relations, in an attempt to avoid creating too complicated an information system that would be unnecessarily troublesome to maintain and update.

5.4 Aggregation levels for the CUBER metadata model

On the basis of the metadata requirements analysis and several brainstorming sessions with the CUBER partners, it was decided to define four aggregation levels, viz. study material, study course, study package and study programme since they reflect the educational reality. This is a significant extension to the LOM schema, and it aims at following the requirements of the CUBER users. Study material was defined as the smallest unit in CUBER, followed in the hierarchy by study course. After study package, the largest unit in CUBER was the study programme. These aggregations were decided to be used in CUBER because *the* *study elements* to be described in CUBER differ from each other in their scope, teaching related activities, purposes, and other significant characteristics. The aggregation levels were used to describe the differences and the relationships between the study elements in CUBER. Aggregation levels were organized hierarchically, and they had both common and individual features. This classification of the aggregation levels intended to describe the existing classification system in higher education, though there were significant discrepancies among the European countries involved in CUBER.

The *descriptions* of the study elements attached on the different aggregation levels varied according to the individual level. The metadata schema offered a possibility to use free-text descriptions in some of its elements, and the level of abstraction in these descriptions should be compatible with the aggregation level of the study element. For example, the description of the contents of a course must be more detailed and concrete than the description of programme’s contents. The description should be most detailed in courses and materials. The level of generalization was intended to be intermediate in packages, whilst the programmes were to be described in general terms. The description of learning targets and outcomes, teaching methods, and examinations moves gradually from concrete and specific towards abstract and general according to the level of aggregation. The aggregation levels may also have some specific metadata elements with specific vocabulary that differs from those of the other aggregation levels.

Table 3. Definition of CUBER Aggregations

Word	Meaning
Material	The term study **material** refers to any kind of resource (digital or non-digital) that contains information used in education. Study **material** includes no teaching activity; its function is to serve as information source for a study course.
Course	The term study **course** refers to a complete unit of instruction that provides the learners with the knowledge or skills required for competence in a subject matter. A study **course** is any academic or vocational course arranged by a course provider. This is the lowest level that can offer credits or recognition within an educational institution. Study **course** usually includes teaching activity and examination.
Package	The term study **package** refers to a collection of study courses. Study packages can offer credits but no official degrees nor certificates, i.e. a study **package** has internally visible outcomes. A study **package** can be part of a study programme.
Programme	The term study **programme** refers to a collection of study courses and/or study packages, and it can lead to an official university degree or a certificate of competence, i.e. a study **programme** has externally visible outcomes.

The CUBER aggregation metadata element was introduced because the aggregations of LOM are different from the aggregations needed in CUBER. First, the level of abstraction was more detailed in LOM than in CUBER. LOM aggregations were much more atomic when compared to the aggregations of CUBER. Second, the content and context of CUBER was different from those of the aggregations of LOM, which would have been more suitable for describing, for example, HTML-documents (i.e. learning objects). Instead, CUBER needed to describe its learning objects or study elements in a much wider context that enabled the description of larger study elements, such as study courses, packages and programmes that were not included in LOM. This context of description also had to pay attention to the specific nature and characteristics (such as examinations, counseling etc.) of these study elements. For the above reasons CUBER metadata defined its own aggregation levels as a central conceptual model of the application domain. However, in order to ensure interoperability and compatibility with LOM, a mapping between the CUBER aggregations and LOM equivalents was created.

5.4.1 Metadata elements for the aggregation levels

Despite the fact that the study elements or learning objects of CUBER on different aggregation levels differ significantly from each other, they still have some common characteristics and common metadata elements. These elements have the same meaning on each level and they must be used coherently in order to avoid misinterpretations and dysfunction of the CUBER system. If a metadata element has a vocabulary, this vocabulary has to be used coherently on each aggregation level. The instructions for using the metadata elements with regard to the aggregation levels are defined in the CUBER metadata specification document's column "Aggregation levels" (see Appendix 1). The following paragraphs and chapters aim to introduce and describe the aggregation levels and their central characteristics on a general level only.

All the aggregation levels included in CUBER have a certain language. They all have a title, which can be both in the local language and in English. All the levels can be subject to charge, and the amount of the fee or payment is given in Euro. The aggregation levels may require some previous skills, knowledge or education, and CUBER will refer to these as prerequisites. The prerequisites can be academic achievements, work experience or specific skills, such as the ability to use certain software. Each aggregation level can have an intended target audience. In addition they have difficulty levels. Content is a common element too, for all levels contain some kind of information or activity and they all are located somewhere. Material can be found e.g. in libraries or on the Web. The courses, packages and programmes are usually located in the country of the provider, albeit they do not have to be located in a certain physical environment; they can be studied virtually as distance learning. All the study elements have providers, which can be academic or other educational institutions. Providers have names and contact information, and the providers can be persons and institutions. These study elements can be joined together by using the relations defined in the CUBER metadata specification. Relation *Has_part* defines the larger entity that the study element belongs to. Relation *Requires* defines the study elements that must have been completed successfully before entering this particular study element (or learning object). The different *versions* of the study elements can be expressed by using the metadata element 2.1 LifeCycle.Version. There is also a possibility to express the version of a study course by implementing the Course_Occasion class in the Class diagram of CUBER metadata.

Material

The term study material refers to any kind of digital or non-digital resource that contains information used in education. (LOM, 2004.) Material can be attached to particular a course and it can be reused in a number of courses. Material includes no teaching activity; its function is to serve as an aid or as a source of information in a course. Study material is the lowest level in the hierarchy, and there cannot be any levels or study elements below it.

Material has to have some kind of content but the content can vary according to the intended usage of the material. Also the type of material can vary. Material can consist of e.g. books, journals, articles, or HTML pages, downloadable files, CD-ROMs, or TV programmes, videos, and DVDs. As someone has created the material, it always has an author or many authors. The author or some organisation may own the copyrights in the material, and there can be conditions or limits for using the material, such as for teaching purposes only. Material has usually been created in some kind of organisation or academic institution and the author may have an official role in that organisation, such as professor or researcher. Material can be developed; i.e. it can be revised or re-edited. Thus several versions or editions of the original material can exist.

Course

The term study course refers to a complete unit of instruction that provides the learners with the knowledge or skills required for competence in a subject. A course may consist of lessons and tests together with associated learning objectives. Study course can be any academic course arranged by a course provider. Study course is the lowest level capable of offering credits or recognition. Study course is also the lowest level to offer teaching activity. Nevertheless, a course cannot offer degrees or certificates, but it can be part of a study programme that offers a degree or a certificate. In addition, a course can be part of a study package. It is should be noted that course can be an independent element; it does not have to belong to any entity. The time frame or the duration of a course can vary from days to months, but cannot exceed one semester or study year.

The learning targets are quite precise, and they can be described as specific skills or pieces of knowledge. The learning outcomes are rather narrow, which means that they must be described in a concrete and detailed manner. The learning outcomes are evaluated, usually through exams. One or more teaching methods can be used in a course, but they are more limited in number than e.g. in a programme. The teaching methods used can be described in detail or at least they can be announced more accurately than in packages or programmes.

The learner needs to enrol in a course, and he needs the enrolment dates and methods in order to be able to enrol. In addition to enrolment procedures, there may be limitations for entering the course, such as limitations to the number of participants. Furthermore, prerequisites can limit the target audience remarkably. Courses can be unique, i.e. courses can be arranged only once, and however they can also be repeated periodically. If the course is repeated, the version has to be announced. Each course belongs to some discipline or at least to some more general subject matter taught at universities. The learner has to study to complete a course successfully, and the study load can be indicated in terms of hours of work required of the learner.

Package

The term study package refers a collection of courses. There must be more than one course in a package. A study package can offer credits or other official recognition, but it is not capable of offering any degrees or certificates. Nevertheless, it can be part of a study programme, which does lead to a certain degree or a certificate. In other words, a study package can offer internally visible outcomes.

Since there are several courses in a package, the learning targets can include many kinds of specific skills and pieces of knowledge. Thus the learning targets must be described in a more abstract manner and the description has to refer to all the main issues of the courses included in the package in question. Accordingly, the learning outcomes are described on a more general level than in courses. The description of learning outcomes must mention the skills and knowledge gained on a general level or as classified into categories.

Packages can have prerequisites, just as courses can; i.e. the package requires some previous studies, experience or skills. The teaching methods can vary remarkably, because there can be many kind of courses in one package. Thus the teaching methods cannot be described in detail but as a generalisation of the main methods. The methods of evaluation, such as exams, can vary within a study package due to the variety of courses contained in the package.

As a unit of instruction, a package is longer than a course. The duration of a package can vary from months to one year. The amount of work required of the learner is announced in hours. Enrolment is required for admission, but there can be other requirements or limits as well. The package can belong to a certain discipline, and the courses of the package can represent individual or various subject matters. The packages offered by the course providers can be arranged only once or they can be repeated. In the latter case the version of the package must be announced, e.g. the starting semester of the package.

Programme

The term study programme refers to a collection of courses and/or packages that gives an official degree or a certificate, i.e. an externally visible outcome. In other words, a programme always requires courses or packages as its building blocks. The study programme is the largest unit in CUBER, and there cannot be any entities above it. Study programmes have the longest time frame; they can last for several years. The amount of work required from the learner cannot be announced in hours because of the length of studying. Instead, the official or average time needed to complete the programme can be announced. The programmes can include courses in one or more disciplines and on different levels of difficulty, but the degree itself can be only in one particular discipline. The entire programme can have its own level of difficulty, such as the Master's or doctoral level.

The learning targets are on a rather general level, and the learning outcomes are broader than in a course or in a package, and they should be described in general terms. The teaching methods can vary remarkably within a programme, for which reason the description should allow using generalised terms. Due to the length and complexity of a programme, the methods used to evaluate students can be diverse. Consequently, only the main alternatives should be mentioned here. Programmes can be offered only once, or they can be repeated. If the programme is repeatable, the versions of the programmes must be distinguishable, because the structure and the content may change from year to year.

5.5 Metadata structure

Metadata is composed of *categories and metadata elements*, which can have *sub-elements*. The structure of CUBER metadata is *hierarchical.* The elements represent the general level of description, and the sub-elements are more detailed in the information they convey. The categories exist only for grouping the related metadata elements together. (LOM, 2004.)

5.5.1 Data elements

Each data element in the metadata schema has been defined in terms of name, explanation, size, order, value space, data type and example. These terms define the characteristics, values and purposes of use that are allowed for the metadata element.

- Name - the name by which the metadata element is referenced
- Explanation - the written definition of the metadata element
- Size - the number of values allowed for the metadata element
- Order - whether the order of the values is significant or not
- Value space - the set of allowed values, e.g. a vocabulary
- Data type - a set of distinct values, e.g. char, string
- Example - an illustrative example of the metadata element

Notice that both the Size and Data type columns in the metadata specification can include smallest permitted maximum values (explained later). Only sub-elements and metadata elements without sub-elements can carry information as values. Categories and metadata elements that include sub-elements cannot have values. These metadata elements can have values indirectly through their sub-elements. (LOM, 2004)

After the requirements analysis in the CUBER project, two extra data elements, or columns, were added to the original LOM schema:

- **Aggregation levels** - defines the aggregation levels on which the metadata element should be used
- **Mandatory** - defines whether it is obligatory to use this metadata element to describe the study element.

These determinants are to be used within the CUBER system only for the special purposes of CUBER.

5.5.2 Definitions

There are some terms in the metadata specification that need definition in order to be used in a coherent manner. The following list defines these terms according to LOM, and they have to be used in CUBER as described here.

- *A category* is a rubric for a group of related metadata elements gathered together. Categories cannot have values; i.e. they do not carry any information.

