Towards normalization in e-learning for collaborative work environments

Theme 1. Information sciences in the digital age: learning, education and research.

Everardo REYES [1-2], Abderrazak MKADMI [1-3]

[1] *Laboratoire Paragraphe, Université Paris VIII.*
2, rue de la Liberté 93526 - SAINT-DENIS cedex 02, France.
Tel. +33.1.49.40.67.58
[2] *Departamento de Comunicación, Tecnológico de Monterrey (ITESM) Campus Toluca.*
Eduardo Monroy 2000. Col. San Antonio Buenavista. CP. 50110. Toluca, Mexico.
Tel. +52.72.22.79.99.90
[3] *Unité de recherche : Bibliothèque numérique : pour la valorisation du patrimoine, ISD, Campus Universitaire de la Manouba, Tunisie*
Tél.: +216.71.600.

everardo.reyes@itesm.mx ; amkadmi@univ-paris8.fr

Keywords: digital documents, collaborative work, e-learning recommendations, learning objects

Abstract:

Contemporary approaches on information and communication technologies and their implication on knowledge domains reveal needs to improve our manners for treating information. The possibilities of XML technologies and hypermedia allow the development of new systems to distribute knowledge in local or global fashion. This perspective enhances the role of authors who have now access to open channels for sharing and working with other authors.

This paper presents an exploration of digital documents in the educational field and its implications within common problematic such as normalization, modeling, sharing and reuse, with the intention to present an integration of two software applications developed at the *Paragraphe Laboratory* of the University of Paris VIII: ICRS and HTT Online.

Taking into account the works lead by W3C, IEEE, and IMS Global, our study aims to link the robustness of XML technologies with the support of collaborative work strategies in distance learning environments where Learning Objects and Content Packages seem to become the common currency between universities and educational institutions.

1. Introduction: an approach to digital documents

Creating and exchanging information has established common practices in most of today's knowledge domains. Either we lead a project on digital libraries or on distance education we deal with the so-called *digital documents*.

New ways of producing and making available information demand to reconsider the notion of document itself. Contrary to traditional documents - or *analog documents* - digital documents are not static objects regarded in terms of its physical form or its editing processes [Chartron, 2002]. In terms of our research, we consider *digital document* an object created by an author that can also be reconstructed by a potential user with the intention to take advantage of it. Current discussions recommend the term *resource* to refer to digital documents and *document* to refer to traditional documents, which should be seen as finished entities [Cruzel, 2004], however we regard digital documents right from their content and we consider that the use of XML technologies for its modeling should be employed.

The field of digital documents is broad. [Le Cosnier, 1997] has proposed a typology consisting of seven categories, sorted by priority: 1) electronic journals; 2) pedagogical documents; 3) public administration documents; 4) heritage documents of great historical value; 5) personal documents; 6) electronic press; and, 7) media and Internet.

Our research mainly focuses on the second category: pedagogical documents for computer-based training (CBT). For [Le Cosnier, 1997], these documents develop its potential through distance and asynchronous environments; they are commonly modular and could be articulated in different ways depending on the learner's path. From this perspective, the employment of learning platforms - or Learning Management Systems (LMS) - is required, yet the development of tools for authoring content is needed for granting best practices and results to acquire, to share and to produce knowledge.

Taking into account the works lead by W3C, IEEE, IMS Global, and SCORM, our study aims to link the robustness of XML technologies with the support of collaborative work strategies in distance learning environments where Learning Objects and Content Packages seem to become the common currency between universities and educational institutions.

This paper presents an exploration of digital documents in the educational field and its implications within common problematic such as normalization, modeling, sharing and reuse, with the intention to present an integration of two software applications developed at the *Paragraphe Laboratory* of the University of Paris VIII: ICRS and HTT Online.

2. Learning and teaching in the digital age

E-learning, as we know it today, has been closely attached to the evolution of information and communication technologies (ICT). The inclusion of educational technology at school allowed teachers to enhance learning from different supports besides the spoken word. In that manner, radio, television, movies, photographs, and, more recently, CD-ROMs, DVD-ROM's and Internet have gained a great success during last two decades.

