An Ontology Framework for e-Learning in the Knowledge Society

Abstract: Efficient knowledge management is essential within the information society. Life long learning as well as the use of new media have lead to e-Learning of different kinds. In order to combine existing resources, a general description of this topic is needed. The semantic web aims at making these meta data machine understandable. In this paper we present our Ontology Framework for e-Learning. After the introduction we review existing approaches and describe our general view of the concepts. In chapter 4 and 5 we present different views of our framework aimed at the intended application areas as material or user centred approaches and end up with the conclusions.

Introduction

E-Learning has become an important topic during the last few years. E-Learning can also be seen as knowledge generation, dissemination and (re)use. The knowledge (based) society is a common term now. More and more information is being published, in particular in the scientific world.

Bricks-and-mortar university libraries have been providing literature for studying and learning for a long time, digital libraries and other content providers now offer additional material for e-Learning. Software companies want to sell their platforms and publicly founded institutions have the mission to educate. The private individual has to ensure life long learning.

At first, legacy learning material has been digitized, e.g. scripts were scanned and an electronic book shelf was offered to a course. But the new media also require new didactics dealing with blended learning, different channels of moderation and presentation [2,7]. This is a time consuming and cost intensive task. Experts have to take the decision: what is relevant for a subject and how to classify it. Libraries use their classification schemes for book indexing, but scientists may use other keywords in their field when they search for information. A student not familiar with the terminology easily gets lost in the wealth of information.

In order to overcome the limitation of one perspective and to relate the content with outside resources, the semantic web offers the functionality of machine understandable meaning. Ontologies are explicit visualizations of concepts in a given context and can be expanded. But at the moment only a few ontologies [8] exist based on the formal logic criteria of computer science. User aspects are also neglected within this development.

Meta Data and the Semantic Web

As the amount of information in the Internet is rapidly increasing, meta data are more and more necessary to classify and describe it. In this paper we lay out different ordering schemes which are widely used. We start chronologically with one meta data set used by libraries, go on with another used by scientific authors, following by two integrated approaches and try to name the component suitable for e-Leaning material. At the end we will discuss ontologies.
Meta data schemes

Digital libraries provide their users mainly with online customer services, like a web interface of OPAC for retrieval, account settings like the extension of the loan period, reservation of books etc. Getting the material itself mostly requires one’s physical presence.

**MARC21** (MAchine-Readable Cataloguing) was introduced in the 60ies for libraries and became standard later on. The main purpose was the allocation of a book in the shelves of the library and the conversion of the catalogue cards in an electronic form. The main bibliographic data are given in a specific order and describe 8 types of material: **Book**, **Continuing resource**, **Computer files**, **Maps**, **Music**, **Sound recording (non music)**, **Visual materials**, **Mixed materials**. For the classification of e-Learning 5 types of material: **continuing resource**, **computer files**, **visual/mixed material**, **maps** and 3 types of record, namely **manuscript**, **computer file**, **cartographic material** can be used out of these sets.

**BibTeX** was designed by Patashnik and Lamport in 1985 as the LaTeX bibliographic format. LaTeX is an open source document preparation system widely used in the academic community. The authors provide their references/citation of other publications entirely character based, so that it can be shared by the community on the Internet. The type of publication can be classified according to 12 different categories: **String**, **Book**, **InBook**, **InCollection**, **InProceeding**, **Proceeding**, **Article**, **MasterThesis**, **PhDThesis**, **TechReport**, **Manual**, **Misc**. As reference to e-Learning material, **Misc** and **InCollection** would be used here.

But in the era of the WWW everybody can become a provider. The **Dublin Core** (DC) Initiative has been started 10 years ago by librarians in order to provide a meta data standard that support a broad range of purposes and business models. Nowadays research communities, corporate knowledge management, e-government and public sector information use 15 core elements for their objectives.

The core elements are: **Title**, **Creator**, **Subject**, **Description**, **Publisher**, **Contributor**, **Date**, **Type**, **Format**, **Identifier**, **Source**, **Language**, **Relation**, **Coverage** and **Rights**.

Resources should be classified by the normal user and should be made accessible over the Internet for online retrieval. For ‘type’ the following 12 subclasses are provided: **Collection**, **Dataset**, **Event**, **Image**, **MovingImage**, **StillImage**, **Sound**, **Text**, **Interactive Resource**, **Physical Object**, **Service** and **Software**, out of which maybe collection, interactive resource and software are useful for e-Learning as an extension of traditional material like books.

