Proceedings of the 3rd WSEAS/IASME International Conference on ENGINEERING EDUCATION, Vouliagmeni, Greece, July 11-13, 2006 (pp233-238)

# Towards Distributed Learning Organizational Memories

MARIE-HELENE ABEL, CLAUDE MOULIN and DOMINIQUE LENNE
University of Technology of Compiègne, CNRS
Heudiasyc
Compiègne - FRANCE

*Abstract:* This paper presents an analyze of a learning organizational memory and some disadvantages that such a centralized application contains. One issue is the reuse of a prototype giving access to learning resources, outside the university where it was created, by teachers of the same domain, even if they are very interested to do so. One attempt for resolving this problem is to integrate distributed learning object repositories. Two candidate technologies are envisaged and presented here. The main consequence of the openness of a memory to a larger group of teachers coming from different universities willing to share resources is the creation of a community of practice. One important task of this community is then the management of semantics used for indexing resources and in particular the life cycle of ontologies.

*Key-Words:* Learning Objects, Ontology, Organizational Memory, Peer-to-peer, Semantic Web Service, Community of practice.

## 1 Introduction

The development of information technology allows to use a growing number of resources for e-learning. Different approaches may be adopted to exploit these resources. They can be stored in learning objects repositories and then reused, combined and adapted in different contexts. They can also be selected and organized in learning memories that can be directly accessed by learners.

In the MEMORAe project [1], we adopted the latter approach and we designed what we called a "learning organizational memory". We started from the idea that a course can be seen as an organization and we chose to ground our learning organizational memory on two ontologies, one is concerning the notions to learn, the other is generic to each formation. Students can navigate through the former ontology in order to find learning resources adapted to their needs. Our approach proved to be very fruitful. We realized a prototype called E-MEMORAe for two courses, one in "algorithms and programming", the other in applied mathematics. In both cases, the prototype was successfully evaluated in learning situations in our university.

The problem is now to see to what extent our work could be reused in other contexts (e.g. other universities). It appears that the memory reflects the pedagogical choices of teachers or domain experts that have contributed to its realization. For the same kind of course, there is obviously a common kernel from one context to another, but it is difficult to reuse the memory without any modification. Pedagogical teams have never exactly the same point of view on a subject matter. In consequence, they need to modify or adapt the memory.

In order to address this problem, we are now examining the possibility to design a "distributed learning organizational memory". The idea is to consider a network of geographically-distributed pedagogical teams involved in the same kind of course (domain and level). A network of this type can be seen as a community of practice that can benefit of a distributed learning organizational memory. In our view, each team of the community would have its local memory containing its own resources, but also accepting resources from remote repositories. In this case, it seems mandatory that the ontology be shared by the community.

In this paper we first present the centralized approach and the MEMORAe project. Then we examine the requirements for a shift from a centralized to a distributed learning organizational memory. We foresee two types of distributed repositories using either a peer-to-peer network or semantic web services. We conclude by a way to manage cooperatively the ontology among a community of practice.

## 2 Centralized Approach

A learning organizational memory differs from a learning object repository (LOR). Both cases present a centralized approach adopted for managing resources.

### 2.1 Learning Object Repository

Over the last few years, many projects aiming at building bases of sharable and reusable learning resources have been launched. These projects rely on a network of contributors that feed the base with collaboratively controlled resources. Conversely, each contributor can benefit from resources brought by other contributors.

The scope of LOR can be restricted to one or several universities or to all universities of a country; it can also be international. When the expected scope is wide, LOR can be based on a consortium of institutions.

Distinction can be made between learning object repositories, which usually group resources from many subject matters, and "thematic resource bases" that contain resources related to only one domain.

Resources restricted to a particular domain are more homogeneous; in this case resources and associated knowledge can be managed more precisely. Within the MEMORAe project, our goal is to let learners directly access the resources of a course memory.

## 2.2 The MEMORAe Project

Following a knowledge engineering approach, resources are organized in a learning organizational memory based on ontologies [2]. The memory is structured around a course that is seen as an organization. This memory is different from a classical organizational memory because its goal is to provide users with a pedagogical content. This content is the result of two different processes:

- the capitalization of knowledge, information and learning resources relating to the learning context (a course unit),
- a pedagogical work concerning the choice and the organization of this capitalization.

The pedagogical content is composed of the notions to learn, the links between these notions and the learning resources indexed on notions.

MEMORAe relies on two ontologies. The first one (domain ontology) describes the concepts of the « training » domain: users types (tutor, secretary), documents types (book, slides for oral presentation, web page, site, etc.), media types (text, image, audio, video), pedagogical characteristics (activity type).