However, current recommendations demonstrate that digitizing learning materials and delivering them to students does not suffice. It is important to conceive those materials from a more public and collaborative perspective, which means creating pedagogical documents that can be used and reused in different contexts. One of the solutions proposed by international organizations are *Learning Objects* (LO).

2.1 Learning Objects

Defining the concept of Learning Object has not been simple; actually, there exist as many definitions as literature on the theme. Furthermore, it is necessary to explicit that Learning Objects are also called differently depending on the context or institution, most notably: "pedagogical objects" [Kenley, 2000], "online learning materials" [MERLOT, 2000], "pedagogical documents" [ARIADNE, 2000], "sharable content objects" [SCORM, 2004].

It has been noted by [Wiley, 2001] that LOs are grounded in the object-oriented paradigm of computer science. This paradigm values the creation of components (called "objects") that can be reused in multiple contexts. Adapting those ideas, the main characteristics of Learning Objects are:

- they are small, structured and standalone pieces that compose a course ;
- they can be used and reused in different contexts;
- they contain high quality information in a manner that facilitates learning and pedagogy;
- they have a stated objective and a designated audience;
- they have ownership and associated intellectual property rights;
- they are transportable (digital entities deliverable over the Internet);
- they are addressable and available for others through metadata.

We agree with the definition of Learning Object proposed by the E-learning Consortium [ELC, 2003]: "A re-usable, media-independent chunk of information used as a modular building block for e-Learning content. Learning objects are most effective when organized by a metadata classification and stored in a content repository such as a Learning Content Management System". And, in the same line, with the five abilities that should be assured by standard bodies and specifications in e-learning [ELC, 2003]: 1) Interoperability; 2) Re-usability; 3) Manageability; 4) Accessibility; and, 5) Durability.

3. Towards normalization in e-learning

E-learning is still in its developing phase. As we have seen, even the term has no official definition and there are plenty points of view. Besides, concepts such as Learning Object, Learning Management System, or Content Package, still cause confusion from one context to another.

The works lead by international organizations that regularize the use and reuse of learning technologies have pointed out specifications that aim to assist best practices for their implementation. Among those organizations we take into account in our research the W3C, the IEEE and the IMS.

3.1. IEEE LOM: defining metadata

Nowadays, the only learning technology standard is the IEEE Learning Object Metadata (IEEE Standard 1484.12), which was approved in 2002 and is based on the IMS LOM (2001).

IEEE LOM defines a base schema distributed in nine general categories to specify a data framework to define the metadata structure of the instructional components. Each category contains different sub-elements (figure 1).

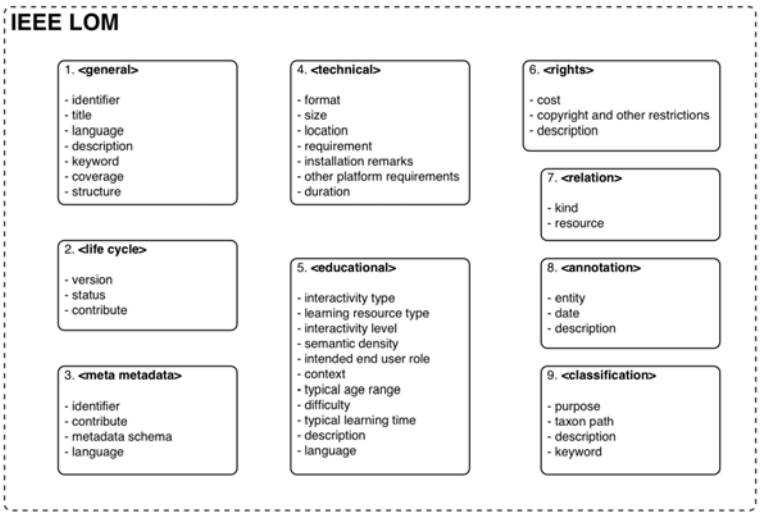

Fig. 1. IEEE LOM elements and sub-elements [IEEE, 2000]

The objective is to guide learners, instructors and software developers to search, evaluate, obtain and use Learning Objects and facilitate their exchange. Nevertheless, in the real practice it has been perceived that IEEE LOM lacks of a well-defined learning perspective.

