A Process of Creating Learning Objects from a Congress Platform

Manuel Blázquez, Miguel Latorre, Gabriel Díaz, Manuel Castro, Jesús Arriaga, Fernando Pescador, César Sanz, Edmundo Tovar, and Tomás Pollán

***Abstract* — This paper presents the work carried out on recovering, manipulating, unificating, labeling and processing of documents developed within the congresses of TAEE (Technologies Applied to the Teaching of Electronics). The process developed consists of the creation of digital objects and documents arising from the extraction of metadata which define each digital object as a learning object. With the joint between a learning object and its metadata some information extrapolation is possible in multiple environments. Thus, these units of information permit the implementation of web environments, the study and analysis of their contents and the creation of file structures for their incorporation in repositories in a massive way.**

***Keywords* — Metadata, repository, LOM, Learning Object.**

I. Introduction

THE Project TAEE (from the Spanish name, Applied Technologies for Learning in Electronics) is an initiative that was founded with the aim to promote the application of learning technologies, as it was done in the UK with the project "*Computer in Teaching Initiative Support Service - CTISS*". In its foundation, TAEE attends to a new concept of project, the "*Service Project*" in order to facilitate and promote other related projects.

Its main purpose consists of providing a platform in an open environment, in which each teacher or group of teachers can design their own experiences and projects according to their needs and possibilities. As a goal, new technologies of mass education in electronics have been applied, with the direct use of educational tools and training resources created primarily from experience in the classroom and laboratory [1].

In order to share information and to present the work in this environment, every two years a congress is held in which faculty and electronics come together. It is the documentation generated in TAEE conferences since 1994 on which the works presented in this paper have been performed, involving the application of these criteria and features to apply to learning objects.

II. Data Treatment Process from TAEE

As a preliminary step, the digitization of every document has been carried out. Subsequently, after a process of separation of each individual document, 964 leading digital objects were obtained from the congresses and in addition, 4441 objects were recovered from these first. Above all, a process of tagging and coding has been undertaken. Given the heterogeneity of the type of objects [2] to handle, it has been necessary to develop a classification method that has been applied as an encoding process based on the congressional structure. The referred code become the name of every file in relation, and is composed by the year of publication, the session at which the exhibition was hosted and the order of presentation of the document contents.

Therefore, at this point the implementation of the phase consisting of individualized treatment of each learning object by means of metadata is required. Through the metadata, which essentially are data that describe the information developed in a given digital learning object, it should be possible to recognize the main information without having to access it, or rather, the metadata should provide insight surface of the documentation sought for recognition of their usefulness. In essence, the implementation of the markup language in the

The authors would like to acknowledge the Spanish Science and Innovation Ministry and the CYTED-508AC0341 "SOLITE – SOFTWARE LIBRE EN TELEFORMACIÓN" Project support. Also would like to acknowledge the support of the Complimentary Action TSI2007-31091-E "Objetos educativos reutilizables (para el EEES en las especialidades en las especialidades de las tecnologías de la información y las comunicaciones)".

Manuel Blázquez is with Electronics and Computer Engineering Department, Spanish University for Distance Education, Madrid, Spain (e–mail: manuel.blazquez.merino@gmail.com).

Miguel Latorre is with Electronics and Computer Engineering Department, Spanish University for Distance Education, Madrid, Spain (e–mail: pelaga@gmail.com).

Gabriel Díaz is with Electronics and Computer Engineering Department, Spanish University for Distance Education, Madrid, Spain (e–mail: gdiaz@ieec.uned.es).

Manuel Castro is with Electronics and Computer Engineering Department, Spanish University for Distance Education, Madrid, Spain (e–mail: mcastro@ieec.uned.es).

Jesús Arriaga is with TAEE Program Committee, Technical University of Madrid, University of Zaragoza, Madrid–Zaragoza, Spain (e–mail: jesus.arriaga@upm.es).

Fernando Pescador is with TAEE Program Committee, Technical University of Madrid, University of Zaragoza, Madrid–Zaragoza, Spain (e–mail: fernando.pescador@upm.es).

César Sanz is with TAEE Program Committee, Technical University of Madrid, University of Zaragoza, Madrid–Zaragoza, Spain (e–mail: cesar.sanz@upm.es).

Edmundo Tovar is with TAEE Program Committee, Technical University of Madrid, University of Zaragoza, Madrid–Zaragoza, Spain (e–mail: etovar@fi.upm.es).