As it can be seen from the examples above, identifying especially e-Learning material in the web is not an easy task. For a more detailed description of these sources **DINI** the German Initiative for Network information came up with a recommendation by different experts. Table 1 lists the 21 elements separated for course and content. First the mandatory fields are assigned to course and put in the second column, the equivalent or identical field for content is put in the lat column in the same row. Note that duration of the content in line 14 is only optional, whereas (Technical) Requirements are mandatory in contrast to course. In line 15 you find type vs. format which are somehow close together, ECTS points and Copyright (17), Classification and Memory size (19) of course do not refer to each and only stand in line for schematic reasons.
Table 1: The DINI e-Learning meta data elements [9] m: mandatory, o: optional

The Learning Object Meta data (LOM) by the IEEE working group 12 provide a more detailed scheme especially for e-Learning material consisting of the following 19 elements: Source, Structure, Aggregation Level, Status, Role (Meta) Role, Type, Name, Interactivity Type, Learning Resource Type, Interactivity Level, Semantic Density, Intended End User Role, Context, Difficulty, Cost, Copyright and Other Restrictions, Kind, Purpose.

Ontologies and the semantic web

Ontologies are the ‘new’ semantic description and being formalized in computer science. Coding and reasoning are the main focus at the moment, the real applications are still missing, though. This needs the input of the communities and should include different tasks and aspects in various areas. Ontology tools, e.g. editors are missing to aid coding the intellectual input of the fields in the correct way.

One definition in computer science: a shared conceptualisation can be given more formally by: An ontology $O$ is a 4-tuple $<C; R; I; A>$, where $C$ is a set of concepts, $R$ is a set of relations, $I$ is a set of instances and $A$ is a set of axioms.

Ontologies have to be implemented as networks of meaning, where from the very beginning heterogeneity is an essential requirement. Expert ontologies form a specific part and should be consistent within the community. So far some initiatives and projects in the academic field come up with the first stable versions in OWL which are documented as well. These are

- **FOAF**\(^1\) (Friend Of A Friend) for communities
  Describing homepages of people, the links between them and the things they create and do
  Coverage: Person as 1 concept with 10 properties
- **SWP**\(^2\) (Semantic Web Portal) Ontology for Scientific Portals
  Coverage: Person (Agent and Organisation), Publication, Conference as the 3 main concepts
  68 classes, 21 data properties, 57 object properties
MarcOnt³ for Digital libraries

Ongoing and still under construction, do mapping of the three formats Marc21, DC, BibTeX

Semantic digital libraries try to integrate user scenarios like shared bookmarks, comments on reading and new forms like web logs and so on. This stresses the community aspects and also provides some feedback loops.

Big commercial e-Learning Platforms like WebCT Vista and Blackboard offer the technical perspective with specific features like forums, assessments and workflow. They are course centred 0and not interoperable. The closed area allows no collaboration between universities using different software. Documentation of interfaces is not officially available, they also lack an import facility of existing meta data and ontologies. The data description is only done at an internal basis.

To close the gap we invite all players to contribute to our interdisciplinary project. We start with the integration of existing schemes, elements and ontologies. Service description and definition will follow. Hereby user and provider can bridge their demand and needs.

The Ontology Framework

In this chapter we present our Ontology Framework for e-Learning. We start with the general consideration of the network and its components. We then present the course-centred view which is the common perspective in literature at the moment. Platforms are currently grouped around a course or topic but they are not user centred [1]. They are using different software and provide a variety of functionalities. Material is hopefully described by some meta data or classified by a whole scheme like DC or LOM [3, 9]. The didactics depends on the material and platform chosen.

After that we switch to the user-centred view as the key feature of the semantic web and take it from there. In the last part we mention the application areas we have in mind and how they fit into our considerations so far.

General View

Within our framework we focus on nine main concepts. The top level ones are: Topic = Subject, Person, Didactics, Platform and Material. Fig.1. shows the semantic network of our first version. The main concepts are in ellipse, subclasses are marked in boxes and HasAttributes are the more detailed relations. The lines do not represent all attributes completely.

Person is also a subclass of Agent in reference to the SWP ontology in chapter 2 and can take several roles. S/he can be a user or provider, student or lecturer, pupil or teacher but also the administrator or a technician. Author is also a key player and an important term in reference to DC:creator. Student will be the main figure in the user-centred view.

Material is part of a Course in correlation with the DINI set and the course-centred view. Learning Objects include scripts, test, book chapters but also multimedia content, instructional content, learning objectives, instructional software and software tools.


Figure 1: Ontology concepts, subclasses and relations for e-Learning
main concepts are in ellipses, subclasses in boxes, HasAttributes are the relations

Course-centred view
The course-centred view shows the perspective of providers and creators of e-Learning platforms. Providers care for the smoothly technical usage of e-Learning platforms. Since providers of (open-source) platforms often enhance the software and infrastructure, e-Learning is presented by the courses and materials forming the centre.

