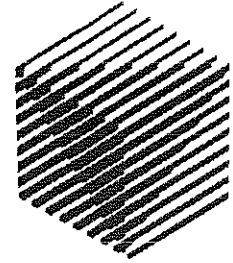European Research Consortium
for Informatics and Mathematics

ERCIM

Ninth DELOS Workshop

Digital Libraries for Distance Learning

Brno, Czech Republic, 15-17 April 1999

NINTH DELOS WORKSHOP

Digital Libraries for Distance Learning

Brno, Czech Republic

15-17 April 1999

CONTENTS

The DELOS Working Group is a part of the ESPRIT Long Term Research Programme (LTR No. 21057) and is managed by ERCIM.

The DELOS Partners are:

ERCIM members:
CLRC, CWI, CNR, FORTH, GMD, INRIA, INESC, SICS, ETH-SGFI, SINTEF Telecom and Informatics, MTA SZTAKI, VTT

Non-ERCIM members:
University of Michigan, USA
Elsevier Sciences, The Netherlands

For additional information, please contact
Costantino Thanos
Istituto di Elaborazione della Informazione, Consiglio Nazionale delle Ricerche
Via Santa Maria 46
I-56126 Pisa
Tel: +39 50 593492, Fax: +39 50 554342, E-mail: thanos@iei.pi.cnr.i
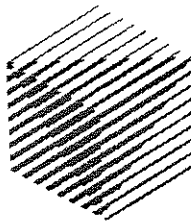
The European Research Consortium for Informatics and Mathematics (ERCIM) is an organisation dedicated to the advancement of European research and development, in the areas of information technology and applied mathematics. Through the definition of common scientific goals and strategies, its national member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry. To further these objectives, ERCIM organises joint technical Workshops and Advanced Courses, sponsors a Fellowship Programme for talented young researchers, undertakes joint strategic projects, and publishes workshop, research and strategic reports as well as a newsletter.

ERCIM presently consists of fourteen research organisations from as many countries:

Web learning environment evolution: integration and open problems

F. Rabitti*, G. Rossini*, P. Savino**, P.Venerosi**

*Istituto CNUCE, Consiglio Nazionale delle Ricerche,Via S. Maria 36
56126 Pisa, Italy
**Istituto di Elaborazione della Informazione, Consiglio Nazionale delle Ricerche
Via S. Maria 46, 56126 Pisa, Italy

Abstract

This paper focuses on the requirements in the evolution of learning environment, as driven by increasing demand of technological integration. Two main directions are identified: the integration between computer mediated communication (CMC) and multimedia technologies and the interaction between web information sources (digital libraries and multimedia information systems) and web training environments. A brief description of our experience, in the integration of web learning environments with digital library technology, introduces the discussion on open problems and new research directions.

1. INTRODUCTION

The many approaches we met in the realization of a web based training (WBT) environment enlightens the enormous shift provoked on distance learning performance by the advent of the Internet and the World Wide Web. Nevertheless, in spite of the flexibility reached to fit different user needs, pedagogical styles are still not adequate to take full advantage of the changes the technological development has produced.

Today, a new evolution of training environment performance is to be expected, due to the integration of different types of technology, and our attention is addressed to the directions that more will be able to improve learning process.

This paper, after a brief outlook of the state of the art of distance learning evolution, tackles its central issue on learning environment requirements, envisaging that their further evolution, from today situation, will consist of training environments open to further technological integration. Two directions are considered:

- the eventual integration between computer mediated communication (CMC) and multimedia technologies
- the interaction between web information sources (digital libraries and multimedia information systems) and web training environment.

A brief description of our experience, linked to the integration of web environment with digital library technology, introduces problems still open and new research directions.

2. THE STATE OF THE ART

In the last decade the widespread introduction of personal computers led to their use in training. Furthermore, the increase of their performance allowed the use of multimedia tools by which to develop interactive courseware, and enabled students to learn at their own pace and in their own time. Nevertheless, some disadvantages of the Computer Based Training (CBT) paradigm were:

- the isolation of students
- the lack of feedback with the instructor and the other students
- the difficulty to distribute up-to-date educational material.

Internet and the World Wide Web (WWW) today help to resolve many of the issue inherent to the previous training paradigm improving:

- easier delivery systems
- distribution of educational materials and on-line revision
- interaction between players.