Pedagogical resources are not organized following the way recommended by the Learning Object Metadata standard in the Educational Category, because we do not agree to associate various activity types like exercise or simulation, with data representation like diagram, figure or graph in the same set. A description of the LOM[1] standard can be found in the document 1484.12.1.

The second ontology (application ontology) specifies the organization of notions which are studied during a course or training session. In the example of an initiation to algorithmic, some notions like data structure or control structure are explained. It is possible, but not mandatory, to consider "tree" and "array" as sub-concepts of the concept "data structure" and to define the relation "uses" between the concepts "data structure" and "iterative structures" (in this case they are the domain and range value of this relation).

These ontologies are not used independently; the second one is necessarily attached to the first one. For example, to express that a document is an introduction to data structures we join the two concepts "introduction" and "data structures" that do not belong to the same ontology. Pedagogical relations like "prerequisite" or "uses" that occur between concepts of the application ontology are defined in the domain ontology. However, specific roles can belong to the application ontology.

Notions are not only chosen because they are related to the course unit, but also because they are the result of a reflection on the course itself. For example, in a memory dedicated to the initiation of programming we defined the notions of "loop" and "array" and the relations between these notions.

Resources have to be selected relying on pedagogical goals. The choice of their indexing terms is related to this goal too. It is not an automatic indexing. The course manager is responsible for the relevance of the links. It is not because a document treats of a notion to acquire that it will be necessary indexed by this notion. The choice is explicit, that is to say that the document must have been evaluated and considered sufficiently adapted to the learning of this notion.

Within the framework of MEMORAe we developed a prototype, called E-MEMORAe[2] and we realized two pilot applications to evaluate our propositions. The first one concerns a course on algorithms and programming at the Compiègne University of Technology (France) and the second one concerns a course on applied mathematics at the University of Picardy (France).

Our objectives within E-MEMORAe are to help the users of the memory to acquire the notions of a given course. To this end, the users have to navigate through the application ontology that is related to the course, and to access to the indexed resources thanks to this ontology.

The general principle is at each step, to propose to the learners, either precise information on what they are searching for, or links allowing them to continue their navigation through the memory. There is no need to use the keyboard for formulating a request, even if the environment allows it.

[1] http://ltsc.ieee.org/wg12/index.html

[2] http://www.hds.utc.fr/memorae/

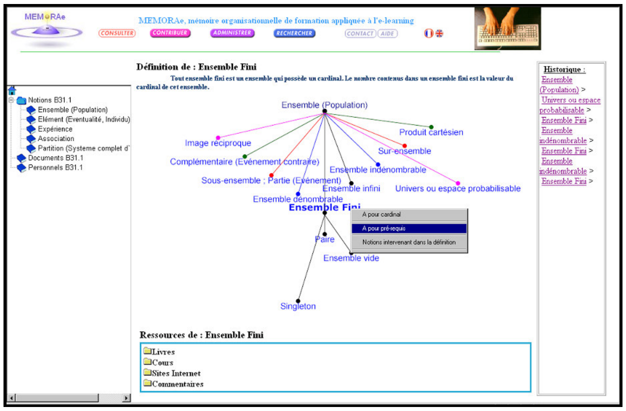

Figure 1: Navigation in the memory.

We defined a usability test[3] in order to see how students use the E-MEMORAe environment. Such a test enables to evaluate learning and memorizing facilities, and the usability of the environment. It also enables to evaluate the types of errors and the satisfaction of the user.

Our objective was to see how E-MEMORAe enables the learners to discover alone new notions to learn. For verifying the understanding of these notions, the learners have to solve some problems and respond to a QCM.

With this test, we can verify the pertinence of our hypothesis on the following points: (i) structuring the content of training by an ontology; (ii) index resources on ontology concepts; (iii) displaying the hierarchy of concepts for facilitating the navigation through the resources; (iv) offering a list of entry points for giving a quick access to the main notions of the course.

The experiments took place at the University of Picardy and the University of Compiègne (France) and were concerning the students attending a master course of statistics or algorithm. Students of the course of statistics were proposed to solve a problem requiring some notions unknown by the students. Students of the course of algorithm were proposed a QCM. They had to use the E-MEMORAe environment to discover the missing knowledge for solving the problem or respond to the QCM.

For each student, the history of the navigation was stored in the memory. We could analyze the way to reach important notions and the resources employed.

After these first experiments, we can conclude that using ontology to index and structure the content of training is a good choice: a majority of students appreciated it. The results obtained by the students show that a majority of them were able to find the indispensable knowledge in a limited time.