- *A CUBER metadata element* is a data element for which the name, explanation, size, order, value space and data type are defined. An example can also be provided. Metadata elements can have values; i.e. they can carry information.
- *A value space* defines the set of values for a given metadata element. In CUBER the value space is usually defined by referring to a given standard or vocabulary. The value space can be enumerated outright as well.
- *A Langstring* is a specific LOM data type that represents phrases in one or several human languages. Multiple semantically equivalent phrases can be included, e.g. translations and alternative descriptions.
- *A smallest permitted maximum* defines the smallest permitted maximum value the application must support for that data type.
- *A reserved data element* is a data element that is not present in data instances.
- *A taxonomy* is a hierarchy of terms arranged from general to specific. It describes and defines a particular classification system in a specific field.
- *A data type* defines a set of distinct values, characterised by the properties of those values and by the operations on those values.
- *A vocabulary* is a list of values that define the value space of a metadata element. The use of the vocabulary is recommended in order to guarantee high interoperability. (LOM, 2004.)

5.5.3 List values

In some instances, a metadata element may contain a list of values, rather than a single value. A list of values must contain at least one value; in other words, the list cannot have a zero-length. A list of zero length cannot be distinguishable from no value, and if a value is intended to be present, a list of zero length cannot be valid as a final value. However, lists of zero length can be used for internal operations of an implementation. If a metadata element with sub-elements contains a list of values, then each of these values shall be a tuple of sub-elements. This means that the value of that metadata element is a list of pairs of the form (sub-element1, sub-element2). (LOM, 2004.)

There can be two kinds of lists: ordered and unordered. In an ordered list the order of the values in the list is significant. For example, the more important attribute can be mentioned first in the list, and a hierarchical classification system proceeds from general towards specific. In an unordered list the order of the values bears no meaning, and values of the list can appear in any order without any loss of information. (LOM, 2004.)

5.5.4 Vocabularies

A vocabulary is *a recommended list of appropriate values for a given metadata element.* Vocabularies have been defined for some of the CUBER metadata elements. Although the vocabulary is meant to be used as defined in the metadata schema, other values that are not present in the vocabulary list may be used as well. (LOM, 2004.)

In most cases it is preferable to use the CUBER vocabularies, for these values have the highest semantic interoperability, which ensures common understanding of the metadata. Some vocabularies of CUBER have been adopted from LOM as such, and some vocabularies have been created specifically for CUBER in order to be able to express the intended meanings and dimensions of the study elements. The vocabularies of LOM have been used whenever possible. In some cases the vocabularies of LOM had to be accompanied with extra values that will be used consistently within CUBER.

5.5.5 Smallest permitted maximum values

The intention of the smallest permitted maximum value is to cover more than 99% of all cases of the values of the data type of the metadata element. Smallest permitted maximum values can be used for two determinants of the metadata elements: size and data type. In addition, there are two cases for which the smallest permitted maximum has a slightly different meaning. First, for the metadata elements with a list value, the applications must support at least that number of entries for the list. The maximum number of entries to be supported imposed by the application must not go under the smallest permitted maximum defined for the metadata element. Second, for the metadata elements with data type Characterstring or Langstring, all applications must support at least that length for the Characterstring value, either directly or contained in the Langstring. The maximum number of characters to be supported must not go under the smallest permitted maximum for the data type of that metadata element. (LOM, 2004.)

5.5.6 Character sets

The aims of LOM and CUBER are to define a conceptual structure for metadata related to learning objects and education. The metadata scheme specifies some external standards to which any Characterstring representation should conform. In the case of non-restricted Characterstring values, reference is made to the repertoire of ISO/IEC10646-1. The decisions to be made that deal with representation shall be taken with a view to support multiple languages. (LOM, 2004.)

5.6 Conformance to LOM

CUBER metadata has been designed to conform to LOM Draft Standard version 6.1, which was the latest version of the draft standard during the metadata development process in CUBER in 2001. Some extensions have been made, for LOM was insufficient for the specific purposes of CUBER. The original LOM metadata elements have not been replaced or abused. They have been taken into CUBER as they were defined in the LOM Draft 6.1. However, not all of the LOM metadata elements are used in CUBER, because they were of little significance for the goals of CUBER. These metadata elements have been left empty, but they have not been removed. The CUBER extensions are added to the LOM Base Scheme as independent metadata elements and sub-elements. Also vocabularies have been modified and added to meet the goals of CUBER, but the vocabularies of LOM have been utilised wherever possible to ensure as high interoperability and conformity as possible.

5.7 Ontology in CUBER

A definite problem emerged during the metadata specification process. The equivocal terms and concepts used within CUBER caused many misunderstandings and misinterpretations due to the variations in the meanings of the terms depending on such factors as nationality or discipline. As there were several nationalities involved in CUBER, there could be several competitive interpretations of the concepts central for the metadata specification. The inevitable need for commonly understood and unambiguous definitions of concepts and the relationships between them was recognized within the project consortium.

In the CUBER project, in which research focused on a European virtual university, one of the most important issues was the lack of a common, shared understanding of educational terminology and concepts among the participating universities and countries. In order to enable seamless and transparent transfer of data in the virtual university, a standard way for both people and computers to communicate all necessary knowledge was needed, with both people and computer systems. The objective of the research on domain ontology was to enable this communication by providing a conceptual model and unified semantics for the purpose of building an information infrastructure for a virtual university.

The goal of ontology building is to simplify reality while retaining fidelity (Lamminaho, 2000). In other words, an ontology provides a vocabulary for users and programmes, with which to communicate about knowledge (Farquhar, Fikes & Rice, 1997), and it is essential for the development and use of intelligent systems. Building an ontology is reported to be extremely difficult, and by its nature it is often occurs in an iterative manner (Jokela, Turpeinen & Sulonen, 2000). At the same time it has been clearly shown that there is a definite need for defining the semantics of the resources (Dominique & Motta, 2000).

In this study, the ontology development process begun by recognising the essential ontology related requirements in higher education of information technology. It has turned out that when matching the educational supply to learners' educational demand a very important aspect is that the level or the specificity of describing the supply (course profile) also has to match to the needs of the learner (user profile).

During the course of the CUBER project, relevant information was analyzed by a multi-professional and interdisciplinary group, and a conceptual model and a domain ontology was developed. In addition, a meta-ontology that describes all of the ontology dimensions and their internal structure and data was defined. These were represented as class diagrams in UML notation. UML notation was chosen because the system developers best understood it and because it provided a visual representation of the information model.

Each of the ontology dimensions was discussed with domain experts, and the most proper ontology structure was chosen after a consensus was reached. Apart from the cases where the ontology clearly had no structure and was flat, both hierarchical and acyclic ontology structures were considered. Typically, many of the ontologies were flat, several had some hierarchy in them and one of them was best implemented as a DAG (Directed Acyclic Graph). In other research in the area of ontologies for educational use (Crampes & Ranwez, 2000), the focus has been mainly in the information system structure, whereas in this study the goal was to find content and structure for the ontologies and a proper conceptual model for the area.

There were many tools to choose from when defining and implementing an ontology. Logic languages could be used, data models could be drawn, for example with UML. In some cases a combination of XML and RDF would feasible, or the ontology could be described with some existing ontology building system, or with an ontology description language. However, at the core of CUBER was a database, not a Web server (or several). This is why it was not chosen to define ontologies as XMLs and DTDs, but rather as UML-diagrams and a database schema generated from them. The main concern was easy implementation of the metadata and ontology as a database installation.

In Figure 2, an example of a simplified hierarchical ontology is presented. In this example, subject can be divided into two dimensions, which can further have subdimensions under them. However, problems arise when the common area of two nodes in different ontology dimension branches need to be defined.

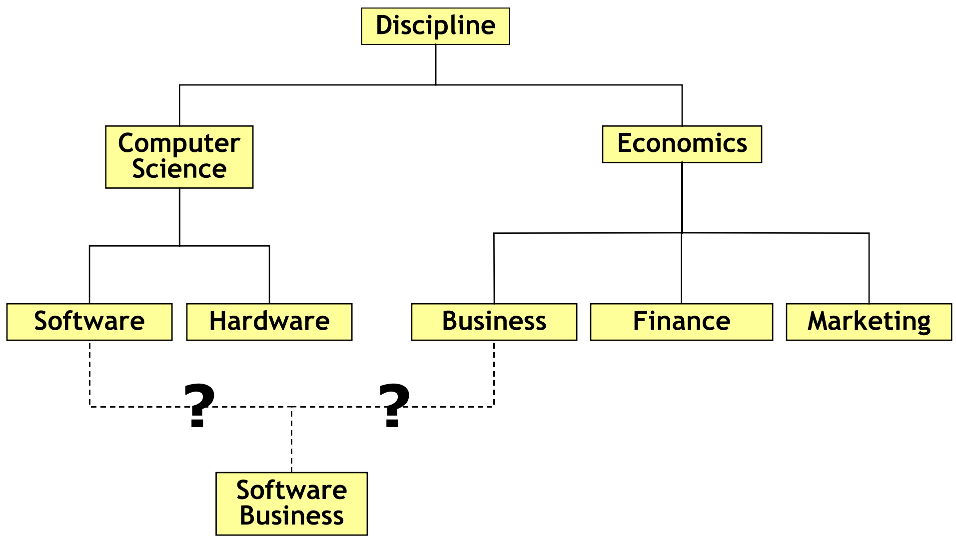


Figure 2. Graphical example of a hierarchical ontology

A directed acyclic graph is a graph where all transitions are unidirectional and there are no loops. The need for ontologies of this sort has been found also in other studies, for example in the Ontolingua Server (Farquhar et al., 1997), and also in the MARC research (Lamminaho, 2000). In a typical case, a course or other object in the virtual university can cover several slightly different subjects. It is practical to be able to tell the system exactly *how much* a course deals with a certain subject. In this way, we make it possible for user queries to be more accurate and we generate better results. The metadata can be given weights, so that it can be partly about one subject area and partly about others.

5.7.1 Meta ontology and conceptual model

The defined set of class diagram representing the central concepts of CUBER and the relationships and rules between them consisted of two parts: the *conceptual model* and the *meta ontology*. These are discussed in the following paragraphs.

The **meta ontology** showed the structure of the ontology in the class diagram. Each appropriate characteristic related to the domain was introduced as dimensions of the ontology. These dimensions were attached to relevant concepts in the conceptual model with associations. Each dimension of the meta ontology was represented as a class with or without an association to itself. If the class did not have an association to itself, the ontology was considered *flat*. When the association was present, the structure was defined as either hierarchical or a directed acyclic graph (DAG) depending on the multiplicities of the association. A many-to-many association represented a graph structure, whereas a one-to-many association represented a tree. The unlikely occasion of a one-to-one association was just an exception of the tree structure, i.e. a tree without any branches. Where applicable, the actual content of the ontology dimensions was presented as separate graphs, or external authorized ontologies were used whenever possible.

The **conceptual model** of CUBER defined the essential concepts of the domain and their attributes. These concepts were also attached to the meta-ontology with associations to the relevant ontology dimensions. The whole conceptual model was divided into parts in order to provide clarity and ease of understanding. The modeling was done with UML (Unified Mark-up Language) and realized with Rational Rose Enterprise Edition. It has to be noted that some additional notations were used, mainly to provide information not available within the Rational Rose notation. This information was vital when the schema was implemented into a relational database schema, as was the case with CUBER. The diagrams contained a few basic elements. The most important concepts of the CUBER domain ontology are *programme*, *package*, *course* and *material*. As an example, some of the main concepts and their relationships are shown in Figure 3. All attributes have been omitted for the sake of clarity.