3.2. IMS CP: distributing learning objects

IMS Global was initiated in 1997 by the non-profit consortium EDUCAUSE, its goal is to detail a set of specifications that build a framework to interchange educational elements. IMS covers six specifications for learning environments: 1) metadata; 2) content packaging 3) assessments design; 4) learner profiles; 5) content sequence, and 6) learning design.

Our research focuses on the Content Packaging Specification [IMS CP, 2003]. The concept of content package implies the distribution of a single and standalone compressed binary file (zip, jar, cab) containing a course, a lesson, a module, or simply a collection of related content objects. A content package bundles content objects with a content organization that is described in a manifest.

The manifest, on its side, is an independent XML file that must be named *imsmanifest.xml*. It describes four parts of the learning object: metadata, organizations, resources and sub-manifests. As recommended by IMS CP, the manifest should be stored at the directory 's top level, from where other sub-folders are created and organized along with the physical files of the learning unit (e.g. PDF, HTML, JPG, MPEG, etc.) (Figure 2)

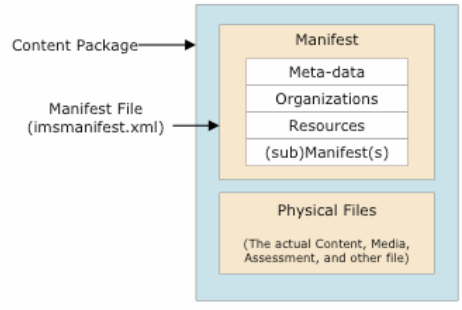

Fig. 2. Conceptual Content Package

Currently, the distribution and integration of packages in different systems is not yet well supported, even if they offer fully compliance to the specification.

3.2. XML: modeling documents

The World Wide Web Consortium (W3C) developed the eXtensible Markup Language in late 90s as a protocol for containing and managing information, however, it is also regarded as a family of technologies that can do everything form formatting digital documents to filtering data.

XML is considered as the future language for electronic exchanges because of its capacity to associate a semantic structure to data (by means of elements and attributes) and the rigid separation of format from content. Furthermore, XML offers the possibility to present data in different formats, to transform its document structure into another (using the eXtensible Stylesheet Language) and to add multimedia assets (e.g. SMIL).

Since the beginning, it has been necessary to accurate practices in order to benefit from semantic possibilities of digital documents. Since XML allows us to describe our own elements, the panorama conceived by Tim Berners-Lee requires an integration of document models (e.g. XML Schemas) with ontologies. In this sense, it becomes important not only to create those models but also to share them.

Current approaches for sharing document models take us to the notion of *repository*. A repository of XML models is a database that facilitates the access to existing document models in a determined area, responding to a certain need. The main idea behind repertories is to foster the free exchange between professional users and to get informed on the work that is being done by others [Mkadmi, 2003].

4. HTT Online and ICRS, an approach to collaborative work environments

ICRS and HTT Online are two software applications created from an educational context at the *Laboratoire Paragraphe* of the University of Paris VIII. This section sketches an approach to collaborative work environments based on reflections on normalization of technology for modeling and assisting the creation of learning materials. By integrating both tools, we propose a hypermedia perspective to simplify the presentation and interaction of actors with digital documents.

4.1. Introducing ICRS: collaborative research of information

The Information Collaborative Retrieval System (ICRS) was developed as a result of field studies and a state-of-the-art in the subject [Mkadmi, 2004]. This prototype is a proposition to users' needs in digital resources. Among the services proposed are:

- Access to a XML models database, with the intention to provide a series of models on determined fields with particular needs. This repository fosters the free exchange of models between professional users, who stay informed on practices and methods lead by other users in specific domains. It also represents a collaborative workspace for assuring the exchange of information.

- Collaborative research and access to a XML structured database that contains documents described according to the structure type defined in the XML repository. The graphical interface supports collaborative work between users and researchers.

The objective of ICRS is to promote collaborative work between users to integrate a virtual community of researchers. By means of repositories, users are able to know what others are doing and to participate in two modes: *reader* (by searching and accessing existing models) and *author* (by creating own models and making them available to others). Documents are on XML format, which facilitates parsing data from other systems.