Tomás Pollán is with TAEE Program Committee, Technical University of Madrid, University of Zaragoza, Madrid–Zaragoza, Spain (e–mail: tpollan@unizar.es).

documentation associated files would help to identify and select information optimally.

Moreover, application of metadata to a learning object facilitates their reusability [3]. Reusability, is taken to mean the ability to use the particular utility of object for its inclusion in a context that doesn't have to be the same scope as the object's original context. However, the formation and creation of learning objects involves some series of initial planning stages [4]. It should be understood that the stages have to be defined depending on the variable complexity of objects, their internal composition, its structure, as a set of properties, unified under the term of granularity. Moreover, together with the granularity, which figures as a quantitative concept, the qualitative features have to be taken into account. Some of these qualitative features can be related with the format of the digital object in an educational context as texts, graphics, photographs, data tables, and so on.

The idea is represented in Fig. 1, which shows a typical learning object, developed as an educational unit. To define its content, the unit has been equipped with an attached metadata file that defines it, but that in turn contains elements that can be reused out of context. To be able to carry out every component, they must be configured as separated units which have their own definitions through their own metadata files.

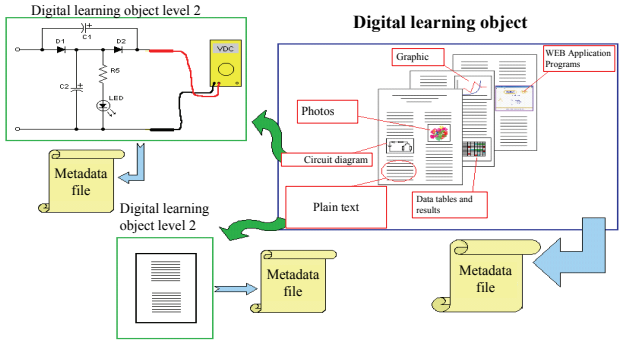

Fig. 1. Learning objects generic disposition, their definitions and relations.

At this point, the documentary extraction final step, each digital object has been classified and standardized by applying the unified coding system, which involves the formal name for all files (pdf, xml, html ....) related to the learning object.

III. The Structure of Metadata Regarding the Extracted Learning Objects from TAEE

To spread the dissemination of digital objects obtained from TAEE, the identification of three parallel models of metadata has been undertaken: a specific model for TAEE first, and two models directly applied from standard IEEE-LOM [5] metadata and Dublin Core [6], [7].

The TAEE specific metadata format, shown in Fig. 2, is based on the IEEE-LOM model, and provides a structure composed, in a first grouping metadata level, by three groups: "*General*", "*Technical*", and "*Ontology*". The first, "*General*", provides the specific TAEE information and therefore fits their subdivision in hierarchical form of ownership ("*Content of the document - Session – Congress*"). Therefore, the metadata associated with each of these subgroups define the nature and origin of the document.

The main metadata fields, in relation with a comparative with the two other standard models, LOM and Dublin Core, are located in the sublabel "Content", where Title, Author, etc, are described.

The second metadata block provides technical matters relating to the original format of the document, the name of the file, the type of educational purpose and licensing rights and the relation of precedence and offspring in a group of dependent objects. Moreover, the "*Annotation*" may contain a string in which the author or agency responsible offers additional comments.

The last block of metadata is based primarily on a specific ontology TAEE, which provides taxonomy of all branches of electronics treated in TAEE. Additionally, it is expected the insertion of additional codes to define them, using the coding WIPO system (from the World Industrial Property Organization), in which objects as patents and industrial designs are classified. Both metadata fields, "*Classification*" and "*WIPO*", allow the introduction of multiple values for better definition of the digital object, in order to identify technical and pedagogical nature of the paper.

The recovery process of metadata for a digital objects has used an ODBC connection database as a temporary storage means for metadata. The choice of such a kind of database has been mainly based on the flexibility to modify and update data, as well as the relational capacity. The latter allows subsequent analysis of the data, so statistical as by application of social networks meta-analysis tools to study relation among authors, organizations, themes and other parameters.