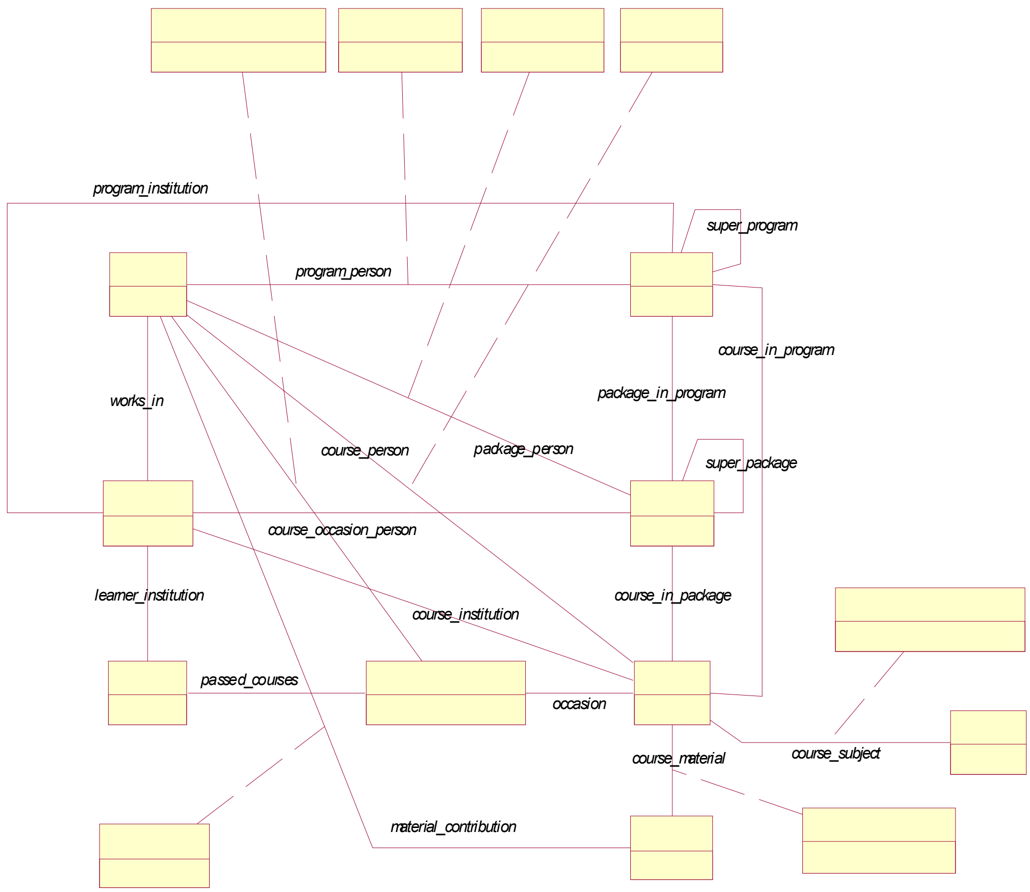


Figure 3. Conceptual model of the CUBER metadata and ontology

Classes represent entities of the modeled domain. In the conceptual model of CUBER, examples of these are *Person*, *Course* and *Institution*. Classes have attributes, which are basically just relevant information about the class. The relationships between classes are called associations. For example, the class *Package* has an association to itself, which means that two instances of this class can have an association. In the example this association has been named *super_package*. This means that a package can be the superpackage of another. In other words, a package can consist of other packages. The multiplicities at the ends of the association (marked with *) mean that a package can have many subpackages and can also be the subpackage of many other packages. The association *package_person* has also some information that cannot be attached to either *Package* or *Person*. This information is not relevant to any package, nor is it to any person. It is relevant only to the specific *combination* of a package and a person. Therefore, an association class *Package role* has been defined. This class and its attributes depend on an association between a certain package and a certain person. This is shown with a dashed line between the association class and the association to which it is attached.

In addition, the two following notations were used that are not found in a standard UML: A *key figure*: When there is a small key in front of an attribute, this means that the attribute is an identifier of the class. This is mainly relevant in the relational database schema, where this attribute is marked as a *primary key*. The value of this attribute has to be unique, i.e. no two instances of this class can have the same value for it. An *asterisk*: An asterisk following an attribute declaration means that this attribute has *multiple values*. Whereas an attribute usually has one value for each attribute, an attribute with multiple values can have many.

5.7.2 Implementation of the class diagram

This chapter describes how the class diagram (conceptual model and ontology) was transformed into a database schema and how the metadata schema has been transformed into an XML schema with a DTD file.

The transformation from the class diagram to a database schema was made in a very straightforward manner due to the fact that the class diagram with its classes and relations is quite close to a database schema. In the field of database management, there are established manners by which to do this. The result is a set of database tables with attributes, and all the information in the class diagram is found in the database schema. This schema is then converted into a suitable normal form. In the CUBER project the third normal form (3NF) was considered appropriate. After normalization, the tables are converted into SQL and can be directly imported into a relational database, such as the one used for the database at the core of the CUBER system. All information in the conceptual model and the ontology was made into CREATE TABLE commands in SQL. This way, the SQL definition is directly executable in a relational database, and all the necessary tables in the conceptual model containing all the information included in the metadata definition are created. The contents of the ontology dimensions were also added into the SQL definition, except for two: Subject and Discipline. These two are classifications rather than ontologies. They are references outside the conceptual model, and are dealt with in the application using the database, not in the database itself. An example of the SQL definition is attached to this thesis as Appendix 3.

In addition to the database solution, an XML schema was made of the metadata for possible future use and exchange of data with other related applications. The XML schema consisted of a DTD file, in which the structure and the tags of the XML file were introduced. The tags were exactly the same column titles found in the CUBER metadata definition, and the XML schema contained the metadata information in a structured form. Further constraints to the metadata content could be made with XML Schemas technology, and in the XML Schemas data types could be defined. The XML definition of the metadata could be used by any application capable of reading the information in the XML. This could be, for example, a Web application, where the semantics of the metadata tags would be defined. Thus, XML enabled the metadata to be used in a Web environment where a standard set of tags was understood.

5.8 The Evaluation of the Metadata Model

Finally, the conceptual model of CUBER was implemented in the form of a relational database and separate user interfaces were built for the end users searching for the information (learners) and for the end users introducing it to the system (learning resource providers). Several usability tests were conducted while user interfaces were iterated, and during the last project months a final evaluation was conducted on the acceptability of the information model itself.

The main concern in the evaluation of the CUBER information model was its ability to describe content providers' educational offers. This issue could only be addressed after there was a final representation of the CUBER model available. Therefore the ideal time for starting the evaluation was after the release of the final system, at the same time when the partners started uploading the information on their educational resources into the CUBER system. (Kautonen & Pöyry, 2002.)

Two kinds of evaluation were used. The user acceptance study was conducted during the CUBER project, and the scenario-based evaluation was performed for this thesis after the end of the project. The user acceptance study concentrates on the course provider group whereas the scenario-based evaluation takes also the viewpoint of the learners into account.

5.8.1 User acceptance study

The purpose of the user acceptance study was not to validate the quality of the system but to explore the acceptance of the information model, to detect possible conflicts and to discover requirements for further development. Therefore, the study concentrated on investigating the original users' information model in relation to the CUBER system's model, revealing the gaps between these two information models, and detecting the effects of their possible conflicts to the use of the system. (Kautonen & Pöyry, 2002.)

The scope of the study was focused on the users of the content provider interface, since they were considered better experienced in educational concepts and educational systems, as well as more experienced in the CUBER information model. The study was carried out in two phases, and it included a questionnaire and a thematic semi-structured interview. First, the participants were asked to fill in a feedback form (questionnaire) on the metadata authoring interface that included the representation of the metadata model. The questions concerned the metadata model, its' general acceptance and details such as understandability and compatibility with the local educational system. Second, the participants were interviewed and the information gathered by the questionnaire was used to guide the conversation towards clarifying questions and acceptance of individual metadata elements. (Kautonen & Pöyry, 2002.)

The questionnaire for the users of the metadata authoring interface consisted of five parts. The first part contained background questions, and the second part consisted of general questions about the web pages that were used to upload metadata descriptions. There were questions such as: "How did you understand the meaning of the elements?" and "Is the order in which the information is presented consistent or confusing?" In the third part of the questionnaire the respondents were asked to evaluate each metadata element with the scale important – usable – not usable – not relevant. In the fourth part there were questions about training and guidance related to the CUBER system. The fifth part of the questionnaire asked the respondents to provide general feedback on the experiences on the CUBER information model and the features of the CUBER system. After filling out the questionnaire, the respondents were interviewed. The thematic interview was based on the questionnaire and its purpose was to discuss the themes of the questionnaire in more detail. (Kautonen & Pöyry, 2002.) The questionnaire is attached to this thesis as Appendix 4.

The number of subjects was considerably small (N=5), and thus the results cannot have widely generalized value. Nevertheless, these individual responses were adequate for revealing the desired information, i.e. how well the information provided by an individual respondent, who represented his/her organization and local educational system, matched or mismatched with the model offered by CUBER. (Kautonen & Pöyry, 2002.)

The results form the feedback form (questionnaire) and the thematic interviews can be summed up to the following findings (Kautonen & Pöyry, 2002):

- Because the users had used the system for relatively short time, they had not used nor needed more advanced functions yet. Therefore they were not able to express their opinion on more complex matters, such as the preferred relations between study elements.

- The study element aggregations are efficient for describing the study elements of the provider organizations, although the supply of an individual organization did not match with the model in full. These mismatches were not considered to cause any sort of defect or obstacle to efficient use of the system.
- The idea of displaying data for aggregations on different abstraction levels is welcomed, since different information is needed to describe different study elements.
- There are some metadata elements, the use and options of which need to be reconsidered. There may also be need to find better-accepted definitions for some individual elements. Nevertheless, there should be more examples of all elements and their usage.

The following metadata elements needed revisions according to the user feedback (Kautonen & Pöyry, 2002):

- Title in English: In some cases there might be conflicts in defining which is the most appropriate title to use; the original or the English one. For this there should be instructions.
- Version: It may not be needed to indicate the version except for the material. For other aggregation levels this element is not so useful.
- Catalog Entry: This option should be reconsidered, but on the other hand some partners are using this element. Evidently more explicit instructions should be given for this element.
- Financing possibilities: This kind of information should be on the educational institutions own web pages, not in the CUBER system itself. A link to the source of information could be provided.
- Structure: The users will need more concrete examples and definitions in order to understand this metadata element.
- Published and other dates: The format of the dates is not appropriate, and the blank entry should be allowed.

- Some elements that have options defined in the ontology will need more options or the options should be even better defined, e.g. “Operating system”.

The results of the acceptance study indicate a need to further define, model, and finally visualize the relations between study elements and other studies. Although the system and its information model were perceived as generally acceptable, better functions on this concept may eventually increase the usefulness and thus usability of the system.

5.8.2 Scenario-based evaluation

In addition to the user acceptance study carried out in the CUBER project, the CUBER metadata model was further evaluated against scenarios based on a user study related to a mobile learning system framework development. This evaluation was conducted for three reasons. Firstly, the user study carried out in the beginning of the CUBER metadata development did not produce detailed enough results as for defining detailed information requirements for the metadata model. Secondly, the LOM standard was missing many elements related to the learning and teaching process, and these elements were added to the LOM schema as extension elements. Thirdly, the number of participants in the CUBER end user acceptance study was quite limited. The evaluation was an analytic one, but based on empiric user data from another study in the area of learning technology.