There are two ways of collaboration in ICRS:

- *Collaborator*: consists of making a collaborative research in its strict sense. This mode considers similar users' profiles to show to them queries and results reciprocally. It is possible to annotate documents and evaluate its pertinence. These annotations may be available for other users, as well.

- *Observer*: this mode deactivates the onscreen presence of other queries, however user's researches, results, annotations and comments remain visible to other users. Observers are meant to be domain experts, persons who would extend their competencies to students or practitioners.

From the main interface of the system, users can choose the mode of collaboration, their profile (educational level and disciplinary field), and, search parameters (keywords, author and text) (Figure 3). Synchronic collaboration is also supported from this interface; users select contacts from a list to send instant messages and to display others' researches sorted by search parameters and the number of results gathered. This functionality allows users

interested in the same subject to click right into other's query to get the same results, avoiding to redo operations.

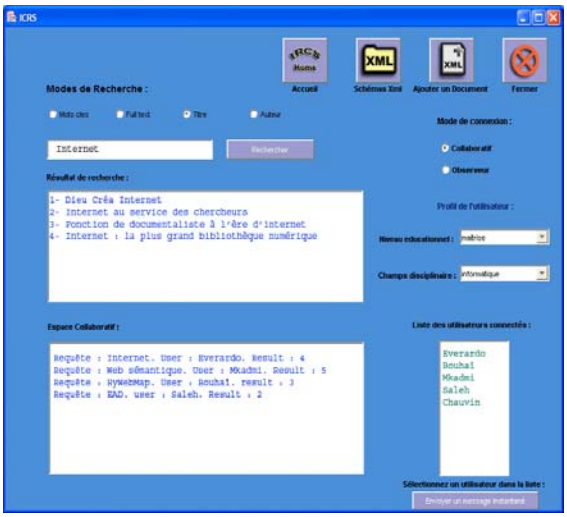

Fig. 3. ICRS: Collaborative research interface.

Results of queries are listed as hyperlinks and ordered by pertinence. They show the title of document and another collaborative interface is launched if clicked.

Annotations and evaluations are two elements that can be easily added to digital documents, this is of great importance when dealing with collaborative environments. Annotations are normally regarded as a technique for appending content (so to share ideas) to a document; on the other hand, evaluations are important indications on the way the author qualifies the document. ICRS can display annotations found in documents and allows to add more, as well as evaluate the content (figure 4). The function of annotations is to emphasize certain parts of the document by underlining them with different colors, depending on the theme. ICRS propose two kinds:

- *Public annotations*: a user highlights some parts that will be visible for all users. These annotations may be in form of comments, of quality judgments or any other that reflects the content of the document.

- *Private annotations*: only the author-reader can see them. They are in form of remark points and/or indications that researchers would work later.

For evaluations, the interface gives a note scale to evaluate the current document, regarding its pertinence and consistency with the query passed. Notes go from 1 (not pertinent) to 5 (very pertinent).

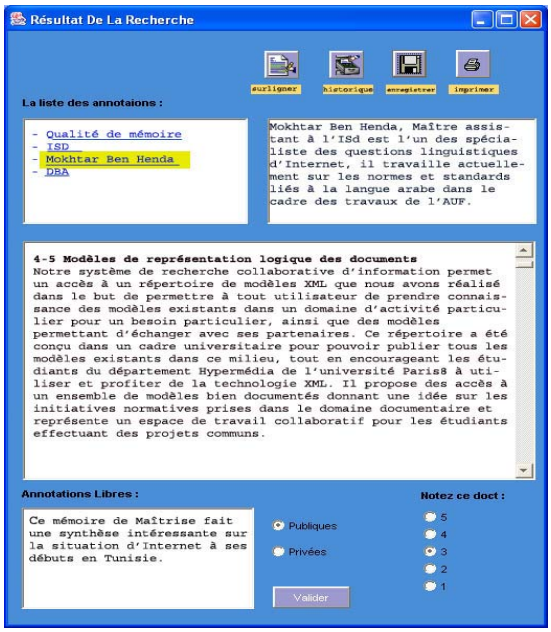

Fig. 4. ICRS : Aggregated value to documents

4.2. Introducing HTT Online: creation of learning materials

HTT Online stands for *HyperTectol Online*, which is the offline system presented in 2004 as an assistant tool for creating distance electronic courses [Reyes, 2004]. However, one of the major constraints was its dependency to Macintosh platform and operating system Mac OS X 10.2. HTT Online represents a sort of evolution of HyperTectol and is meant to assist the creation of Learning Objects in form of pedagogical digital documents and to distribute them as content packages. HTT consists of an interface that allows authoring content direct from a web browser (Gecko browsers suggested), which makes it platform-independent. In addition, one of the main improvements from the offline version is that HTT Online outputs documents defined in XML format, which allows a semantic treatment of information and facilitates its transformation.

