A process of creating learning objects from a congress platform

Manuel Blázquez, Miguel Latorre, Gabriel Díaz, Manuel Castro
Dep. Ingeniería Eléctrica, Electrónica y de Control
Universidad Nacional de Educación a Distancia
Madrid, Spain
manuel.blazquez.merino@gmail.com, pelaga@gmail.com, gdiaz@ieec.uned.es, mcastro@ieec.uned.es
Jesús Arriaga, Fernando Pescador, César Sanz, Edmundo Tovar, Tomás Pollán
Comité de Programa TAEE
Universidad Politécnica de Madrid – Universidad de Zaragoza
Madrid - Zaragoza, España
jesus.arriaga@upm.es, fernando.pescador@upm.es, cesar.sanz@upm.es, etovar@fi.upm.es, tpollan@unizar.es

***Abstract* — This paper contents the work carried out about 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 in the creation of digital objects and documents arising from the extraction of metadata that define each digital object as a learning object. With the joint learning object-metadata is possible the extrapolation of information in multiple environments. Thus, these units of information let the implementation of web environments, the study and analysis of their contents and the creation of file structures for incorporation massive repositories.

***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 UK with the project"*Computer in Teaching Initiative Support Service - CTISS*". In its foundation, TAEE attends to a new concept of project, the "*Project Service*" in order to facilitate and promote other projects.

Their main purpose consists in 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 to 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 congresses since 1994 on which it has been performed the works presented in this paper, involving the application of these criteria and characteristics that apply to learning objects.

II. The data treatment proccess from TAEE

As a preliminary step, the digitization of all congressional documents 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, it has undertaken a process of tagging and coding. Given the heterogeneity of the type of objects [2] to handle, it has been necessary a classification method to be independent of the type of object that has been applied to an encoding method based on the congressional structure, i.e., including the name of the files a document relating to a code that determines 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 it's required the implementation of the phase consisting of individualized treatment of each learning object by means of metadata. Through the metadata, which essentially are data that describes the information that develops in a given digital learning object, it should be possible the recognition of the documentation 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 documentation associated files would help to identify and select information optimally.

Moreover, application of metadata to a learning object facilitates their reusability [3]. For reusability, it's meant the ability to use in an integrated way the particular utility of object for its inclusion in a context that doesn't have to be the same scope as the object original context. However, the formation and creation of learning objects involves a series of initial steps of planning [4]. It should be understood that these difficulties lie in the variable complexity of objects, their internal composition, its structure, etc. This set of properties is unified in terms of granularity of digital objects. Moreover, together with the granularity, which serves a quantitative concept, it has to take into account the qualitative features, i.e. allowing for the heterogeneity in the components of a digital object,

which will lead to the existence of formats of educational units as diverse as text, graphics, photographs, data tables, and so on.

The idea has being represented in Figure 1, which shows a typical learning object which is an educational unit equipped with a 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 in turn have its own definition through its own metadata file.

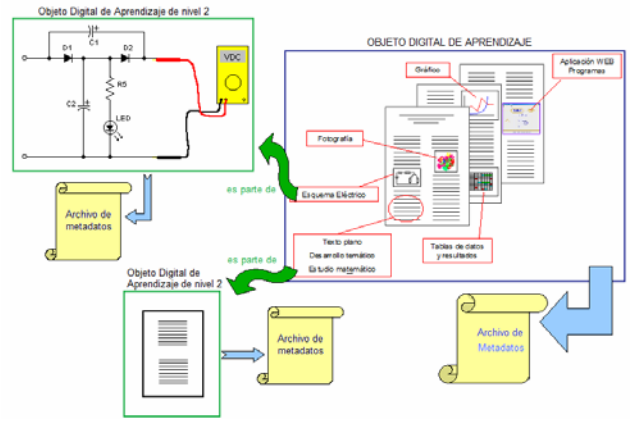

Fig. 1 – Learning objects generic disposition, their definitions and relations

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 Figure 2, is based on the IEEE-LOM model, and provides a structure that at first grouping level metadata, divided them into three groups: "*General*", "*Technical*", "*Ontology*". The first, "*General*", caters for those that are specific TAEE metadata and therefore fits their subdivision in hierarchical form of ownership ("*Content of the document - Session – Congress*"), and therefore, the metadata associated with each of these subgroups define the nature 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.