Mostakhdemin-Hosseini (2004) introduces several scenarios related to a mobile learning system in his licentiate’s thesis. These use scenarios illustrate the concrete ways in which the different user groups might really utilise the mobile learning system, including the context in which the system is used. The user groups of this mobile learning system consist of university students, course assistants, and lecturers, and the scenarios presented in the thesis are based on a user study conducted at the Helsinki University of Technology. The use scenarios describe how the different users interact with a mobile learning system that is used directly for the university’s teaching activities. (Mostakhdemin-Hosseini, 2004.)

However, in this thesis the focus is on a system, more specifically on the CUBER system, that is intended to provide information on the educational supply in order to support the decision-making process of the potential students. Due to these fundamental differences in the systems' purposes, the scenarios developed for mobile learning cannot be directly used in the evaluation of the CUBER system's information model. The tasks to be completed with the CUBER system and the mobile learning system are not directly comparable with each other. Despite of that, the scenarios offered a valuable opportunity to gain additional information based on an empirical user study, because the scenarios provided in the thesis were on a general level enough. Thus the use scenarios were analysed in order to identify the requirements and needs for information that a system (such as CUBER) would have to provide in order to be able to describe the learning and teaching activities defined in the use scenarios. The analysis was done by reading through the scenarios several times. Two questions were asked from the scenarios: 1) what are the teaching and learning actions and 2) what kind of information is needed to support these actions. Then the information requirements for completing the tasks described in the scenarios were listed by answering the second question. The requirements from all scenarios were integrated into one list. These very concrete requirements were categorised as a result of a simple content analysis. These information requirements should be taken into account in the CUBER metadata model in order to be able to *describe* the learning related activities, because the intent of the CUBER system was to communicate descriptive information (with the help of metadata) about the educational supply from the universities to the students.

After identifying the information requirements from the scenarios, a comparison was made between these requirements and the CUBER metadata model in order to find out whether each requirement had a counterpart in the information model, such as a metadata element, or several. The purpose of the comparison was to evaluate the CUBER metadata model against the information requirements that were derived from real empiric user data. The scenarios described from both students' and teachers' perspective the users' activities with the mobile learning system, such as registering to a course, scheduling the course activities, defining the passing requirements for a course, and contacting the course staff etc. (Mostakhdemin-Hosseini, 2004). The information requirements and the results of the comparison are presented in Table 4. In addition, the origin of each metadata element is indicated in order to evaluate the need to expand the LOM standard.

Table 4. Comparison between the scenario-based information requirements and the CUBER metadata elements

Information requirement	CUBER metadata element(s)	Origin (LOM/CUBER)
1. Registration information	5.16 Educational_Enrolment	CUBER
2. Contact information of course staff	1.10. General_Provider, 5.17 Educational_StudyGuidance	CUBER
3. Assignment or exam related to a course	5.15.2 Educational_Evaluation_Method, 5.15.3 Educational_Evaluation_Number	CUBER
4. Requirements related to passing the course	5.15.2 Educational_Evaluation_Method	CUBER
5. How to return the assignments	-	-
6. Communication between students and course staff	5.17 Educational_StudyGuidance	CUBER
7. Schedule of the course tasks and assignments	1.13 General_Date	CUBER
8. Lecture schedules	1.13 General_Date,	CUBER
9. Information on how to access to course/lecture material	4.3 Technical_Location	LOM
10. Format of course/lecture material	4.1 Technical_Format, 5.2 Educational_LearningResourceType	LOM
11. Course website address	4.3 Technical_Location	LOM
12.Equipment needed for the course	4.4 Technical_Requirements, 4.9 Technical_Description	LOM, CUBER
13. Software used on the course	4.4 Technical_Requirements, 4.9 Technical_Description	LOM, CUBER
14. Mode of working on the course (e.g. group work)	5.12 Educational_TeachingActivity	CUBER
15. Version or updating of the course website or material	2.1 LifeCycle_Version, 2.2 LifeCycle_Status	LOM
16. Grading of the course	5.15 Educational_Evaluation	CUBER

The results of the comparison and the evaluation indicate that the CUBER metadata model succeeded well in fulfilling the information requirements derived from the scenarios. Only one of the requirements (5. How to return the assignments) could not be expressed with the CUBER metadata elements, unlike all other information requirements. However, information on how to return course assignments should be found from the course web-page itself, not from the CUBER system that is intended only to provide rather general descriptive information on the educational supply. Thus it can be concluded that the CUBER metadata model meets the users' needs well. This result is well in line with the result from the acceptance study of CUBER, and the results of these two evaluations support each other.

As for extending the LOM standard, the results of the evaluation seem to justify the adding of own CUBER metadata elements to the LOM schema. Only four of the 16 information requirements could have been fulfilled only with the LOM metadata elements. CUBER extensions were needed in 11 cases either to provide a completely missing metadata element or to complement the existing LOM elements. The conclusion is that the LOM standard needed to be extended by the CUBER metadata elements in order to meet the users' needs. Here it should also be noted that the CUBER metadata development process was able to take into account the users of the system despite the shortcomings of the user study.

However, the evaluation based on the scenarios was limited to only to only two aggregation levels used in the CUBER system, namely Material and Course. This is because the scenarios were only describing actions that included these aggregation levels defined in CUBER. Moreover, it was not possible to evaluate the relations between the educational elements in CUBER, because of the limited scope of the scenarios. If further investigated, the CUBER metadata model should be evaluated in real use situations with real data in the system.

5.9 Summary of the results

The results of this study are summarised in this chapter by answering the research questions defined in Section 4 of this thesis. Most of the results have already been presented in the other chapters of Section 5, but for the sake of clarity each research question will be answered here concisely.

1. Using metadata in the context of virtual university:

a. What kinds of user requirements are there for the metadata?

First of all, it is challenging to define the user requirements for metadata, because the users may not be able to clearly articulate their needs and expectations. The user requirements in this study were related to the basic need of being able to find from the Internet the learning resources that would correspond the criteria of the students. For the second end user group this basic need was to be able to describe the educational offer in an efficient manner. As the user requirements may be vague, continuous involvement (if not even co-development) with the end users was required in this study so that the users' needs could be met.

b. Are the current and emerging educational metadata standards sufficient for virtual learning environments?

The current and emerging educational metadata standards may not be sufficient or suitable as such for the various virtual environments for learning. The standards are always high-level, global descriptions of the domain, and they often lack the elements needed in local applications that require very specific metadata elements. This was the case in this study conducted as a part of the CUBER project. The metadata schema developed is an application profile of LOM, including the structure and most elements of LOM but also specific CUBER extensions and modifications.

c. What kind of metadata is needed in virtual university?

In virtual university there is a need for mainly semantic metadata or educational metadata that is specific to the context of use. The CUBER metadata schema (see

Appendix 1) includes the metadata elements that were considered necessary or useful for this virtual university application.

d. How should metadata be structured in a virtual university application?

The metadata should be structured so that it reflects the reality of the virtual university field. In this case this refers to the structural similarity with the higher education studies in Europe. For example, the structures such as study courses and degrees have to be expressed with the help of metadata, which means that the metadata must be structured accordingly. In this thesis the CUBER aggregations are introduced as a conceptual structure for the educational metadata while the LOM structure for metadata is used as the technical structure.

2. Using ontology in the context of virtual university:

a. What kind of ontology is needed in virtual university?

As the field of virtual university is rather limited, a domain ontology is needed. It should be integrated with the metadata schema of the domain.

b. What kinds of requirements are there for ontology?

As ontology provides semantics for the metadata, it should be able to define the meanings of the metadata elements. Moreover, the relationships between the concepts have to be defined in the ontology.

c. How should ontology be structured in a virtual university application?

As ontology is in a way a conceptual model, it should represent the reality as it is, even though in a simplified manner. Thus the domain ontology of a virtual university should be structured according to the structure of the learning objects.

3. Conceptual modeling and information representation in the context of virtual university:

a. How can metadata be modeled and represented?

In this study it was considered feasible to model and represent metadata in the form of a table that contains all metadata elements, their explanations, constraints on use, and other attributes. This format was chosen for two reasons: first, it followed the LOM standard, and second, it was easily understandable for the persons involved in the project.

b. How can an ontology be modeled and represented?

In this study a feasible way to model and represent an ontology was using UML class-diagrams. Formal and semi-formal ontology representation languages would have been available, but they were abandoned for several reasons. The system development team building the CUBER database was not familiar with these languages, whilst most of the whole project team was able to work with the UML diagrams. Aside with this practical solution supporting the system development, UML offered a means to visualise the ontology. This was considered important because this way the structure of the ontology and relationships between the concepts could be shown.

c. Is the conceptual model developed in this study understood and accepted by its intended end users?

The CUBER metadata information model was relatively well understood and accepted by the end users according to the evaluation of the model. However, there are some items for improvement and further research, such as the visualisation of the metadata. According to the scenario-based evaluation, the requirements of the users are met, and the extensions to the LOM schema can be justified.

6 Discussion and conclusions

In this study, research was carried out in the emerging field of e-learning and virtual universities, the focus being on the enabling information infrastructure for bringing higher education to the Internet. Classification, organization and retrieval of learning resources are amongst the most important questions when creating a virtual university. The use of metadata and ontology with clearly defined conceptual models enhance the discovery of learning resources and enable educational institutes to offer their courses for large, even worldwide audiences.

Firstly, the use of a standardized metadata model is essential in improving the information systems in the fields of e-learning and virtual universities by unifying the descriptions and classifications of educational resources. However, the use of metadata alone does not solve the problems. A domain ontology integrated into metadata is needed to form a basis for common semantics and shared understanding of the domain concepts and their relationships.

Secondly, another key finding of this study is the need to modify general standards to meet the needs of a specific application area. In this case, a general level e-learning metadata standard was extended but the compatibility with the standard was maintained by creating an application profile compliant with the standard. The application profile was not only extending the standard but also further defining its concepts on a more detailed level. This way the standard was interpreted locally, whilst at the same time care was taken of the interoperability with other systems using the LOM standard.

Thirdly, in this study it proved useful to develop a conceptual model in UML class diagrams instead of using some formal ontolofy definition language. This conceptual model described graphically the central concepts of CUBER by integrating the CUBER metadata model and the domain ontology. These diagrams show the central concepts, and the relationships and rules between them. This kind of conceptual model serves both the system development and the representation of the domain specific metadata and ontology.

The key contributions of this study include the following:

- **An application profile of the LOM (Learning Object Metadata) standard developed by the IEEE.** In this study an application profile of LOM standard was created in order to meet the specific needs of the application domain, viz. higher education in Europe. The application profile uses the LOM standard as the basis for the metadata schema, but some extensions and refinements have been made due to the fact that a standard is on too a general level to be able to express all the details of a specific field. Especially the pedagogic metadata elements related to the learning context and process were lacking from LOM and they were added to the CUBER metadata schema.

- **A domain ontology for a virtual university information system.** The concepts and terms used in LOM standard and CUBER metadata were analyzed in an expert group consisting of researchers and practitioners from several disciplines and nationalities. As a result the semantics of the concepts and terms was defined in order to form a basis for common, shared understanding. In addition, the relationships between the central concepts were defined.

- **A conceptual model for e-learning in higher education.** A conceptual model for e-learning structures in higher education was developed by integrating the CUBER metadata model and the domain ontology into UML class diagrams. These diagrams show the central concepts, and the relationships and rules between them.

- **Introduction of user-centred activities into metadata modelling.** In this thesis the perspective of user centeredness and usability was introduced for developing a metadata model. The metadata model was tested and evaluated, and as the result the metadata model was accepted by its users, and it met the information requirements for describing the activities of a virtual university. The metadata model and interface were evaluated by the actual users of the system, i.e. representatives of several European universities. The evaluation of the CUBER metadata proved that the metadata model was accepted in general. Most metadata elements were regarded as useful or very useful, while only few elements were found useless or difficult to understand. However, the representation, such as visualization of the information should be improved in order to facilitate the usability of the metadata model.