HTT has been conceived for distance education environments where faculty members have experience with learning platforms and need assistant tools for creating their own learning material. From this point of view, our system considers two main kinds of users: teachers and learners, in other words, *authors* and *readers*.

The creation of digital documents is based on a window management system, where each window is presented inside the browser (simple HTML layers controlled with JavaScript). In this sense, authors may load several windows at the same time for reading previously released content, authoring new documents, or linking between documents. The text-based content of windows is added and edited through a word processor-like tool, which defines the element markups in semi-automatic way (e.g. the author selects headers to select if they head a document, a section or a paragraph) (figure 5)

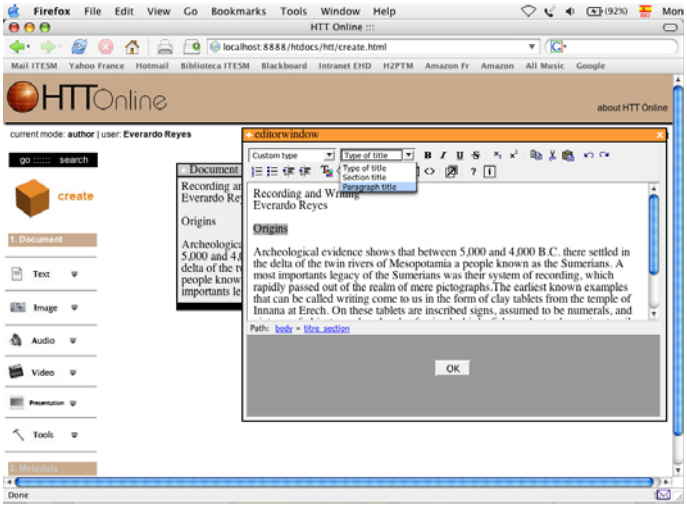

Figure 5. HTT Online, window manager system and interface for editing documents

During the creation it is possible to include media assets: images, audio, video and animations. Asset files are stocked inside the final content package unless they are described as URL addresses. Moreover, any window is capable of loading URLs, making them to appear as virtual windows.

Every document created with HTT is output in XML format and follows the structure proposed by A. Mkadmi [Mkadmi, 2004] for defining scientific documents and technical reports. The model regards four parts in a document: metadata, content of the document, footnotes, and bibliography references. Anyhow, our current version focuses on the content definition (figure 6).

HTT Online offers a metadata editor in compliance with the IEEE LOM standard. The editor is represented through HTML forms in two flavors: shortcut (most used elements) or extended (almost eighty elements to be filled in are presented). This interface (figure 6) generates two files: the *xmlmanifest.xml* and the *start.html*, both files located at the folder's top level.

Readers (students) access the system in read-only mode. They are able to search for learning objects and verify metadata upon the following parameters: name of the author (instructor), title of document, subject, date, id entry, and type of media asset included.

Finally, the *start.html* file allows readers and authors to have a quick view of the content of the learning object, based on the metadata manifest. Furthermore, access to learning material can be directly done from here. For all documents found, users are able to choose from a series of stylesheets in order to display the document according to his/her characteristics.