[3] http://www.usabilis.com/gb/whatis/usability.htm

## 2.3 Disadvantages of a Centralized Approach

In a collegial and collaborative work environment, academics will typically share their teaching experiences and resources. However, academics often work in isolation and the dissemination of ideas and practices amongst academics who are physically separated across different departments, campuses and institutions becomes problematic. Without easy channels of communication, academics are not always aware of the latest educational innovations and technologies and often fall into the trap of 'reinventing the wheel' [3].

A learning organizational memory should be a solution to this problem but in order to several academics accept to use it, it is necessary to adopt a distributed architecture. Indeed, each academic needs to have the possibility to manage its own resources and to develop its own part of ontology, even if he wants to share some of them.

In the case of a centralized architecture, all resources and knowledge are stored on a server. This one is managed by only one university. The all ontology and indexing resources must be shared by all academics. There are not real possibilities to develop different points of view.

It seems therefore interesting to shift towards a distributed approach.

# 3 Decentralized Memory

People often want to contribute to common projects, even with a slight engagement, but do not want to be overloaded when managing their resources. They generally agree to maintain local knowledge repositories (KR) and to give an open access to their documents. This structure allows knowledge to be managed locally and autonomously where it is created and used. Local indexing also allows several points of view on the same resources present on the net.

We consider two technologies for managing KR that could be loosely linked together. A first approach is to regard peers inside peer-to-peer systems as a way of storing documents. The second one is to encapsulate local resources by web services. The main problem in both technologies remains the organization of semantics.

## 3.1 Peer-to-Peer

A peer-to-peer network is composed of participants (referred as peers or nodes) giving access to personal resources and eventually to non personal resources in case of replication of resources. The main characteristics of peer-to-peer systems are the ability to pool together and harness large amounts of resources, self-organization, load balancing, adaptation and faulttolerance. The pure peer-to-peer systems are distributed systems without any centralized control or hierarchical organization, in which each node runs software with equivalent functionality and responsibilities and all communication is symmetric [4].

Many projects (Gnutella[4], Freenet [5], SWAP[5]) implement this concept and involved different technologies and specific protocols [6, 7]).

When discovering a resources, peer-to-peer applications need to determine the node that stores a data item. A node must either route queries to its neighbors or respond to queries if file is contained locally. Information retrieval is performed by using appropriate algorithms [8] which provide for typical filtering.

Peer-to-peer networks present some issues. Among them the semantics of the net itself, document replication, security and authentication and anonymous queries,.

Pure peer-to-peer systems tend to be inefficient; for example, current search in Gnutella consists of flooding the network with query messages. Another important source of inefficiency is bottlenecks caused by the very limited capabilities of some peers.

For solving the problem of the numbers of nodes to be visited when a query is sent by a peer, recent content addressable networks, such as, Chord [6], Pastry[4], and Tapestry[6], offer an administration-free and fault-tolerant distributed hash table (DHT) that maps "keys" to "values". Section 4.2 proposes a document indexing based on DHT keys.

### 3.2 Web Service

Web services are accessible local applications encapsulating legacy systems. Centralized registries (UDDI[7] or ebXML[8]) store service descriptions allowing eventual clients to discover the functionalities of the services and to invoke them if they are interested with.

In our hypothesis, web services are designed to deliver semantic indexed resources. They could receive a request and return resource Urls. Their installation and management is a little bit more heavy than a peer to peer solution, but common technical approaches could be easily found.

For a client, finding appropriate resources requires two steps: discover the right service and retrieve from the service documents proposed in answer to a particular request. In consequence, services have to be first semantically described, using for example an OWL/S [9] description, i.e. containing elements belonging to an ontology. Then, services have to answer to semantic requests.

In this solution where services play the role of distributed repositories, it seems necessary to design semantic web services and not only basic web services. Another consequence is the maintenance of semantic service registries for storing service descriptions. Two kinds of semantic engines appear mandatory in this model: the registry of services is itself an engine able to select one or several services from a semantic queries and each web service must contain an engine for selecting resources from a request.

Thus, semantics must be separated between general descriptions of services and indexing of resources. Different solutions may be proposed. The most general one is to give the same description for all services. For example, services are described as having the objective of delivering learning objects. Then requests based on RDF[9] indexing of resources (using for example RDQL [10] or SparQL [11] format) are sent to the services. Another solution is to let service providers give a more specific description of their services for a first filtering of their resources.