Defining a thorough metadata schema with an ontology proved to be an interesting but a very challenging and demanding task. Expertise from different fields of science was needed in order to solve the multidisciplinary problems; the team consisted of researchers from educational science, software engineering and computer science. This kind of a multidisciplinary team was successful because many complementary viewpoints were needed, as was the case in the whole of the project. Using teams with members from various fields of science enabled to broaden the perspectives and to produce innovative solutions with real added value. The importance of user-centered approach in the metadata development was recognised during the research, and in the future it should be integrated into the development process more closely. User studies should be given more resources and importance so that the metadata model would really meet the needs of the users.

In a multi-cultural project like CUBER even the politics of education played an important role; the differences between the educational systems and the interests of the different nationalities became challenging starting points for the study. The field of education is very sensitive with regard to the policies of national governments. It is a real challenge to be able to combine the possibly conflicting views into a compromise that would be satisfactory for all participants. Actually, not only the policies and educational systems differ, but also the whole culture of education and accordingly the concepts, processes and assumptions. This adds extra complexity into the process of developing an educational information system for the use of many nationalities.

The use of existing and emerging standards is highly recommended on the basis of experiences from this study. In order to enable interoperability with other information systems may be crucial for the success of the system, and standards play a significant role in enhancing interoperability. Especially in the internationalizing educational market, the possibility to co-operate and exchange information may be crucial. Using a general and widely known metadata standard such as LOM is preferable, but an application profile tailored to the context specific requirements may be needed. However, this does not compromise the benefits of using a standard, provided that the application profile is designed carefully.

The future research should focus on the representation of metadata and related information models. For example, the visual representation of the information in the user interfaces can be studied and developed in order to enhance the understandability of the metadata schema. Moreover, designing a user interface that supports the representation and understandability of the metadata, ontology, and conceptual model is a topic for future research. In addition, investigating the usability requirements of a metadata model would be needed.

References

ARIADNE 2004. The ARIADNE Foundation Web site. <http://www.ariadne-eu.org/index.html> [Checked Dec-08-2004]

Benjamins, V.R., Fensel, D., Decker, S. & Gomez Perez, A. 1999. *$(KA)^2$: building ontologies for the Internet: a mid-term report*. Int. J. Human-Computer Studies (1999) 51. pp. 687-712.

Berners-Lee, T., Hendler, J. & Lassila, O. 2001. *The Semantic Web*. Scientific American, May 2001.

Bradley, N. 2000. *The XML Companion. 2nd Edition.* Addison Wesley.

Britain, S. & Liber, O. 1998. *A Framework for Pedagogical Evaluation of Virtual Learning Environments.* JISC Technology Application Programme.

Carroll, J.M., Chin, G., Rosson, M.B.& Neale, D.C. 2000. *The Development of Cooperation: Five Years of Participatory Design in the Virtual School.* Proceedings of DIS 2000, New York. Pp. 239-251.

Crampes, M. & Ranwez, S. 2000. *Ontology-Supported and Ontology-Driven Conceptual Navigation on the World Wide Web*. Proceedings of the 11th ACM Conference on Hypertext and Hypermedia, 2000.

CUBER. 2004. The CUBER project. <http://www.cuber.net> [Oct-28-2004]

Davies, J., Fensel, D. & van Harmelen, F. 2003. *Towards the Semantic Web – Ontology Driven Knowledge Management.* John Wiley & Sons Ltd.

Domingue, J. & Motta, E. 2000. *PlanetOnto: From News Publishing to Integrated Knowledge Management Support*. IEEE Intelligent Systems, May/June 2000. Pp. 26-32.

DCMI 2004. Dublin Core Metadata Initiative. <http://dublincore.org/> [Dec-08-2004]

Duval, E. 2001. *Standardized Metadata for Education: a Status Report.* Proceedings of ED-MEDIA 2001, Tampere, Finland. Pp.458-463.

Duval, E., Hodgins, W., Sutton, S. & Weibel, S.L. 2002. *Metadata Principles and Practicalities*. D-Lib Magazine April 2002. [Dec-08-2004] <http://www.dlib.org/dlib/april02/weibel/04weibel.html>

ECTS 2004. European Credit Transfer System's official website. < http://europa.eu.int/comm/education/programmes/socrates/ects_en.html> [Dec-08-2004]

Farquhar, Adam, Fikes, Richard and Rice, James. 1997. *The Ontolingua Server: a tool for collaborative ontology construction*. Int. J. Human-Computer Studies (1997) 46. Pp. 707-727.

Faulk, S. 1997. *Software Requirements: A Tutorial.* Software Requirements Engineering, 2nd edition. Edited by H. Thayer et al. IEEE Computer Society Press.

Fensel, D. 2001. *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce.* Springer Verlag.

Fensel, D., van Harmelen, F. & Horrocks, I. 2003. *OIL and DAML+OIL: Ontology Language for the Semantic Web.* In Davies, J., Fensel, D. & van Harmelen, F. (edited by) *Towards the Semantic Web – Ontology Driven Knowledge Management.* John Wiley & Sons Ltd. pp.11-31.

Gilliland-Swetland, A.J. 2000. *Introduction to Metadata, Setting the Stage.* <http://www.getty.edu/research/institute/standards/intrometadata/> [Dec-08-2004]

Gruber, T.R. 1993. *Toward principles for the design of ontologies used for knowledge sharing*. Padua workshop on Formal Ontology, March 1993. <ftp://ftp.ksl.stanford.edu/pub/KSL_Reports/KSL-93-04.ps> [Oct-28-2004]

Guarino, N.1997. *Understanding, building and using ontologies*. Int. Journal of Human-Computer Studies (1997) 46. pp. 293-310.

Heinrich, E. & Anders, Frederic. 2003. *Enriching Document Collections through the Writing of „Stories“.* Proceedings of ICALT 2003, Athens, Greece. Pp.101-105.

Hendler, J. 2001. *Agents and the Semantic Web*. IEEE Intelligent Systems Journal March/April 2001. <http://www.cs.umd.edu/users/hendler/AgentWeb.html> [Dec-08-2004]

Holzinger, A., Kleinberger, T. & Müller, P. 2001. *Multimedia Learning Systems based on IEEE Learning Object Metadata (LOM).* Proceedings of ED-MEDIA 2001, Tampere, Finland. pp.772-777.

Hyvönen, E. 2001. *Semantic Web – kohti uutta merkitysten Interetiä.* University of Helsinki, Department of Computer Science. Semantic Web Kick-Off in Finland 2nd November 2001.

IEEE/LTSC a. IEEE Learning Technology Standards Committee (LTSC)/IEEE P1484.12 Learning Object Metadata Working Group. 2003. *Draft Standard for Learning Object Metadata.* < http://ltsc.ieee.org/wg12/index.html> [Oct-28-2004]

IEEE/LTSC b. IEEE Learning Technology Standards Committee (LTSC)/IEEE P1484.3 Glossary Working Group. 2003. *Draft Standard for Information Technology --*

Learning Technology – Glossary. <http://ltsc.ieee.org/archive/harvested-2003-10/working_groups/> [Dec-08-2004]

IMS 2004. IMS Global Learning Consortium, INC <http://www.imsglobal.org/metadata/> [Dec-08-2004]

IQ-Form 2004. IQ-Form – Verkko-opiskelun ohjaus- ja arviointijärjestelmä. Virtuaaliyliopistohanke. <http://www.edu.helsinki.fi/iqform/default.htm> [Dec-08-2004]

Jokela, S. 2001. *Metadata Enhanced Content Management in Media Companies.* Acta Polytechnica Scandinavica. Mathematics and Computing Series No. 114. Helsinki University of Technology: Doctoral thesis.

Jokela, S. & Turpeinen, M. 2000. *Ontology Development for Flexible Content.* HICSS-33, Minitrack on System Support For Electronic Business On Internet.

Järvinen, P. & Järvinen, A. 1996. *Tutkimustyön metodeista.* Opinpaja Oy: Tampere.

Kassanke, S., El-Saddik, A. & Steinacker, A. 2001. *Learning Objects Metadata and Tools in the Area of Operations Research.* Proceedings of ED-MEDIA 2001, Tampere, Finland. Pp.891-895.

Kauppinen, Marjo. 2002. *Requirements Engineering*. Qure project. SoberIT, HUT.

Kauppinen, M., Kujala, S., Aaltio, T. & Lehtola, L. 2002*. Introducing Requirements Engineering: How to Make a Cultural Change Happen in Practice.* IEEE Joint International Conference on Requirements Engineering, Essen, Germany, September 2002.

Kautonen, H. & Pöyry, P. 2002. *CUBER Metadata Authoring Interface.* Internal report of the CUBER project. [Printed report available from the author.]

Kotonya, G. & Sommerville, I. 1998. *Requirements Engineering – Process and Techniques*. John Wiley & Sons Ltd. England.

Krämer, B. J. 2000. *Forming a Federated Virtual University Through Course Broker Middleware*. LearnTEC 2000, Karlsruhe, February 2000 <http://www.fernuni-hagen.de/DVT/Publikationen/Papers/LearnTEC2000.pdf > [Dec-08-2004]

Lamminaho, V. 2000. *Metadata specification: Forms, Menus for Description of Courses and All Other Objects.* CUBER project: Deliverable D3.1. [Available from author.]

LearningCircuits 2002. <http://www.learningcircuits.org/glossary.html> [Dec-08-2004]

Liu, J., Chan, S., Hung, A. & Lee, R. 2002. *Facilitators and Inhibitors of e-Learning.* In Virtual Environments for Teaching and Learning. Series on Innovative Intelligence – Vol. 1. Edited by: L.C. Jain, R.J. Howlett, N.S. Ichalkaranje & G. Tonfoni. World Scientific. pp.75-109.

LOM 2004. Learning Object Metadata standard (LOM) of IEEE. <http://ltsc.ieee.org/wg12/index.html> [Dec-08-2004]

Marquardt, M.J. & Kearsley, G. 1999. *Technology-Based Learning, Maximizing Human Performance and Corporate Success.* American Society for Training & Development. St. Lucie Press.

Metodix 2004. <www.metodix.com> [Oct-28-2004]

Mostakhdemin-Hosseini, Ali, 2004. *Framework for M-Learning Systems Based on Educational Components.* Helsinki University of Technology, Department of Information Processing Science. Licentiate's Thesis.

Mäkelä, A. & Battarbee, K. 1999. *Applying Usability Methods to Concept Development of a Future Wireless Communication Device – Case in Maypole.* Proceedings of 17th International Symposium on Human Factors in Telecommunications. Copenhagen, Denmark, May 4-7.1999. Pp. 291-298.

Monthienvichienchai, R., Sasse, A. & Wheeldon, R. 2001. *There's Trouble in Paradise: Problems with Educational Metadata Encountered during the MALTED Project.* Proceedings of ED-MEDIA 2001, Tampere, Finland. pp.1347-1353.

Männistö. T. 2000. *A Conceptual Modelling Approach to Product Families and their Evolution.* Acta Polytechnica Scandinavica. Mathematics and Computing Series No. 106. Helsinki University of Technology. Doctoral thesis.

Nevgi, A. & Tirri, K. 2003. *Hyvää verkko-opetusta etsimässä.* Kasvatusalan tutkimuksia. Suomen kasvatustieteellinen seura.