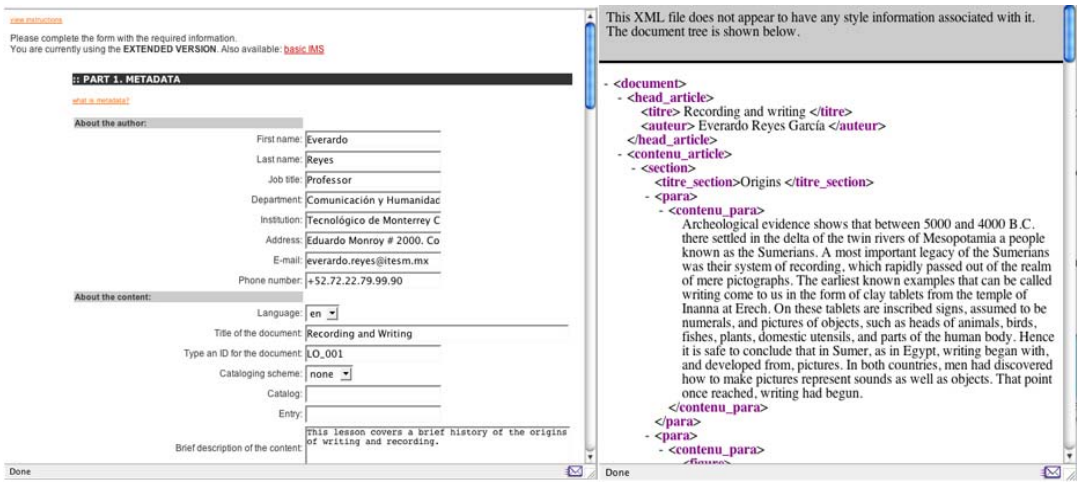

Figure 6. Metadata editor and output of XML document

4.3. Integration of ICRS with HTT Online

The employ of the systems presented above implies a first step towards semantic pedagogical learning materials, where XML technologies combined with standards and norms are the cornerstone for their development. In this part we analyze the integration of ICRS and HTT Online from a collaborative work in distance education environments perspective.

So far, one of the major projects on creating educational elements and attributes in XML format for education is the Educational Modeling Language (EML), developed at the Open University of Netherlands in 1998. The EML is a notational method for e-learning environments based on a pedagogical meta-model and since 2000 it has been used to describe courses [Koper, 2003]. On February 2003, when the final specification of IMS Learning Design (LD) was approved, EML has not been maintained any longer.

Consequently, IMS LD is based on ELM. Its objective is to allow authors to create various types of educational designs, following a consistent notation, that can be implemented homogeneously in several courses or learning contexts. The implementation is made through three levels:

- *Level A*: contains the vocabulary to support pedagogical diversity,
- *Level B*: adds attributes and conditions to level A (this allows personalization and more elaborated interaction sequence based on students' portfolios), and
- *Level C*: adds notifications to Level B.

In spite of this, disadvantages have been related to its recent implementation, i.e. it does not consider student characteristics or student evaluation, which requires design with other IMS specifications (e.g. IMS Learner Information Package Specification, IMS Question and Test Interoperability), and does not give guidance to define educational contents.

As it was noted before, the system ICRS offers to instructors and domain experts an interface to work together to construct XML models for documents. In addition, every author can closely observe the work of others with the intention to make improvements on his own research. The importance of such system in e-learning environments is to promote a channel for collaborative work to exchange document models that could be used in specific situations. For digital documents the model is the most important part; syntax can be transformed anytime with XSLT to suit specifications, but models need human interaction with serious consideration for future transportability.

Once a document is found or annotated in ICRS, its treatment in HTT Online enhances the content with media assets, learning standard metadata and delivered mode in form of content package. In the same way, importing digital documents created in HTT into ICRS is achieved transparently, according to the Mkadmi model.

In order to benefit from semantic documents in e-learning, it is fundamental to promote the use of document models in faculty members as an approach to normalization. Collaborative work as a technique for granting interchange may promote greater advancements not only in local situations (e.g. laboratories, universities), but also in global circumstances (e.g. countries, consortiums).

5. Conclusions and perspectives

Contemporary approaches on information and communication technologies (ICT) and their implication on knowledge domains reveal needs to improve our manners for treating information. The possibilities of XML technologies and hypermedia allow the development of new systems to distribute knowledge in local or global fashion. This perspective enhances the role of authors who have now access to open channels for sharing and working with other authors.

From this context we have presented in this paper the integration of two software applications, developed at the *Laboratoire Paragraphe* of the University of Paris VIII, to promote collaborative work based on reflections on normalization of technology for modeling and assisting the creation of learning materials: ICRS and HTT Online. By integrating both tools, we propose a hypermedia perspective to simplify the presentation and interaction of actors with digital documents.