Often, no dedicated didactical concept is associated with the platforms. Human interaction and intellectual input is needed to keep the technology alive. Users are authors and teachers. They do not need to be the same person, a teacher can easily (re)use other authors’ materials. Authors create material, which may be used in different courses with different didactical concepts. Teachers order own and external material and form it into a course, using a didactical concept. The didactics depends on the level and the subject of the individual course.

Courses are then taken by students, chosen by level and subject. Depending on the concept and the didactics, students are able to interact with the course and the material. This may be very passive, like only downloading files from the platform or simply from a homepage without any other action. But they might also be active in a forum, give feedback to the teachers and/or the authors who then might change or adjust their material. Students may also contribute with their own material and documents to the platform.

It would be a real enhancement for courses and materials, or for overall teaching, if independent referees would evaluate them. Unfortunately, this is only commonly accepted for scientific publications, today. In the last years though, students started to evaluate their courses, supported by their university.

To be able to use e-Learning offering, students and teachers need support from their institution, e.g. their university. They need support for the platform, but also the technical
infrastructure like computer pools and projectors, (W)LAN and more. Both groups depend on further education as well, like media training or the usage of authoring tools. A modern university should be able to provide all of that.

Often, the platform is offered by the university as well. In some cases though, private companies are responsible for the provisioning and the support of e-Learning platforms.

Furthermore, libraries have started to collect, provide and archive not only print media, but also digital media. As a consequence libraries catalogue and archive e-Learning courses and materials from e-Learning platforms. For this, special sets of meta data are needed. Mostly, these meta data are based on LOM, Dublin Core or PICA.

User-centred view
Starting with the user the following view is provided. A person with the main role student has already knowledge and experience. So a certain level and subject can be defined. Attending a new course will focus on improving specific skills for further education. To use material a technical infrastructure is needed. Detailed descriptions will help to fulfil the requirements and to choose the best fitting.

For the user, accessibility is the important part. S/he will not care about didactics so far as her/his needs are satisfied and the necessary information is delivered out of an information pool. Using a personal profile can enhance semantic based retrieval and gain best individual results. In the future software agents can also provide specific offers by regularly checking all available background information.

Application aeras: LiLi and HU
We plan to test our framework on two examples: First LiLi [10], a free collection of online material for physics [4] and second the library science department of the Humboldt University of Berlin [11] which offer distance postgraduate studies. At the moment both platforms are material/course-centred.

LiLi is a collection consists of links to e-Learning material for physics which is described by a set of meta data, comments, and ratings. It can be used by students for self-study. Trying to understand a certain topic in physics, students may search for scripts and visualisations in LiLi. Lecturers may search in LiLi in order to complete their lectures with visualisations or examples. Both groups use the same meta data and comments for the search and the choice of appropriate material.

While using LiLi, both groups are also asked to give comments and ratings on the individual entries. By this, community aspects are taken into account which provides useful input to further applications and the reuse of the material. Not only comments but also whole entries, consisting of links and meta data to online material can be inserted into LiLi by teachers and students and also by authors themselves. Only a registration to the system is needed for this.

LiLi is strongly connected to “physik multimedial”, a platform for physics’ courses. A function to add links and comments from LiLi directly into the courses which are offered and organized by this platform, is given. On the other hand, lecturers are requested automatically to inscribe the material they add to their course also to LiLi, completed by meta data.

This means, that from both sides a connection between the courses in the platform and material that is described in LiLi, is given.

We showed that connections between persons, material, topics, meta data, courses and a platform are already given in LiLi and “physik multimedial”. These connections can be coded in formal RDF descriptions. Furthermore, LiLi is an OAI data provider. By this, the information provided by LiLi could easily be retrieved by libraries, archives or portals.
Secondly the library science department of the Humboldt University of Berlin [11] will be the test case for a distance postgraduate study. Here a subscribed platform is offered to the students. The course contains slots in Berlin and online self study units with a final exam and a university degree.

In a next step the evaluation of the framework will help to improve its further development and will be the followed by the formal description and coding. The implementation will hopefully lead to a prototype application in this area.

Figure 2: Screenshot of the e-Learning platform LiLi

Conclusion

In this paper we presented the theoretical concepts of an ontology framework for eLearning and gave background information concerning the knowledge society. We have described the way to coding and implementation of this framework.

Notes
1  http://www.foaf-project.org/
2  http://sw-portal.deri.org/ontologies/swportal.html
3  http://www.marcont.org/
References


[10] LiLi <http://www.physik-multimedial.de/lili/golili/lili.php>

[11] Department of Library and Information Science, Department of Distance Study at the Humboldt University to Berlin <http://www.fbiw.hu-berlin.de/startseite/willkommen_e>

Meta data schemes
DCMI Dublin Core Metadata Initiative <http://dublincore.org/>
MARC Machine-Readable Cataloging record <http://www.loc.gov/marc/>

Ontologies
MarcOnt September 2005 <http://www.marcont.org/>