Niemi, H. 2002. *Empowering Learners in the Virtual University.* In Theoretical Understandings for Learning in the Virtual University. Edited by: Niemi, H. & Ruohotie, P. Research Center for Vocational Education, University of Tampere. Pp. 1-35.

Noy, F. & McGuinness, D.L. 2001. *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05, Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

Mohan, P. & Brooks, C. 2003. *Learning Objects on the Semantic Web.* Proceedings of ICALT 2003, Athens, Greece. Pp.195-199.

Patnaik, D. & Beeker, R. 1999. *Needfinding: The Why and How of Uncovering the People's Needs.* Design Management Journal 1999.

Rumbaugh, J. 1994. *Getting started – Using use cases to capture requirements.* Journal of Object Oriented Programming, Vol. 11, No.1 September 1994.

Ryan, S., Scott, B., Freeman, H. & Patel, D. 2000. *The Virtual University – The Internet and Resource-Based Learning.* Kogan Page.

Simon, B. 2001. *Do e-learning standards meet their challenges?* UNIVERSAL Project. <http://www.rz.uni-frankfurt.de/neue_medien/standardisierung/simon_text.pdf> [Oct-28-2004]

Sowa, J. F. 2000. *Knowledge Representation. Logical, Philosophical and Computational Foundations.* Brooks/Cole, Thomson Learning.

Strijker, A. 2001. *Using metadata for re-using material and providing user support tools.* Proceedings of ED-MEDIA 2001, Tampere, Finland. Pp.1823-1824.

Stojanovic, L., Staab, S. & Studer, R. 2001. *E-learning based on the semantic web.* In the Proceedings of WebNet2001 - World Conference on the WWW and Internet, Orlando, Florida, USA, 2001.

Teare, R., Davies, D. & Sandelands, E. 1999. *The Virtual University – An Action Paradigm and Process for Workplace Learning.* Cassell.

Turpeinen, M. 2000. *Customizing News Content for Individuals and Communities.* Acta Polytechnica Scandinavica. Mathematics and Computing Series No. 103. Helsinki University of Technology. Doctoral thesis.

Twigg, C. & Oblinger, D. 1996. *The Virtual University*. <http://www.educause.edu/ir/library/html/cnc9723/cnc9723.html> [Oct-28-2004]

VanDurm, R., Duval, E., Verhoeven, B., Cardinaels, C. & Olivié, H. 2001. *Extending the ARIADNE Web-Based Learning Environment.* Proceedings of ED-MEDIA 2001, Tampere, Finland. Pp.1932-1937.

Wason, T. 2001. *Dr. Tom's Metadata Guide.* IMS Global Learning Consortium. <http://wason.home.mindspring.com/TDW/Papers.htm> [Dec-08-2004]

WEINSTEIN, P. 1998. *Ontology-Based Metadata*: Transforming the MARC Legacy. Proceedings of the Third ACM Digital Library conference, Pittsburgh, PA, USA, June 1998.

Wiendieck, G., Preisendanz, M. & Filla, M. (2000). *Final report and model of user behaviour.* Deliverable D1.6 of the CUBER project. [Available from author.]

Appendix 1. The CUBER Metadata Schema

CUBER METADATA SCHEMA version 2.2
Based on LOM 6.1.

Note: The metadata elements written in normal **font are original LOM elements.**
The metadata elements written in *italic* **font are original LOM elements, but NOT used in CUBER.**
The metadata elements written in bold font are CUBER's extensions to the LOM.
CUBER has also added two new columns to the table of LOM schema: "Aggregation levels" and "Mandatory". These will be used only for the CUBER system.

Nr	Name	Explanation	Size	Order	Value Space	Data Type	Example	Aggregation levels	Mandatory
1	General	General descriptive information about the learning object as a whole.	1	N/A	-	-	-	-	-
1.1	Identifier	A globally unique label for identifying the learning object.	1	N/A	-	Reserved.	Not Used.	-	-
1.2	Title	Name of the L.O.	1	N/A	-	Langstring (smallest permitted max: 1000 char)	-	All	Yes

1.3	CatalogEntry	Defines an entry within a catalog assigned to this L.O.	Smallest permitted max. 10 values	No.	-	-	-	-	-
1.3.1	Catalog	The name of the catalog (i.e. the listing identification system).	1	N/A	Repertoire of ISO/IEC 10646-1	Characterstring (min-max: 1000 char)	ISBN, ARIADNE	All	No
1.3.2	Entry	Actual string value of the entry within the catalog defined in 1.3.1.	1	N/A	-	Langstring (smallest permitted max: 1000 char)	-	All	No
1.4	Language	The primary human language(s) used within this L.O.	Smallest permitted max: 10 items	No	LanguageID =Langcode (´-´Subcode)*, ISO 639, ISO 3166	Characterstring (smallest permitted max: 100 char)	"en" "en-GB" "de" "fr-CA" "it"	All	Yes
1.5	Description	A textual description of the content of this L.O.	Smallest permitted max: 10 items	No	-	Langstring (smallest permitted max: 2000 char)	-	All	Yes for Course, Package and Programme
1.6	Keywords	Keywords or phrases describing this L.O.	Smallest permitted max: 10 items	No	-	Langstring (1000 char)	-	All	Yes
1.7	*Coverage*	*The span of such things as time,*	*Smallest permitte*	*No*	-	*Langstring (smallest*	*(en, Circa,*	*NOT USED IN CUBER!*	-

		culture, geography or region that applies to this L.O.	*d max: 10 items*			*permitted max: 1000 char)*	*16th century France)*		
1.8	Structure	Underlying organisational structure of this L.O.	1	N/A	Collection Mixed Linear Hierarchical Networked Branched Parcelled Atomic	Vocabulary; see **Ontology**	-	Material Package Programme	No
1.9	*Aggregation Level*	*The functional granularity of the L.O.*	*1*	*N/A*		*Vocabulary*	-	*Not used in CUBER*	-
1.10	**Provider**	**Information about the provider or organiser of the L.O.**	**N**	**No**	**-**	**-**	**-**	**-**	**-**
1.10.1	**Institution**	**The name and other information about the institution that organises or provides this L.O.**	**N**	**No**	**V-card**	**Langstring (smallest permitted max: 1000 char)**	**University of X**	**Course Package Programme**	**Yes**
1.10.2	**Person**	**Information about the**	**N**	**No**	**-**	**-**	**-**	**-**	**-**

		persons related to the L.O.							
1.10.2.1	**Role**	**The roles of the persons involved in providing the L.O.**	**1**	**N/A**	**Administrator Advisor Assistant Contact person Examiner Lecturer Teacher Tutor**	**Vocabulary; see Ontology**	**-**	**Course Package Programme**	**Yes**
1.10.2.2	**Information**	**Information about the person related to the L.O.**	**1**	**N/A**	**V-card**	**Langstring (smallest permitted max: 1000 char)**	**-**	**Course Package Programme**	**Yes**
1.11	**CUBER Identifier**	**A label for identifying the L.O. Valid only within CUBER.**	**1**	**N/A**	**Repertoire of ISO/IEC 10646-1**	**Characterstring (min-max: 100 char)**	**-**	**All**	**Yes**
1.12	**CUBER Aggregation**	**The functional granularity of the L.O.s included in CUBER.**	**1**	**N/A**	**0=material 1=course 2=package 3=programme**	**Vocabulary; see Ontology**	**-**	**All**	**Yes**
1.13	**Date**	**The time span or important dates of the L.O.**	**N**	**-**	**-**	**-**	**-**	**-**	**-**
1.13.1	**Begin**	**The begin date of the L.O.**	**1**	**N/A**	**-**	**Date**	**-**	**All**	**Yes**

1.13.2	**End**	**The end date of the L.O.**	**1**	**N/A**	**-**	**Date**	**-**	**All**	**No**
1.13.3	**Kind**	**The nature of the contribution or action required with regard to the dates announced.**	**1**	**N/A**	**Enrolment Exam period Publishing time Study period Other**	**Vocabulary; see Ontology**	**-**	**All**	**Yes**
1.14	**Title in English**	**Name of the L.O. in English.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**	**Java program-ming**	**All**	**Yes**

Nr	Name	Explanation	Size	Order	Value Space	Data Type	Example	Aggregation levels	Mandatory
2	LifeCycle	This category describes the history and current state of this L.O. and its contributors.	1	N/A	-	-	-	-	-
2.1	Version	The edition of this L.O.	1	N/A	-	Langstring (smallest permitted max: 50 char)	3.0, 3.1, 1.2 alpha	Material Course	No

2.2	Status	The state or condition of this L.O.	1	N/A	Draft Final Revised Archived Unavailable	Vocabulary; **see Ontology**	-	All	Yes
2.3	Contribute	This element describes the people and organisations that have affected the state of this L.O.	Smallest permitted maximum 30 items	No	-	-	-	-	-
2.3.1	Role	Kind of contribution. *At least the author(s) of the L.O. should be described*.	1	N/A	**-**Author -Editor **-**Publisher **-**Content provider -Graphical designer -Instructional designer -Initiator -Terminator -Technical implementer -Educational validator -Technical validator -Script writer **-**Unknown	Vocabulary; **see Ontology**	-	Material	Yes

2.3.2	Entity	Information about the people and organisations contributing to this L.O.	Smallest permitted max: 40 items	Yes	Vcard	Characterstring (smallest permitted max: 1000 chars)	-	Material	Yes
2.3.3	*Date*	*The date of the contribution.*	*1*	*N/A*	-	*Date*	-	*NOT USED IN CUBER!*	-
2.4	**Recurrence**	**This element indicates whether the L.O. is unique or repeated periodically.**	**1**	**N/A**	**Repeated One-time study element Every 3 or 6 months Every year Last occasion**	**Vocabulary; see Ontology**	**-**	**All**	**No**

Nr	Name	Explanation	Size	Order	Value Space	Data Type	Example	Aggregation levels	Mandatory
3	MetaMetadata	This category describes the M.D. record itself.	1	N/A	-	-	-	-	-
3.1	Identifier	A globally unique label that identifies this M.D. record.	1	N/A	-	Reserved	-	-	-

3.2	CatalogEntry	This element describes an entry within a catalog given to the M.D. instance.	Smallest permitted max: 10 items	No	-	-	-	-	-
3.2.1	Catalog	The name of the catalog (i.e. listing identification system).	1	N/A	Repertoire of ISO/IEC 10646-1	Characterstring (smallest permitted max: 1000 char)	ARIADNE	All	No
3.2.2	Entry	Actual string value of the entry in the catalog.	1	N/A	-	Langstring (smallest permitted max: 1000 char)	(en,KUL532)	All	No
3.3	Contribute	Describes the people and organisations that have affected the state of this M.D.	Smallest permitted max: 10 items	Yes	-	-	-	-	-
3.3.1	Role	Kind of contribution	1	N/A	Creator Validator	Vocabulary, **Ontology**	-	All	Yes
3.3.2	Entity	Information about the people and organisations contributing to this M.D. instance.	Smallest permitted max: 10 items	Yes	Vcard	Characterstring (smallest permitted max: 1000 char)	-	All	Yes

3.3.3	Date	The date of the contribution.	1	N/A	-	Date	-	All	Yes
3.4	Metadata Scheme	The name and version of the authoritative specification used to create this M.D. instance.	Smallest permitte d max: 10 items	No	Repertoire of ISO/IEC 10646-1	Characterstrin g (smallest permitted max: 30 char)	LOM-1.0	All	Yes
3.5	Language	Language of this M.D. instance. Default value for all the Langstring values in this M.D. instance.	1	N/A	LanguageID= Langcode (´-´Subcode)*, ISO 639, ISO 3166	Characterstrin g (smallest permitted max: 100 char)	"en" **Default in CUBER is English**.	All	Yes