In one hand, ICRS proposes an interface for collaborative research of information. It consists of two main services: access to a XML models database (repertory for free exchange of models between professional users) and access to a XML structured database that contains documents described according to the structure type defined in the XML repertory. ICRS also supports annotations and evaluation to the documents found, which aggregate value to documents.

In the other hand, HTT Online is an online tool for creating learning objects in form of pedagogical digital documents and to distribute them as content packages. HTT consists of an interface that allows authoring content direct from a web browser. In addition, HTT Online outputs documents defined in XML format, which allows a semantic treatment of

information and facilitates its transformation, it also takes into account IEEE LOM for metadata definition and IMS CP for content packaging.

The integration of both systems has the intention to be a first step towards normalization in e-learning in collaborative work environments, where pedagogical digital documents are meant to have their own educational structure described in XML format. The main advantage of such paradigm is to render documents understandable both by humans and computers.

Currently, we cannot discuss field experiences because the implementation is in its prototype stage and will fully start the next academic year at the *Département Hypermédias* of the University of Paris VIII; however, further work in our research is concerned with a transparent compatibility between systems (e.g. importing document models from ICRS into HTT Online) and with document models in the e-learning domain.

6. References

[ARIADNE, 2000] ARIADNE (2000). *Alliance of remote instructional authoring and distribution networks for Europe* [On-line] Available: http://ariadne.unil.ch/

[Chartron, 2002] Chartron, G., Guyot, B., Lafouge, T., et al. (2002). Le document numérique : un objet fédérateur de recherche en sciences de l'information. *Documentaliste Sciences de l'information*, 39 (6), December 2002.

[ELC, 2003] E-learning Consortium (2003). *Making sense of learning specifications & standards: A decision maker's guide to their adoption* [Electronic Version] Saratoga Springs, NY: The Masie Center.

[IEEE, 2000] IEEE LOM (2000). *LOM working draft v4.1* [Electronic Version]. Available: http://ltsc.ieee.org/doc/wg12/LOMv4.1.htm

[IMS CP, 2003] IMS CP (2003). *IMS Content Packaging specification v1.1.3*. [Electronic Version] Available: http://www.imsglobal.org/content/packaging

[Kenley, 2000] Kenley, R. (2000). *Pedagogical objects. Case studies towards a different perspective on collaborative education*. [Electronic Version] Paris: École d'Architecture. Available: http://www.elia-artschools.org/ejhae/Issue0/Kenley.pdf

[Koper, 2001] Koper, R. (2001). *Modelling units of study from a pedagogical perspective. The pedagogical meta-model behind EML*. Available: http://eml.ou.nl/introduction/docs/pedmetamodel.pdf

[Cruzel, 2004] Cruzel, S.L. (2004). Documents, ressources, données : les avatars des documents numériques. *Information-Interaction-Intelligence*, n° 1, 2004.

[Le Crosnier, 1997] Le Crosnier, H. (1997). *Les bibliothèques numériques*. October 1997.

[MERLOT, 2000] MERLOT (2000). *Multimedia educational resource for learning and on-line teaching* [Electronic Version] Available: http://www.merlot.org/

[Mkadmi, 2004] Mkadmi, A. (2004). *Recherche collaborative d'informations : repenser l'architecture des SRIs à l'ère numérique*. (Doctoral thesis). Paris: Université Paris VIII.

[Mkadmi, 2003] Mkadmi, A., Bouhaï, N., Langlois, M. (2003). Partager des modèles XML:Quel intérêt ? *BBF*, n°5, September 2003.

[Reyes, 2004] Reyes, E., Saleh, I. (2004). HyperTectol, an assistant and authoring tool for using multimedia in Learning Objects creation. *Proceedings of the 5th. International Conference on Information Technology Based Higher Education and Training*. Istanbul: IEEE.

[SCORM, 2004] SCORM (2004). *SCORM 2004 Overview*. Alexandria, VA: ADL. [Electronic Version] Available: http://www.adlnet.org

[Wiley, 2001] Wiley, D. (2001). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. Wiley (ed.), *The Instructional Use of Learning Objects*. [Electronic Version] Available: http://reusability.org/read/chapters/wiley.doc