IV. Automatic Proccess of XML Metadata Files Generation from TAEE Digital Objects

Once the database is full of metadata by which learning objects are defined, it's time to implement an automatic proccess of conversion to XML [8] files. To start the metadata download job, a block programmed in Java, the "*metadata Shuttle*", has been developed whose specific purpose consists of the generation of XML files from the database and which uses a XML template to put them into. The way it works consists of the recognition and transformation of every metadata, used as strings. So, the Shuttle takes the string and searches for a decoy that has been placed between opening and closing metadata tags. The program essentially seeks between strings and decoys, and makes a complete overlap in the field name from tables in the origin database. Given the coincidence of the field, the decoy is replaced by inserting the corresponding recorded value.

The Shuttle program uses an iterative two-tier structure: first, searches the decoys in every field value of the record in the XML template and as a second step, fills and closes the file. In a new database record, the XML template is

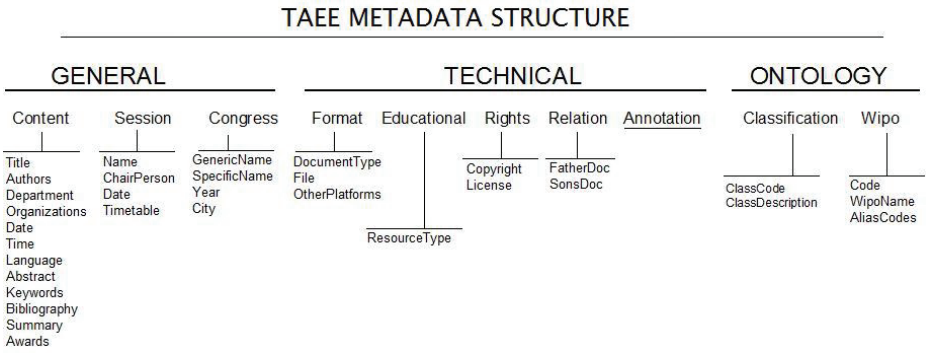

Fig. 2. Metadata Structure TAEE.

open and a new process of search and fill begins. The Shuttle generates as many XML files as there are records that have to read in the database. In a short time, it yields a set of files that will be located in the file structure.

The components of the application to generate XML files consist of a configuration file and a relational database itself as input elements and a metadata template as an output. The organization of files in directories is divided equally, that is, there is a directory where configuration files are located, a directory as Java source code and a directory provided for the XML file location generated in the process of creation.

This file structure is the same for every conference and shall contain, separately, the elements of composition of documents. While the papers are separated as learning objects, metadata files will have a distinguished place for each set so TAEE metadata, as LOM metadata or Dublin Core metadata. For the generation of the latter, two modified separate blocks have been created in the Shuttle to generate XML files based on templates that meet those LOM and Dublin Core requirements [9]. In this modified block, the programming structure is very similar to the original, although it has specifically been changed in order to the manipulation of multiple information metadata strings, converting these into substructures repetitive paths as required standards. A general view of the proccess is depicted in Fig. 3.

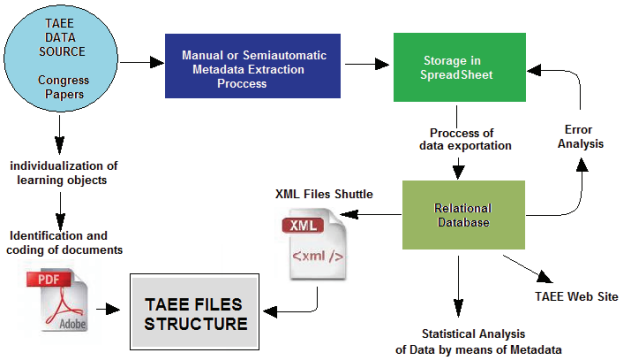

Fig. 3. Full Proccessed documentation of treatment of TAEE.

At this point it can be said that, regardless of adding some TAEE original document of some interest to the community concerned, the metadata have been extracted and created in different formats (spreadsheet, database, XML files). It is conceivable that such a multiplicity of data is unnecessary but is justified for two reasons: first, the multiplicity generates consistency and security in the information; and second, it's just a question of memory availability to maintain such parallel metadata structures. Because of its applications, its growth is limited and can include the documentation in any kind of repository in a quick way, giving flexibility to spreading processes. Certainly, the LOM-Dublin Core interconversion is ensured by the resemblance between the labeling of both standards and the amount of metadata that is used in these standards.

Once XML files and learning objects have been generated, the process of creating digital objects from TAEE is finished. After the complete file structure and metadata objects have been placed, the next phase begins, with the goal to spread the contents of TAEE online.

