# Technical Infrastructure of the COSMOS Portal

doi:10.3991/ijet.v3i1.761

N. Doulamis[1], C. Psaltis[1], A. Doulamis[2], S. Sotiriou[3] and M. Sotiriou[4]
1 National Technical University of Athens, Athens, 2 Technical University of Crete, Chania,
3 Ellinogermaniki Agogi, Athens, 4 iKnowHow S.A., Athens, Greece

***Abstract*—This paper presents the main operations and technologies implemented in the framework of the EU funded COSMOS project. COSMOS introduces an advanced web repository which allows teachers and students to search, retrieve, access educational content and re-use educational material for creating earning activities through a specifically designed web interface incorporating innovative technological solutions. The repository is based on an IEEE LOM representation of the content which supports educational scenarios and learning activities as well. The architecture also supports tools for describing and managing digital content rights, which are interoperably represented using the Creative Commons Rights Expression Language (ccREL).**

***Index Terms*—Drupal in Education, Educational Metadata, Educational Technology, Web Technologies**

## I. INTRODUCTION

At the European Council in Lisbon in March 2000, the European Union set itself the ambitious objective of becoming "*the most competitive and dynamic knowledge-based economy in the world*". In this perspective, the European Council underlined that there is an augmented need for Europe's education and training systems to adjust to the needs of the knowledge-based society. The penetration of innovative technological applications with Internet access in education may facilitate the accomplishment of this goal.

Today, the recent advances on Web and computer technology can transform classrooms to research laboratories. More especially, multimedia technologies, computer graphics and visualization tools make teaching of Science more attractive and fascinating [1], [2]. The recent evolution in software and hardware, like for example the development of low cost multimedia capturing devices (video cameras, scanners, etc) or the implementation of sophisticated video compression methods, produces a huge amount of content [3]. It is anticipated that more than 1 Exa-bytes (millions of Terra-bytes) of new information are annually created in the Web and the capacity of the Internet in 2010 will reach the amount of Yota-Bytes ($10^{24}$ bytes) [4]. However, the generated content is unstructured (untagged) and thus all this information is actually lost, unless appropriate mechanisms are developed for the organization and management of multimedia content making multimedia information easily searchable, accessible and retrievable.

Traditional search engines, like Google and Yahoo, are not able to discover educational multimedia material suitable for specific learning needs. For example, it is quite difficult to find multimedia content which describes the first Newton's law appropriate for the first classes of high school.

To overcome this difficulty, recently, many educational repositories have been developed with the purpose of archiving educational materials and organizing them under an intelligent framework, which will make them easily accessible and searchable from interested users (e.g., teachers/pupils). This target is promoted by the European Union as Reflected in the eEurope 2008 Action Plan "Developing a better understanding of the role of science in society and bringing science and scientific subjects closer to the citizen is expected to help increasing young people's interest in science and scientific careers". Furthermore the "i2010" framework [5] has among others as main objective to "Create a "borderless European information space" including an "internal market for electronic communication and digital services".

The development of Portals and Repositories aims to offer young people the opportunity to use scientific content from a wide range of resources and access digital services. Finally it supports the provision of key skills to future citizens and scientists (collaborative work, creativity, adaptability, intercultural communication). One of the first attempts towards the direction of developing an educational repository is the Discovery Space Portal, which is supported by the Dspace project [6]. The portal aims at developing a virtual thematic park that exploits astronomical data in order to simulate teaching of Science either in schools or in universities. However, the data in Dspace are not stored based on a structured, sophisticated framework, making content access not efficient, especially in case that a huge amount of information is stored in the repository. Another interesting attempt is the Colorado educational simulator, over which a set of virtual educational scenarios are graphically created to present different physical laws [7], [8]. Again, the stored data in the Colorado interface do not follow a particular structure (metadata schemas) which allows for easy data interchange among other similar databases.

To address these difficulties, the Institute of Electrical and Electronics Engineering (IEEE) supports a standard, called IEEE Learning Object Metadata or IEEE LOM [9], which permits metadata structuring of educational material so as to achieve easy searching, accessing and exchanging of educational content. IEEE LOM is an extensive XML schema which encodes learning metadata used for describing educational and learning activities. The web repository of MERLOT (Multimedia Educational Resource for Learning and Online Teaching [10]) portal follows the IEEE LOM standard. Merlot supports different types of learning objects, such as simulation, animation, drill and practice, lecture/presentation and so on. However, it does not adopt a common learning enrichment strategy for encoding the metadata structure so as to guarantee easy sharing of the educational content among different educational systems [11].

Another problem of the aforementioned approaches is that educational content metadata are not enough for providing interoperability between different educational repositories and easy content sharing. This is due to the fact that different institutions, schools and educational organizations use different educational vocabularies for describing similar educational content. The aforementioned problem is more evident at European level than in USA, due to the heterogeneity of the European learning systems. For this reason, we need common educational content enrichment strategies being developed throughout Europe. In addition, we need software tools and repositories that allow for the users to a) tag educational content with common metadata structures, b) encode these content using the IEEE LOM standard for interoperability, c) search, retrieve and browse content of interest and finally d) re-use existing content through intelligent authoring tools.

All the aforementioned aspects are addressed in the framework of COSMOS project [12], which allows for a common metadata strategy using common Science vocabularies at pan European level. In this paper, we present the main operations, technologies and functionalities used in the COSMOS portal.