The Internet provides, in fact, a flexible way to distribute educational material world-wide instantaneously, and the web may represent an obvious medium for delivering just in time (JIT) training to students; resources are

97

available by clicking on the links of the hypertextual navigational model and a web page can be realized just to represent an integrated and coherent working environment.

Independent cross-platforms allow one to use powerful tools such e-mail, newsgroup, IRC, etc. enabling, through a variety of formats, student-student and student-teacher interaction. When integrated within an appropriate management system these tools can provide collaboration between students and instructors, student progress tracking tools, and individually customized course materials.

Furthermore, the dynamic and interactive nature of WWW has reinforced the capability to interact with multimedia materials, enabling new ways of conceiving course to fit educational goals; the recent advent of second generation browsers, like Hot-Java makes WWW interactive by incorporating multi-media applications that can be programmed, run live and distributed in a portable manner [1].

Three issue have been involved in the main changes Internet and WWW have provoked on distance learning paradigm.

- The technological pull

Internet has promoting a powerful shift in many research areas including interaction. In educational communication field, Computer Mediated Communication (CMC) tools, allowing people to interact each other both synchronously and asynchronously in private, broadcast, and group discussion communication are Internet-based electronic communication technology. Browsers are typical WWW-based tools, shifting versus synchronous capabilities.

- The pedagogical issue

The pedagogical paradigm, stressed by the infusion of technology into the learning environments, requires a shift in teaching/learning styles. The problem to face now, is how to exploit the new technological potentiality for more advanced learning environments, where WBT tools are applied and combined only when they enhance the learning process [2]. This means for trainers to work in team to optimize resources for a given target and to be acquainted with new technology capabilities. In most situations WBT courses are still delivered as in traditional classroom setting, taking poor advantages of the new Internet capabilities.

- The role of players

The mission of teachers, supported by sophisticated authoring tools, do evolve from the role of lecturer to that of guide and facilitator. They remain responsible of the didactic choices, but will spend much more time in applying tool capabilities to create learning situations, to provide further communication opportunity, and to support student diversity.

Students, on the other side, free of the traditional control on their study time/place, can better organize how to spend their time choosing under their own responsibility studying and learning styles. They can share time among the study, the search of supplementary material, the collaboration with other students and the teacher, 'choosing for themselves a range of different communication and presentation paths enhancing their understanding".

The increased flexibility of studying/learning activity and the real- time feed-back are, indeed, the effective advantages the players meet approaching a web training environment

3. LEARNING ENVIRONMENT REQUIREMENTS AND THEIR EVOLUTION

In general terms, a web training environment can be conceived as a virtual space where user interactions are performed using a combination of tools offered by the environment, with the purpose of enhancing the learning process. Some conditions, indeed, must be satisfied. A pre-requisite is the availability of a suitable infrastructure enabling the use of Internet technologies, such as network connections, student access equipment, technical support personnel, and administrative capabilities. A general condition for a successful use of WBT is to endow the training development environment with publishing technologies and the essential communication and interaction systems for courseware delivering [3].

In this context, particular Computer Mediated Communication technologies, both asynchronous and synchronous, may be used by the target developers to enable specific functionality, within their economic, environmental, and pedagogical constraints. Sometimes hybrid systems are preferred combining courses delivered on CD-ROM with data streaming from the Internet. In most cases, the set of WBT functions are implemented by specific authoring tools, interaction and collaborative tools, and multimedia repositories.

Recent WBT tools, entirely web based, aim at a flexible development space, where user needs and tools are evaluated, implementing a strategy for an effective tool exploitation. Unfortunately, the range of Internet tools used in these systems is relatively small. The number and type of components and tools that can be used in a learning environment is much larger than what is available in today WBT systems, making impossible to follow a veritable strategy when designing a learning environment.

The further evolution we envisage from today situation is the realisation of open training environments able to integrate technologies coming from technological areas not directly focused on training and education. Two directions, in our opinion, will characterise this evolution and change the web training scenario: on one side,

the integration between Computer Mediated Communication technologies and multimedia technologies, and, on the other side, a stricter interaction between web information sources (digital libraries and multimedia information systems) and web training environment.