Nr	**Name**	**Explanation**	**Size**	**Order**	**Value Space**	**Data Type**	**Example**	**Aggregation levels**	**Mandatory**
4	Technical	This category described the technical requirements and characteristics of this L.O.	1	N/A	-	-	-	-	-
4.1	Format	Technical data type(s) of this L.O.	Smallest permitte d max:	No	MIME types based on IANA	Characterstring (smallest permitted max:	Video/mpg , Text/html,	Material	Yes

			40 items		registration (see RFC2048) or "non-digital"; **see Vocabulary & Ontology**	500 char)	Application /x-toolbook, Book		
4.2	*Size*	*The size of the digital L.O. in bytes.*	*1*	*N/A*	*ISO 646, but only the digits ´0´…´9´*	*Characterstring (smallest permitted max: 30 char)*	-	*NOT USED IN CUBER!*	-
4.3	Location	A string to access this L.O. Physical location of the L.O. Exact location or method of locating.	Smallest permitted max: 10 items	Yes	Repertoire of ISO/IEC 10646-1	Characterstring (smallest permitted max: 1000 char)	Http://host/id	All	No
4.4	Requirements	Describes the technical capabilities required to use this L.O.	Smallest permitted max: 40 items	No	-	-	-	-	-
4.4.1	Type	The technology required to use this L.O.	1	N/A	Operating system Browser	Vocabulary; **see Ontology**	-	Material Course	No
4.4.2	Name	Name of the technology required to use this L.O.	1	N/A	If Type =Operating system, then: PC-DOS MS-Windows	Vocabulary; **see Ontology**	-	Material Course	No

		<u>Note1:</u> The value of this element can be derived from 4.1 Technical.Format automatically.			MacOS Unix Multi-OS None If Type=Browser, then: Any Netscape Internet Explorer Opera				
4.4.3	Minimum Version	Lowest possible version of the required technology to use this L.O.	1	N/A	Repertoire of ISO/IEC 10646-1	Characterstring (smallest permitted max: 30 char)	-	Material Course	No
4.4.4	*Maximum Version*	*Highest version of the technology known to support the use of this L.O.*	*1*	*N/A*	*Repertoire of ISO/IEC 10646-1*	*Characterstring (smallest permitted max: 30 char)*	-	*NOT USED IN CUBER*	-
4.5	*Installation Remarks*	*Description of how to install this L.O.*	*1*	*N/A*	-	*Langstring (smallest permitted max: 1000 char)*	-	*NOT USED IN CUBER*	-
4.6	*Other*	*Information*	*1*	*N/A*	-	*Langstring*	-	*NOT USED IN*	-

	Platform Requirements	*about other software and/or hardware requirements.*				*(smallest permitted max: 1000 char)*		*CUBER*	
4.7	*Duration*	*Time a continuous L.O. takes when played at intended speed.*	*1*	*N/A*	*-*	*Date*	*-*	*NOT USED IN CUBER*	*-*
4.8	**Material Size**	**Size of a digital or a non-digital L.O.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**	**150 pages 11700 words 50 kB 1H 15Min**	**Material**	**No**
4.9	**Description**	**Further description on the technical characteristics of the L.O.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**	**Guidelines and commands for using Unix.**	**All**	**No**

Nr	Name	Explanation	Size	Order	Value Space	Data Type	Example	Aggregation levels	Mandatory
5	Educational	This category describes the key educational and pedagogical characteristics of	1	N/A	-	-	-	-	-

		this L.O.							
5.1	*Interactivity Type*	*The flow of interaction between the learner and this L.O.*	*1*	*N/A*	*Active Expositive Mixed Undefined*	*Vocabulary*	*Active: Exercises Expositive: Documents*	*NOT USED IN CUBER*	-
5.2	Learning Resource Type	Specific kind of L.O. NOTE: This element will be used for <u>material-level only</u>; it makes no sense to use this element for other aggregation levels with this vocabulary.	Smallest permitted max: 10 items	Yes	Exercise Simulation Questionnaire Diagram Figure Graph Index Slide Table Narrative text Exam Experiment Problem-statement Self assessment	Vocabulary; **see Ontology**	-	Material	No
5.3	*Interactivity Level*	*The degree of interactivity between the learner and the L.O.*	*1*	*N/A*	*Very low Low Medium High Very high*	*Vocabulary*	-	*NOT USED IN CUBER*	-
5.4	*Semantic Density*	*Amount of information conveyed by this L.O. as compared to its size or duration.*	*1*	*N/A*	*Very low Low Medium High*	*Vocabulary*	-	*NOT USED IN CUBER*	-

		its size or duration.			*Very high*				
5.5	*Intended End User Role*	*Principal users for whom this L.O. was designed.*	*Smallest permitted max: 10 items*	*Yes*	*Teacher Author Learner Manager*	*Vocabulary*	-	*NOT USED IN CUBER*	-
5.6	Context	The principal environment within which the use of this L.O. is intended to take place. **The original LOM vocabulary has been replaced by CUBER´s own vocabulary.**	Smallest permitted max: 10 items	No	**DL0 General studies DL1 Basic, Bac. DL2 Intermediate, Bachelor DL3 Advanced, Master DL4 Post-graduate, L/D DL5 Vocational, further education**	Vocabulary **See Ontology**	-	Course Package Programme	Yes for Course, Package and Programme
5.7	*Typical Age Range*	*Age of the typical user of this L.O.*	*Smallest permitted max: 5 items*	*No*	-	*Langstring (smallest permitted max: 1000 char)*	-	*NOT USED IN CUBER*	-
5.8	*Difficulty*	*How hard it is for the target audience to work through this L.O. in relation to the educational level*	*1*	*N/A*	*Very easy Easy Medium Difficult Very difficult*	*Vocabulary*	-	NOT USED IN CUBER	-

5.9	Typical Learning Time	Typical time it takes to work through this L.O. (e.g. hours, days, weeks, months)	1	N/A	-	Time, Date	**Used in CUBER only when ECTS not available.**	All	No
5.10	*Description*	*Comments on how this L.O. is to be used.*	*1*	*N/A*	-	*Langstring (smallest permitted max: 1000 char)*	*Teacher guidelines that come with a textbook.*	*NOT USED IN CUBER*	-
5.11	*Language*	*The human language used by the target group of this L.O.*	*Smallest permitted max: 10 items*	*No*	*LanguageID =Langcode (´-´Subcode)*, ISO 639, ISO 3166*	*Characterstring (smallest permitted max: 100 char)*	*"en" "en-GB" "de" "fr-CA" "it"*	*NOT USED IN CUBER*	-
5.12	**Teaching Activity**	**Description of principal teaching activities used for this L.O.**	**1**	**N/A**	**-**	**-**	**-**	**-**	**-**
5.12.1	**Teaching Method**	**This sub-element describes the principal teaching method used for this L.O.**	**1**	**N/A**	**Face-to-face** **Distance (www-based)** **Distance (independent)** **Mixed face-to face and distance** Undefined	**Vocabulary; see Ontology**	**-**	**Course** **Package** **Programme**	**Yes**
5.12.2	**Dependence on time**	**This sub-element**	**1**	**N/A**	**Given schedule**	**Vocabulary; see Ontology**	**-**	**Course** **Package**	**Yes**

		describes the dependence on time of this L.O.			Negotiable schedule No time-restrictions Undefined			Programme	
5.12.3	Dependence on place	This sub-element describes the dependence on place of this L.O.	1	N/A	Given place Negotiable place No place-restrictions Undefined	Vocabulary; see Ontology	-	Course Package Programme	Yes
5.13	ECTS Credits	This element describes the ECTS credits of this L.O.	1	N/A	-	Characterstring (smallest permitted max: 30 char)	-	Course Package Programme	Yes for Course
5.14	Dedication	This sub-element describes how intensively the learner must work.	1	N/A	Part-time Full-time Mixed (part&full) No time limits Undefined	Vocabulary; see Ontology	-	Course Package Programme	No
5.15	Evaluation	This element describes the principal method(s) and amount of evaluation for this L.O.	1	N/A	-	-	-	-	-
5.15.1	Assessment	This sub-element describes the	1	N/A	Formal assessment Informal	Vocabulary; see Ontology	-	Course Package	No

		assessment related to this L.O.			**assessment** **Final assessment** **Continuous assessment** **Several assessments** **No assessment** ***Undefined***				
5.15.2	**Method**	**The principal method(s) of assessment for this L.O.**	**N**	**No**	**Exam with attendance** **Electric exam in distance** **Exercises** **Assignment** **Participation** **Presentation** **Essay** **Seminar paper** **Portfolio** ***Undefined***	**Vocabulary; see Ontology**	**-**	**Course Package**	**No**
5.15.3	**Number**	**The number of tasks or exams that form the basis for evaluation for this L.O.**	**1**	**N/A**	**ISO 646, but only digits ´0´…´9´**	**Characterstring (smallest permitted max: 30 char)**	**-**	**Course Package**	**No**
5.16	**Enrolment**	**This element contains information on**	**1**	**N/A**	**-**	**Langstring (smallest permitted**	**-**	**Course Package Programme**	**No**

		the enrolment, e.g. method of enrolment.				**max: 1000 char)**			
5.17	**Study Guidance**	**This element describes the guidance or tutoring provided for the learner.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**	**-**	**Course Package Programme**	**No**
5.18	**Pre-requisites**	**This element describes the skills required in order to take this L.O.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**	**Good computer literacy; Fluent German.**	**Course Package Programme**	**No**
5.19	**Degree**	**The official degree related to this study element.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**		**Course Package Programme**	**Yes for Programme**

Nr	Name	Explanation	Size	Order	Value Space	Data Type	Example	Aggregation levels	Mandatory
6	Rights	This category described the intellectual property rights and the conditions of use for this L.O.	1	N/A	-	-	-	-	-

6.1	*Cost*	*Whether use of this L.O. requires payment.*	*1*	*N/A*	*Yes* *No*	*Vocabulary*	-	*NOT USED IN CUBER*	**-**
6.2	Copyright	Copyright and other restrictions on the use of this L.O.	1	N/A	Yes No	Vocabulary	-	Material	No
6.3	Description	Comments on conditions of use of this L.O.	1	N/A	-	Langstring (smallest permitted max: 1000 char)	-	Material	No
6.4	**Cost in EURO**	**The amount of payment in EURO.**	**1**	**N/A**	**-**	**Characterstring (smallest permitted max: 100 char)**	**100 EURO**	**All**	**Yes**
6.5	**Financing**	**The financing possibilities or grants available for the learner.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**	**-**	**All**	**No**

Nr	**Name**	**Explanation**	**Size**	**Order**	**Value Space**	**Data Type**	**Example**	**Aggregation levels**	**Mandatory**
7	Relation	This category describes the relationships	Smallest permitte d max:	No	-	-	-	-	-

		between this L.O. and other L.O.s	100 items						
7.1	Kind	Nature of relationship between this L.O. and the target L.O. <u>NOTE1</u>: Relations based on Dublin Core.	1	N/A	IsPart **HasPart** IsVersionOf HasVersion IsFormat Of HasFormat References IsReferencedBy IsBasedOn IsBasisFor **Requires** IsRequiredBy	Vocabulary; **See Ontology**	**<u>NOTE2:</u> HasPart and Requires are used in CUBER. Optional, Additional, Compulsory and Exchangeable can be used too, but only in free text.**	All	Yes for package and programme
7.2	Resource	The target L.O. that this relationship references.	1	N/A	-	-	-	-	-
7.2.1	Identifier	Unique identifier of the target L.O.	1	N/A	-	Reserved	Not used.	-	-
7.2.2	*Description*	*Description of the target L.O.*	*1*	*N/A*	-	*Langstring (smallest permitted max: 1000 char)*	-	*NOT USED IN CUBER*	-
7.2.3	CatalogEntry	Defines an entry within a catalog	Smallest permitte	No	-	-	-	All	Yes for package and

		assigned to this L.O.	d max: 10 items						programme
7.3	**Dependencie s**	**Description of dependencies between the study elements.**	**1**	**N/A**	**-**	**Langstring (smallest permitted max: 1000 char)**	**-**	**Course Package Programme**	**No**