V. Dissemination of TAEE Learning Objects

The process of dissemination of digital objects from TAEE has been carried out simultaneously in two stages. The first is the use of metadata for the automatic creation of the pages that make up the TAEE site. This process has used the same visual format for the formation of each congressional page, each of which offers direct access to every learning object and its metadata.

Furthermore, as a second step the bulk of digital objects has been carried out in our university's institutional repository, eSpacio-UNED, [10] located in the web site http://e-spacio.uned.es as a means for the storage and distribution of objects. Thus, the components associated with the digital object are: a component of LOM metadata, a component of Dublin Core metadata derived from the LOM description, a component linked to the URL of the digital object itself and a component of the relationship of that object to other TAEE objects. This means that every digital object will have an own configuration collection, whose handling and visualization is possible by assigning a persistent URI to every one of the components of the object and the object itself. This allows the access to both learning object and its components, not only from the environment of the repository, but external to it, too. By "*external*", it's meant to be incorporated into the elements likely to be found by the searching engines available on the web as Google. To make both stages a matter of joint, as an alternative to TAEE web, it has been developed a user interface to access TAEE data located in the

Fig. 4. User interface homepage to access to TAEE information located in repository.

repository. In that interface, XML data has been used as site backbone. By means of "adhoc" style sheets tranformation and cascade style sheets, it has been possible to conclude the interface attending to the repository philosophy. An illustration of user inteface homepage is shown in Fig. 4.

VI. CONCLUSION

The content of this document offers the work done on the processing of information generated in TAEE congresses. The general objectives consist of the creation of a unified and structured information environment from the documentation generated in TAEE. XML files have been created in connection with every digital object extracted from the general documentation, and have been used in further utilities as a common Web environment, which favours the thematic consultation of the documentation, the adaptation of TAEE documentation to metadata standards, the promotion of online presence through repositories hosting, the performance of a content analysis of digital objects through which it's likely to identify the path TAEE has taken through years of existence, the relations between the organisms and agencies who have taken part in TAEE and the methods and procedures carried out in greater profusion in the field of education.

The achievement of the objectives aims TAEE to be enhanced as an Internet social network by adopting the measures applied to documentation and adaptation to the appropriate standard formats. Thus, it aims to provide the digital learning objects handled to the public and interested academic personnel, allowing the reuse of materials generated as a source of knowledge.

ACKNOWLEDGMENT

In memoriam of Tomas Pollan, a good friend and a best colleague.

REFERENCES

[1] J. Arriaga, A. Carpeño, T. Gordillo, "Del objeto de aprendizaje a la implementación de una asignatura. Un caso práctico," Universidad Politécnica de Madrid, *Congreso TAEE* 2006.
[2] Desarrollo de un esquema de metadatos para la descripción de recursos educativos: El perfil de aplicación MIMETA. Miguel Ángel Marzal García-Quismondo, Javier Calzada-Prado, Aurora Cuevas Cerveró. Revista española de Documentación Científica nº 29. Octubre-Diciembre 551-571, 2006. ISSN 0210-0614.
[3] M. Latorre, et al. "A good practice example on Learning Object Reutilization". *DIEEC, UNED* Spain, 2008.
[4] Marie-Claude Côté, (TBS), Cpt. Peter Hope (DND). "Presentation to the GOL Metadata Working Group". Government of Canada E-learning Metadata Application Profiles. January 20, 2004
[5] Learning Technology Standards Committee of the IEEE. "Draft Standard for Learning Object Metadata" IEEE 1484.12.1-2002. 15 July 2002.
[6] Sitio Web de DCMI. http://dublincore.org/index.shtml. Consultado el 27 de febrero de 2009.
[7] Bravo García, A.. Microformatos Dublin Core. Artícle from Web site http://www.alzado.org/. Consultado el 7 de marzo de 2009.
[8] Gutiérrez Gallardo, J.D. "XML – Manual imprescindible". Anaya Multimedia. ISBN 84-415-1576-X.
[9] J. I. Rouyet and V. Martín V. "A comparative study of the metadata in SCORM and Dublin Core". Universidad Pontificia de Salamanca.
[10] Proyecto e-Spacio UNED. Repositorio digital institucional de la UNED. Memoria descriptiva. Biblioteca de la UNED. URL: http://e-spacio.uned.es/fez/index.php. Última visita:14/4/2009.