3.1 INTEGRATION BETWEEN CMC AND MULTIMEDIA TECHNOLOGIES

Computer Mediated Communication and multimedia technology integration, in fact, is far from being realised. Multimedia technologies have a growing role in the training process: multi-format data embedded into didactic material; repositories archiving multimedia educational objects; multimedia systems fostering information streams toward learning environment. CMC components, instead, (almost the synchronous ones), are not yet fully exploited in the most known WBT systems, and an effective integration process among Multimedia and CMC technologies, inside learning environments, didn't take place.

Some examples of integration like the video education on demand services and the interactive and dynamic software visualization, are the result of advanced researches, but they shouldn't hide other integration perspectives much closer to the current players expectations. We refer to the necessity to carry out the conversion of conventional courses into multimedia, to enrich existing CMC technology by means of multimedia tools, and furthermore to realize a distributed multimedia software environment supporting players in their production and use of multimedia courseware.

Such integration is also a challenge from a socio-economic point of view: predictions are of a growth of the market of multimedia in "training and education", pushing to the combination and integration of Computer Mediated Communication technology with multimedia technology. The synergy of both technologies will determine the future of the WBT industry.

This last consideration leads to point out the need of some guide to select tools and applications: at first, precise user requirements have to be stated, than, the creation of a sound and robust classification of tools features and technical requirements will help companies and developers.

The accessibility of information about new products is, indeed, a precondition for increasing competitiveness. We look, however, at the competition level where products aren't chosen for their brightening visibility on the net, but for their correspondence to the most granular user needs.

3.2 INTEGRATION BETWEEN DIGITAL LIBRARIES AND WBT ENVIRONMENTS

The requirement of a stricter interaction between learning environment and web information sources, relies fundamentally on two facts: first of all, on the consideration that all teaching and learning activities have always been improved by the access to a variety of information sources; and secondly on the fact that the search of appropriate sources of information throughout the net is today possible by means of gathering tools and digital libraries technology, which constitute the highest effort made to organise the information on the net for an easier search and an effective manipulation.

In fact, formal, informal, and professional learning [4] necessitates to interact with different domains of information, both general and specific; this opportunity is now being offered by the enormous quantity of information one can access browsing the web, and by the increasing availability of digital libraries, providing to satisfy many facets of the information requests "breaking down the physical barriers between resources".

The experience like that held at Virginia Tech [5], where courses have been integrated by an hypermedia digital library of Computer Science Literature, demonstrates the effectiveness of using digital libraries improving learning, and reducing the time/place distance between players and appropriate information. Furthermore, the growing range of input formats, lets digital libraries offer more granular information than users have never acquired before.

Adding digital library capability to a learning environment means for the players to be inserted in an international circuit of scientific information, with the opportunity to access material directly from their educational posts; to accelerate the dynamism of education; to allow for more integration of the different styles of learning and teaching. [6]

4. OUR EXPERIENCE

This last aspect was considered in our experience developed in the context of a project for the technology transfer towards Mediterranean Countries, funded by the Consiglio Nazionale delle Ricerche (CNR). Our work was essentially concerned with two kind of technologies: the digital library technology for the diffusion of specific scientific and technological themes of interest for the Mediterranean Countries, and the Web Based Training applications. Furthermore, we considered the case of web training environments where education and digital libraries technologies were considered together.

An implementation of a learning environment was, in fact, realized around a course using TopClass by WBT Systems, where the educational system server and a digital library server, previously installed, were integrated [7].

At first, it was considered the elementary need of students, instructors, and administrators to access reference material to deepen the understanding of questions raised during their work sessions. TopClass learning environment was then provided with a link to the digital library server previously installed, allowing teachers and students, in every instant (both at course composition phase and reading phase), to ask the local repository any kind of questions. The link provides the whole visibility of digital library functions, allowing users to formulate consistent queries.

Furthermore, since digital library logic allows for forwarding questions from local repository to the regional repositories and then at world-wide level, the chance to have a pertinent answer is expected to be reasonably high. In addition, another connection with the digital library, directed particularly to the students, was implemented enabling immediate and transparent access to the new material added to these repositories: a query relating to the topic of the course was formulated and entered by the instructor, then made persistent in the system, enabling students to get all news of interest just clicking a button. In this case, given the total transparency of searching execution, no knowledge of digital library is required.