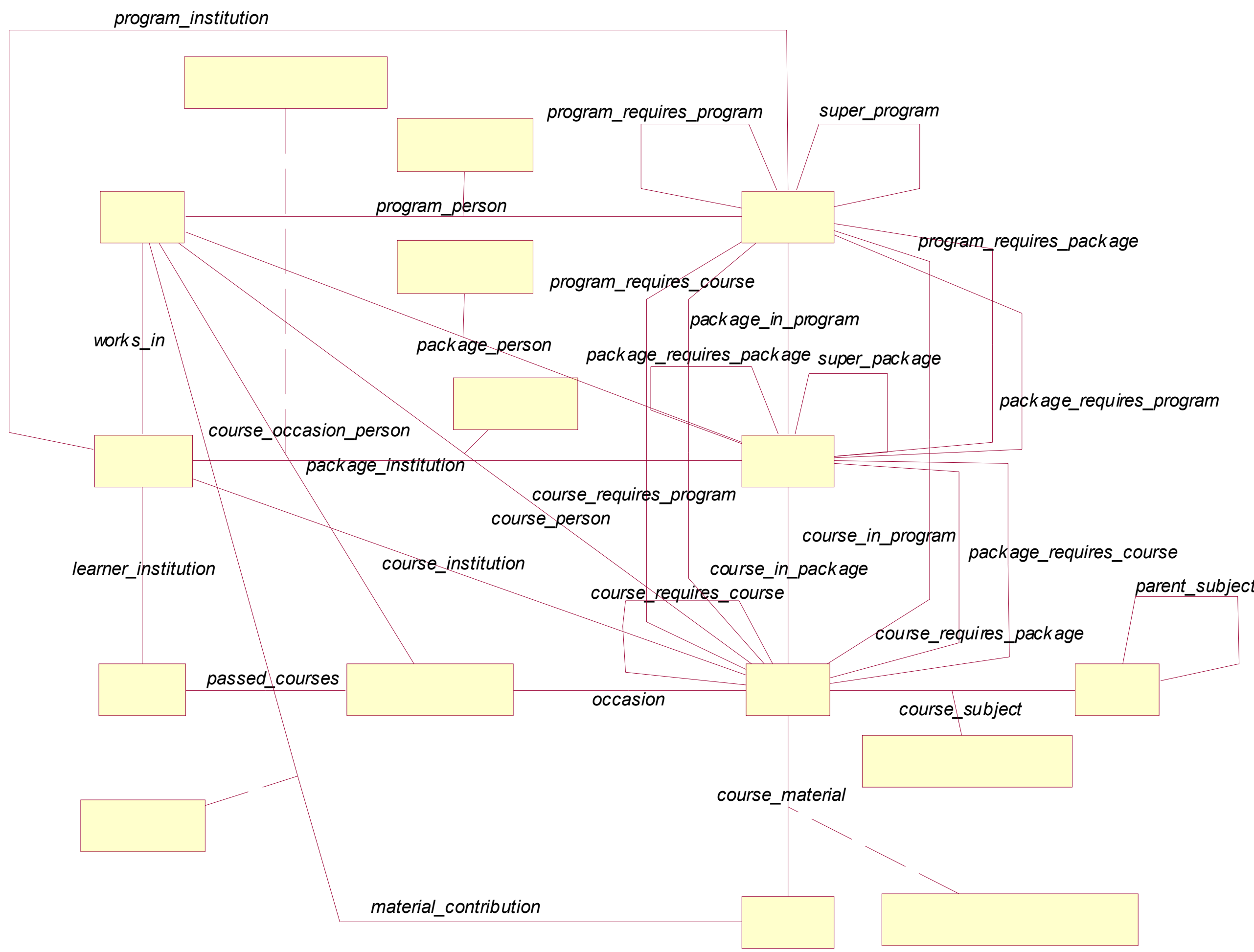
Nr	Name	Explanation	Size	Order	Value Space	Data Type	Example	Aggregation levels	Mandatory
8	Annotation	This category provides comments on the educational use of this L.O.	Smallest permitte d max: 30 items	No	-	-	-	-	-
8.1	Person	The person who created this annotation.	1	N/A	V-card	Characterstring (smallest permitted max: 1000 char)	**Informatio n from the V-card will be defined separately.**	All	No
8.2	Date	Date this annotation was created.	1	N/A	-	Date	-	All	No
8.3	Description	The content of this annotation.	1	N/A	-	Langstring (smallest permitted max: 1000 char)	-	All	No

Nr	Name	Explanation	Size	Order	Value Space	Data Type	Example	Aggregation levels	Mandatory
9	Classification	This category describes where this L.O. falls within a particular classification system.	Smallest permitted max: 40 items	No	-	-	-	-	-
9.1	Purpose	The purpose of classifying this L.O.	1	N/A	Discipline Subject Idea Prerequisite Educational objective Accessibility restrictions Educational level Skill level Security level	Vocabulary	**Discipline and Subject will be used in CUBER.**	Course Package Programme	Yes
9.2	TaxonPath	This element describes a taxonomic path in a specific classification	Smallest permitted max: 15 items	No	-	-	-	-	-

		system.							
9.2.1	Source	The name of the classification system.	1	N/A	Repertoire of ISO/IEC 10646-1	Langstring (smallest permitted max: 1000 char)	(en, ACM) (en, ARIADNE)	Course Package Programme	Yes
9.2.2	Taxon	This element describes a particular term within the taxonomy.	Smallest permitte d max: 15 items	Yes	-	-	-	-	-
9.2.2. 1	Id	The identifier of the Taxon.	1	N/A	Repertoire of ISO/IEC 10646-1	Characterstring (smallest permitted max: 100 char)	320, 4.3.2, BF180	Course Package Programme	Yes
9.2.2. 2	Entry	The textual label of the Taxon.	1	N/A	-	Langstring (smallest permitted max: 500 char)	(en, Medical sciences)	Course Package Programme	Yes
9.3	*Description*	*This is the description of the L.O. relative to the Classification Purpose (9.1) of this classification.*	*1*	*N/A*	-	*Langstring (smallest permitted max: 2000 char)*	-.	*NOT USED IN CUBER*	-
9.4	*Keywords*	*The keywords descriptive of the L.O. relative to the Classification*	*Smallest permitte d max: 40 items*	*Yes*	-	*Langstring (smallest permitted max: 1000 char)*	-	*NOT USED IN CUBER*	-

		Purpose (9.1) of this specific classification.							

Appendix 2. Example of the conceptual model in CUBER



Appendix 3. Example of the SQL Definition

```
CREATE TABLE Course_occasion_role (
  role varchar(20) CHECK (role IN ('Administrator','Assistant','Contact
person','Examiner','Lecturer','Teacher','Tutor','Advisor')),
  username varchar(10),
  start_date date,
  course_identifier varchar(15),
  PRIMARY KEY (role, username, start_date, course_identifier),
  FOREIGN KEY (username)
    REFERENCES Person,
  FOREIGN KEY (start_date)
    REFERENCES Course_occasion
  FOREIGN KEY (course_identifier)
    REFERENCES Course
);

CREATE TABLE Course_role (
  role varchar(20) CHECK (role IN ('Administrator','Assistant','Contact
person','Examiner','Lecturer','Teacher','Tutor','Advisor')),
  username varchar(10),
  course_identifier varchar(15),
  PRIMARY KEY (role, username, course_identifier),
  FOREIGN KEY (username)
    REFERENCES Person,
  FOREIGN KEY (course_identifier)
    REFERENCES Course
);

CREATE TABLE Package_role (
  role varchar(20) CHECK (role IN ('Administrator','Assistant','Contact
person','Examiner','Lecturer','Teacher','Tutor','Advisor')),
  username varchar(10),
  package_identifier varchar(15),
  PRIMARY KEY (role, username, package_identifier),
  FOREIGN KEY (username)
    REFERENCES Person,
  FOREIGN KEY (package_identifier)
    REFERENCES Package
);

CREATE TABLE Program_role (
  role varchar(20) CHECK (role IN ('Administrator','Assistant','Contact
person','Examiner','Lecturer','Teacher','Tutor','Advisor')),
  username varchar(10),
  program_identifier varchar(15),
  PRIMARY KEY (role, username, program_identifier),
  FOREIGN KEY (username)
    REFERENCES Person,
  FOREIGN KEY (program_identifier)
    REFERENCES Program
);
```

Appendix 4. Feedback Form for the CUBER Metadata Information Model

Background information:

- Your email address (optional):
- Your country:
- Organisation type: University/Open university/Distance teaching university/Polytechnic/College/Vocational training/Other
- Knowledge domain (e.g. biology):
- Do you have previous experience in entering study descriptions to a similar system? Yes/No
  - If yes, which system?
- Is there a specific field of knowledge that you or your organisation would like to find from or enter into CUBER (e.g. ICT, Marketing, Medicine, etc.)?

General questions about the Add/Edit pages

Select the option that is closest to your opinion.

Grouping of element (‘Basic Information’, ‘General Information’ etc.) is
Consistent O O O O Confusing

The order in which the information in presented is
Consistent O O O O Confusing

Did you find/use the descriptions (‘Lexicon’ pop-up window) provided of elements? Yes/no?

How did you understand the meaning of the elements? Which was the most important quidance for you in general?

Element titles	Very important	O	O	O	O	Not important
Definitions	Very important	O	O	O	O	Not important
Examples	Very important	O	O	O	O	Not important

Comments about understanding the meanings of the elements:
Suggestions for improving the descriptions are also welcome.

Questions about Metadata details

From the viewpoint of your organisation, how important do you consider the following information? You may select more than one option per row.

	Important	Usable	Not usable	Not relevant
Title in English				
Original title				
Language				
Keywords				
Version (material and course)				
Description in English				
Original description				

Web address (URL)				
Catalogue entry				
Cost				
Financing possibilities				
Structure (material)				
Published (material)				
Enrolment period				
Enrolment description				
Study period				
Exam period				
Key persons				

Degree (package and programme)				
Subject				
Discipline				
Difficulty				
ECTS Credits				
Typical learning time				
Teaching method				
Dependence on time				
Dependence on place				
Dedication				
Study guidance				
Assessment				
Evaluation method				
Recurrence				

Operating system				
Browser				

Description of technical requirements
Copyright (material)
Conditions of use (material)
Included material (course)
Pre-requisites
Other pre-requisites
Other study element dependencies
Included packages or courses (packages and programme)

Status
Availability status of the record

If you had aby answer in ‘Not usable in this format’, please comment:
Suggestions of improvement are also welcome.

Questions about training and guidance

Finding the Guide section was
Easy O O O O Difficult

Finding the needed information from the Guide section was
Easy O O O O Difficult

Which guidance options did you use?
-Printable ‘User Instructions’ document
-On-line instructions, e.g. ‘Common tasks’
-Site map
-CUBER Lexicon
-Help

Did the guidance help you with the problem?
Yes, perfectly O O O O No, I couldn’t solve the problem

Comments about these Guidance options and their use:

General Feedback

-Are there any features missing in this system you would need?
-Have you had some special experiences with the system you’d like to let us know?
-Please leave your contact information if you’d like us to reply to you on these issues.