Semantic engines used for discovering resources generally need adapted algorithm using for example the subsumption of concepts. It is not clear where these algorithms could be implemented: at the level of repositories, inside web services or be reserved to client applications.

## 4 Indexing

### 4.1 Semantic Indexing

In a very technical sense, a document can be seen merely as a collection of words. Document indexing aims to provide access to a document starting from one or more words, and then retrieving all documents in which these words occur**.** Accordingly, it generates the so called inverted indices in which each word points to a collection of documents. Such index is inverted in the sense that instead of presenting a document as a collection of words, it rather presents each word as a collection of associated document links.

Unlike standard keyword indexing designed to generate inverted indices containing every word that occurs in a document, semantic indexing associates documents only with items from a predefined vocabulary, the so called controlled vocabulary. Actually, the term "semantic" is justified only if such controlled vocabulary is provided by ontology, i.e. as labels associated with formal concepts, which are

4 http://en.wikipedia.org/wiki/Gnutella
5 http://swap.semanticweb.org/public/theproject.htm
6 http://p2p.cs.ucsb.edu/chimera/
7 http://www.uddi.org/
8 http://www.ebxml.org/

9 http://www.w3.org/RDF/

defined by a hierarchy, and the relations that hold between them.

Semantic indexing requires to differentiate the concept level and the individual level. We claim that words found in documents have to be considered as labels of individuals, instances of concepts occurring in an ontology. The individuals themselves are not elements of the ontology. For example, if a document contains the word Italy, we would like to index this document on the concept of country associated with the individual whose a label is “Italy”. Italy is a word occurring in the document, Country not.

### 4.2 Indexing in peer-to-peer systems

In this section, we propose a way of indexing resources compliant with recent content addressable networks. We define an index key as a couple containing a concept and a value. This definition lets the user choose the ontology containing the concept. The value is defined as a string associated with a language. We have built an OWL [12] representation of an entry key but for concision reasons we give the following example using a simplified syntax. The example is extracted from the domain of Geography:

Key1: (ex:Z-Country[10] ; :en “italy”)

The concept namespace is given here with the prefix ex representing any ontology and the chosen language (English) is represented by :en.

At a upper level, following this previous definition, for indexing a document concerning some countries, we must see country as an individual of a concept. As country is itself a concept of an ontology, it is then considered as an instance of its meta-class. Using the OWL formalism, this meta-class would be owl:Class. The key corresponding would be:

Key2: (owl:Class ; ex:Z-Country)

An index entry may contain several keys. For example for indicating that a document is concerning the country of Italy and more precisely the region of Sardinia, we can build the following keys:

Key3: (ex:Z-Country ; :en “italy”)

Key4: (ex:Z-Region ; :en “sardinia”)

If the documents were written in Italian we would have:

Key3: (ex:Z-Country ; :it “italia”)

Key4: (ex:Z-Region ; :it “sardegna”)

## 5 Community of Practice

Current Knowledge Management solutions still focuses on centralized knowledge repositories generally organized around ontologies. Dealing with resources in such a way presents many problems concerning costs of maintenance and hardly support dynamic updates.

Manville and Foote [13] define a Community of Practice as a ” a group of professionals informally bound to one another through exposure to a common class of problems, common pursuit of solutions, and thereby themselves embodying a store of knowledge” where knowledge is considered as the capacity to act on information; a capacity which increases with improved community activity.

The concept of Community of Practice has turned out to provide a useful perspective on knowing and learning. A growing number of people and organizations in various sectors are now focusing on communities of practice as a key to improving their performance. Communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly.

Wenger and Lave [14] has articulated the nature of the practices from which the term community of practice derives its name. “…Such a concept of practice includes both the explicit and the tacit. It includes what is said and what is left unsaid; what is represented and what is assumed. It includes language, tools, documents, images, symbols, well-defined roles, specified criteria, codified procedures, regulations, and contracts that various practices make explicit for a variety of purposes.” [15]

Because ontologies reflect a “shared world view”, codifying “well-defined roles”, “specified criteria” and “codified procedures”, they can serve as symbolic tools within a community of practice supporting communication and knowledge sharing [16].

The communities we consider here concerns academics [17] and are generally established in a face-to-face setting. Once an ontology has been constructed by the initial group of the community, it is used to describe web documents. It may appear insufficient after first tests. The community of practice responsible of the evolution of the ontology must define a protocol for insuring a appropriate lifecycle of the ontology [18]. Three roles emerge: domain experts in charge of proposing modifications like teachers specialists of the application domain, ontology experts insuring a coherent representation of the ontology usable by each LOR and the ontology manager in charge of the publication of updated versions of the ontology.