We were interested also to extend the functionality of learning environment relating to the labs creation, by naming and annotating them for an easy retrieval and reuse, but our trial surprisingly failed because of the poor capability of the full-text search system, able to work on the lecture units only. This lack of prevision leads us to understand how search and reuse problems are not still sufficiently taken into account inside the training environment, and the need of an integration with multimedia content retrieval technology.

5. OPEN PROBLEMS AND FUTURE RESEARCH DIRECTIONS

In fact, to look at the world wide educational material, as a virtual distributed archive of courses interacting with digital libraries is today only an attractive hypothesis: the educational information domain has not yet benefited of the digital libraries research efforts and requires a qualified and adequate indexation: some researches in this sense are however in progress [8]. The problem is burdened also by the fact that even the basic terminology and concepts for representation, organization, and management of knowledge are still unsatisfactory. The reason is that a great part of the distance education information domain has been upset by the Internet advent, so the Information Representation (IREP) tools can't be repackaged but expect to be redefined starting from the current situation.

Traditional media disseminating knowledge have yet concepts and techniques to guarantee quality of information they carry, and methods for classification, retrieval and query to ensure relevance of search results. In some of these cases the problem may be how to repackage these tools on the net.

On the contrary, educational information throughout the Internet consists only on a part of traditional media such as textual information or static multimedia documents; from now on, prediction are of course material constituted "of video-clip presenting the lecturer and a sequence of slides that may involve pictures, video, animation, simulations with spoken or written explanations". This kind of dynamic TMDs (truly multimedia documents) has a 'temporal dimension', depending on the streams of information concurring to their composition. Multimedia authoring represent current dynamic TMDs area; investigations are in progress on the structure of such documents, but not adequate retrieval techniques have till now been exploited.

Reasoning on this emergence in order to focus some open problems and several aspects of research directions requires to distinguish three main questions: the access aids, the kind of repositories to collect TMDs and the reuse problems.

On a pragmatic and methodological level we envisage two types of initiatives concurring to facilitate the access to the educational hypermedia.

The first care concerns the terminological issue. Technical natural terminology represents, indeed, the first and more effective vehicle of communication inside a new information domain, when concepts aren't yet not fully operative. Today, the lack of specific classification schemes on web educational topics goes together with the inadequacy of the most known and authoritative classification schemes to incorporate systematically the most significant aspects of distance learning issue. Elsewhere, the appreciable efforts of the WBT systems evaluation centres to individuate the main topic aspects to be rated, didn't take useful results from a terminological point of view.

To proceed with terms standardization along with a categorization process, capturing the multifaceted aspects of the educational domain (covering features, course development, and course administration), means to endow users with a pragmatic IREP tool, to be used when other retrieval techniques couldn't be employed. In particular we refer to the use of index terms to select an information domain of interest (search starting query) against which to exploit other IR techniques or a combination of them. In addition, we have in mind the case of searching files containing software or labs, that necessarily relies on the use of technical terminology, whether controlled vocabularies, or taxonomies.

In our concern, all the initiatives supporting this aim, like that promoted by the Galaxy Scientific Corporation, Georgia, USA, in order to provide a classification scheme for the web training environment, must be strongly sustained. Great efforts are also expected from corporate institutions like, ACM, ASIS, INSPEC etc..

The lack of visibility of the sites where educational material are delivered is another problem to face and may be considered preliminary to that of the veritable information access. Such a question, depending upon the inadequacy of current Web structures to support easy paths to reach sites of interest, inhibits users to get the more common information they need (i.e. where, and which courseware are delivered on-line) and represents a different case from the more specific user need to access information even by content. To distinguish the specific requirements of the two situations, may lead to suggest different methodological paths to satisfy them.

In fact, the former aspect, we are now concerned with, relies on the necessity of 'grouping' information about material on the net: not single document meta description will be in question thereby, but an automatic generation of a network of meta-data expressly developed for the infrastructure they serve. This low-cost support will conduct users, interested to the on-line courseware, directly to their web page; not more it is expected to obtain. The more costly effort to create veritable educational digital libraries are not in question here.

What we need for a similar support is a set of meta information to improve visibility of collections, notifying the sites where distance learning is effective and courseware are delivered on-line, along with other information user requires to locate preferred sites. To do it, we envisage a combination of meta-information added on the top of the author-generated material to facilitate grouping (recent XML managing meta information may help to do this), linked to an ad hoc structure deserving, such as an organized index, to drive user to satisfy the meta-information they need, world-wide.