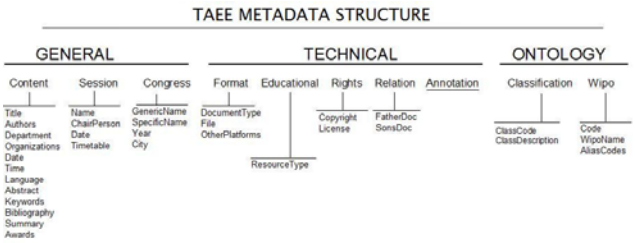

Fig.2 - Metadata Structure TAEE

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 offer any additional comment.

The last block of metadata is based primarily on a specific ontology TAEE, which provides a taxonomy of all branches of electronics treated in TAEE. Additionally, it is expected the insertion of codes to name 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 digital objects has used a database ODBC as temporary storage of metadata. The choice of such databases has been based mainly on the flexibility in modifying and updating, as well as available relational capacity. The latter allows subsequent analysis of the data, so statistically 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 define, it's time to implement an automatic proccess of conversion to XML [8] files. To start the metadata download job it has been developed a block programmed in Java, the "*metadata Shuttle*", whose specific purpose consists in the generation of XML files from the database and which uses a XML template to put them into. The way it works consists in the caption 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 registry value.

The shuttle program uses an iterative two-tier structure: first up the XML file by replacing the decoys for field values of a record, and as a second step, fills and closes

the file opening a new one on the template, which is cloned. The Shuttle program generates XML files as many records 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 relational database itself as input elements and 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 congress 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 separates blocks have been created in the Shuttle to generate XML files based on templates that meet those LOM and Dublin Core requirements [9]. In these 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 has been depicted in figure 3.

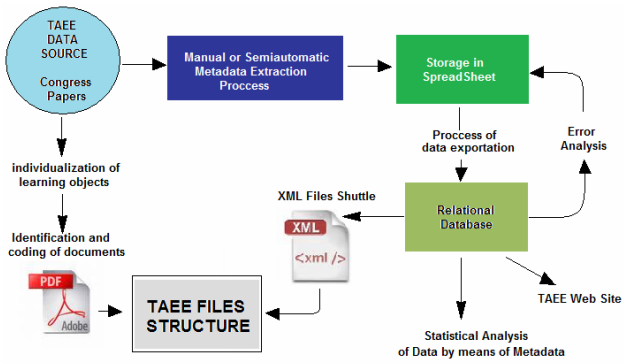

Fig.3 - Full Proccessed documentation of treatment of TAEE

At this point we can say that regardless of adding some TAEE original document of some interest to the community concerned, it has all the metadata extracted and originated in different formats (spreadsheet, database, XML files-TAEE ,-LOM XML files, XML files, Dublin Core). 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 availability of memory, its growth is limited and is minimized and offset by the ability to save conversion effort to try to include documentation in repositories of one sort or another. 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 generated the XML files and learning objects, 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. THE DISSEMINATION OF TAEE LEARNING OBJECTS

The process of dissemination of digital objects from TAEE was 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.

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 mean for the storage and distribution of objects. The integration of the resulting objects has generated in the environment of the repository a number of elements making up every digital object structure. 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 everyone 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 so foreign to it. By "*foreign*", 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 an user interface to access TAEE data located in the 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 has been shown in figure 4.

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

VI. CONCLUSION

The content of this document offers the work done on the processing of information generated in TAEE congresses. The general objectives consist in the creation of a unified and structured information environment from the documentation generated in TAEE. XML files has been created in connection with every digital object extracted from the general documentation, and it has 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 hosting on documentation in repositories, the performance of a content analysis of digital objects through which it's likely to identifie the path TAEE has taken, the relations between the organisms and agencies who has taken part of TAEE, as a social network and the methods and procedures carried out in greater profusion in the field of education.

The achievement of the objectives aims TAEE to enhace as 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

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)".

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

REFERENCES

[1] Arriaga, J., Carpeño, A., Gordillo, T. "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] Latorre, Miguel 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] Rouyet, J.I and 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.