## 4 Conclusion

In this paper we have presented a project which has seen the constitution of an organizational learning memory. We have also presented the prototype which has been implemented and validated in actual situations, confirming the justness of our initial hypothesis.

[10] We write the prefix “Z-“ when it is necessary to give the concept id.

However, we are now confronted to the reuse of our prototype by other groups of teachers. The difficulties are not coming from technical considerations, but from pedagogical choices we had adopted. In order to address this problem, we tried to put in evidence the disadvantages of our modeling and mainly the fact that we have designed a centralized system.

Then we have examined the requirements for introducing a distributed aspect into learning organizational memories. We have foreseen two types of distributed technologies that could be used - peer-to-peer network and semantic web service. Obviously, further researches and experiments are necessary for bringing a satisfying answer to the issue we have presented.

We have not limited our investigation to the sole problem of technology because the issue presents an organizational dimension. It seems that any solution leads to a cooperative and distributed management of the knowledge occurring in the system. Researches on communities of practice seem promising for maintaining the semantic shared elements that such a solution could require and in particular the ontologies.

*References:*

[1] Lenne, D., Abel, M.-H., Benayache, A. and Moulin, C., E-MEMORAe: an e-Learning Environment Based on an Organizational Memory. in *World Conference on Educational Multimedia, Hypermedia and Telecommunications, Ed-Media'05*, (Montréal, Canada), pp. 4651-4658, 2005.

[2] Abel, M.-H., Barry, C., Benayache, A., Chaput, B., Lenne, D., and Moulin, C. Ontology-based Organizational Memory for e-learning. *Educational Technology & Society Journal*, Vol. 7, Issue 4, 2004.

[3] Sheard, J., Carbone, A. From Informal to Formal: Creating the Australian Computing Education Community. Proceedings of the sixth conference on Australian computing education *ACE2004* - Volume 30 CRPIT '04, 2004.

[4] Druschel P., Rowstron A., "PAST: A Large-Scale, Persistent Peer-to-Peer Storage Utility," *hotos*, Eighth Workshop on Hot Topics in Operating Systems, pp. 0075, 2001.

[5] Clarke I., Sandberg O., Wiley B., and Hong T.W. Freenet: A distributed anonymous information storage and retrieval system. In *Workshop on Design Issues in Anonymity and Unobservability*, pp. 311-320, ICSI, Berkeley, USA, July 2000.

[6] Stoica, I., Morris, R., Karger D., Kaashoek M.F., Balakrishnan, H. Chord: A scalable peer-to-peer lookup service for internet applications, Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications, San Diego, California, United States, pp: 149 – 160, 2001.

[7] Kubiatowicz J., Bindel D., Chen Y., Czerwinski S., Eaton P., Geels D., Gummadi R., Thea S., Weatherspoon H., Weimer W., Wells C., and Zhao B., Oceanstore: An architecture for global-scale persistent storage. In *Proc. ASPLOS*, pp. 190–201, November 2000.

[8] Zeinalipour-Yazti D., Kalogeraki V., Gunopulos D., Information Retrieval Techniques for Peer-to-Peer Networks, Computing in Science and Engineering Volume 6 , Issue 4, pp. 20-26, July 2004.

[9] http://www.daml.org/services/owl-s/1.1/

[10] RDQL - A Query Language for RDF, http://www.w3.org/Submission/RDQL/

[11] SPARQL Query Language for RDF, http://www.w3.org/TR/rdf-sparql-query/

[12] Web Ontology Language overview http://www.w3.org/TR/owl-features/

[13] Manville, B. and Foote, N. “Harvest Your Workers' Knowledge”. Datamation, July 7, 1996.

[14] Lave, J. and Wenger, E., *Situated Learning: Legitimate Peripheral Participation*, Cambridge University Press, Cambridge, UK, 1991.

[15] Wenger, E. Communities of Practice: Learning, Meaning, and Identity. Cambridge University Press, Cambridge, UK, 1998.

[16] Domingue, J., Motta, E., Buckingham, S., Vargas-Vera, M., Kalfoglou, Y. and Farnes, N., Supporting Ontology Driven Document Enrichment within Communities of Practice. Proceedings of the 1st international conference on Knowledge capture K-CAP '01, Victoria, British Columbia, Canada, October 2001.

[17] McLoughlin, C. (1999) Culturally responsive technology use: developing an online community of learners. British Journal of Educational Technology, 30(3): 231-243.

[18] Preece, J. (2000) Online Communities: Designing Usability, Supporting Sociability: Wiley.