This kind of application will be fully based on the hypertextual model of web using HTML, XML, augmented by an auxiliary database structure of links enabling forwarding requests.

The logic of the hypothesis is inferred from librarian experience, where it is usual to distinguish data necessary to locate documents supporting associative search, like the set of data enabling to construct directories, from all others including auxiliary data that deserve to create true catalogues, searchable even by content. Furthermore, the utility to create structures of link separate from documents, instead of over-structuring hypermedia in order to diminish the abundance of 'go to' inherent to hypertextual model, is heavily and authoritatively sustained [9], [1].

To search for information on the net throughout well tagged collections of courseware capable of cross-search by means of suitable indexes requires, indeed, an other kind of effort. The path seems to be similar to that being performed within the Digital Library Project, sharing the paradigm of federated database architectures. Queries will be submitted against the federated schema and transformed into transparent requests interoperating among the databases involved.

Therefore, digital libraries experience deals with material such text-image and graphics; in the web educational domain, indeed, documents are composed by media (such as text, graphic, and image) and temporal media (such as audio, video, and music); thereby, the synchronization among media in the same document has to be resolved along with its operability among dynamic and/or multimedia data contained in the federated collections.

Considering the question of collection containers, in the functioning of learning environments, we note that WBT systems use rough object stores or a set of data bases with specific task; but their organization isn't oriented to further integration.

The choice of well suited containers becomes pre-eminent, instead, when we look to the perspective of interoperating TMDs collections. A number of technical problems arise when concerning the nature of multimedia document: how they be represented in their atomic components and at the whole to handle content-based multimedia data retrieval; or how to define a global access to multimedia databases geographically distributed. The area of databases results newly stressed, and hypothesis for the short-medium period concerning global schemas of distributed objects stores go together with others focusing "on the integration of multimedia databases for use in a distributed web-environment" employing metadatabase and search agent.

The reuse problem poses still an other question: to look at repositories both as containers delivering information on-line and as places supporting other user activities such as information management and sharing. Reuse, in fact is based both on the need of search and access capability and on mechanisms to manipulate different objects in order to create new TMDs. This question, apart the case where materials are reused at the whole (the need is that of an easy upload capability), involves the integration of different technologies. Video education on demand services are an example. The perspective nowaday, is that of extending the dimension of documents from pure container of information into digital information object, 'active compound documents', combining information content together with executable code to introduce active behaviour [10]. The need of integrated work support environments may extend the scope of educational digital libraries and increases the requirement of some architectural framework handling all the aspects treated and more, giving consistency to any web learning environment design.

REFERENCES

[1] A.D. Marshall and S. Hurley. The Design, Development and Evaluation of Hypermedia Courseware for the World Wide Web. *Multimedia Tools and Applications*, 3, 5-31 (1996).

[2] P.B. Lawhead et al. The Web and distance learning: what is appropriate and what is not. Report of the ITiCSE'97 Working Group on the Web and Distance Learning. *SIGCSE/SIGCUE ITiCSE'97*, 27-37.

[3] S. Hartley et al. Enhancing teaching using the Internet. Report of the Working Group on the World Wide Web as an interactive teaching resource. *SIGCSE/SIGCUE ITiCSE '96*, 218-228.

[4] G. Marchionini and H. Maurer. The Roles of Digital Libraries in Teaching and Learning. *Communication of the ACM*, Vol. 38, No 1, 67-75 (1995).

[5] URL: http://ei.cs.vt.edu/

[6] M.C. Buzzi, P. Venerosi, F. Rabitti and P. Savino. The Role of Digital Libraries in Network-Based Training. *Euro-Med 98 Conference*, Cyprus, March 1998.

[7] G. Rossini. L'utilizzo del sistema TopClass per l'apprendimento basato sul Web. *I.E.I., CNR, Pisa, Nota Interna B4-22*, 1998.

[8] H. Maurer. Web-Based Knowledge Management. *Computer*, March, 122-123 (1998).

[9] ARIADNE, Educational Metadata Recommendation Summary, Working document, Version 2.0, April 1998. URL: http://ariadne.unil.ch/

[10] C. Nikolaou, and M. Marazakis. System Infrastucture for Digital Libraries: A Survey and Outlook. *SOFSEM'98, Lecture Notes in Computer Science 1521*, Springer, 186-203 (1998).