In COSMOS, content description is performed using metadata representation tools through XML-based schemas of the IEEE LOM standard [9]. The description is classified with respect to domain specific vocabularies while multi-lingual searching on them is supported. The proposed approach creates fascinating opportunities to interact with the digital content in novel ways. Using the COSMOS portal a user is able to i) search for educational content in several European languages, ii) retrieve learning activities and iii) construct new content by combining existing ones [13]. To fulfill these objectives, COSMOS portal integrates state of the art World Wide Web technologies through the open source content management system (CMS) Drupal [14]. Furthermore, tools for describing digital intellectual property rights (IPR) are supported. IPR are formulated based on the publishers' needs. This means that the user is able to select the license under which he or she would like to publish new content. The possible selections range within the Creative Commons family of licenses [15]. The specific license is interoperably represented using the Right Expression Language of the MPEG-21 [16]. In particular, in the framework of the project, we support the ccREL (Creative Commons REL) specification [17], [18]. ccREL is based on an RDFa expression of triplets and it is embedded into the HTML of the content [19], [20].

This paper is organized as follows: Section 2 describes issues that concern the design phase of COSMOS portal. An overview of the COSMOS portal is given n Section 3. The upload content functionality of COSMOS is described in Section 4, while the search and retrieval capabilities in Section 5. Intellectual Property Rights are discussed in Section 6. Educational use cases supported by COSMOS are mentioned in Section 7. Finally, conclusions are drawn in Section 8.

## II. DESIGN OF COSMOS PORTAL

A concrete design step ensures limited development problems and low development time. The main objective of the design is to build a system stable, secure and fast with an open architecture which can be flexible, easily upgraded and ported to a number of hardware platforms.

The web server selected for COSMOS Portal is Apache [21]. The Apache server is a well known web server developed and maintained by an open community under the auspices of the Apache Software Foundation. The application is available for a wide variety of operating systems and it is currently the most popular HTTP server on the Web. In particular, the survey of [22] shows that Apache served 49.12% of all websites. The second basic software component is the database server which stores the entire portal data. In particular, MySQL [23] was selected a relational database management system (RDBMS). The program runs as a server providing multi-user access to a number of databases. Both systems were selected since i) they present many advantages compared to other approaches, ii) they are open source and free and iii) they have been widely used and tested.

The last step of the design process was to select the most appropriate developing strategy and the software tools that would offer the best results. Concerning the development strategy, Rapid Application Development (RAD) [24] and Agile methodologies [25] were selected as the most effective alternative. These involve iterative development of prototypes which are continuously adopted and upgraded. These strategies offer working prototypes quickly and effectively tailor the final product to the exact needs of the users. Even if a detailed user's needs requirements analysis has been successfully conducted, it is possible that additional needs will come up during development. These needs are very difficult to incorporate and deal with if methodologies other than RAD-Agile are used.

Today RAD-Agile is a standard development strategy mainly due to the wide variety of available tools that support it. These tools are platforms of different technologies and ready to be used as modules. Tools of this type that better suit the purposes of the COSMOS Portal are called Content Management Systems (CMS).

### *A. CMS Selection for COSMOS portal*

A content management system (CMS) is a computer application used to create, edit, manage, and publish content in a consistently organized fashion. CMSs are frequently used for storing, controlling, versioning, and publishing industry-specific documentation such as news articles, operators' manuals, technical manuals, sales guides, and marketing brochures. The content managed may include computer files, image media, audio files, video files, electronic documents, and Web content. CMSs are widely used for educational purposes mainly in organizing curricula online and for distance learning.

To select the most appropriate CMS a survey of all the available solutions, the advantages and disadvantages of each one was conducted. CMSs can be roughly categorized in the following types according to their main type of use:

- blogging (e.g., WordPress)
- social networking ( e.g., Elgg)
- wikis (e.g., Mediawiki)
- education (e.g., Moodle)
- general use (e.g., Plone)

Keeping in mind that COSMOS Web Portal would be a complex platform for educational purposes that would also facilitate community aspects and collaboration, the most suitable choice were the general use CMSs which offer greater flexibility and more options in creating exactly what you need. The educational brunch was also a possible option; unfortunately CMSs in this category are mainly targeted to small traffic sites, like school specific sites, and thus lack the robustness and features to fuel a high traffic site with a multitude of functionalities like the COSMOS Portal.

The three most prominent general use CMSs are Drupal [14], Joomla [26] and Plone [27] with a continuous presence in the prestigious Packt CMS Awards [28]. They all share similar characteristics, but they also have significant differences. To better illustrate some of these refer to Tables I and II.

TABLE I. BASIC CHARACTERISTICS OF DRUPAL, JOOMLA AND PLONE.

| CMS | Client side coding | Server side coding | Web server support | Database server support | Lice-nse |
|---|---|---|---|---|---|
| **Drupal** | CSS, XHTML, Java-script, AJAX | PHP | Apache, IIS | MySQL, PostgreSQL | GNU GPL |
| **Joomla** | CSS, XHTML, Java-script, AJAX | PHP | Apache | MySQL | GNU GPL |
| **Plone** | CSS, XHTML, Java-script, AJAX | Python | Zope (can be connected to Apache) | ZODB ( can be connected to MySQL) | GNU GPL |

TABLE II. MORE CHARACTERISTICS OF DRUPAL, JOOMLA AND PLONE.

| CMS | Add-ons | Multi-lingual | Taggi ng | Custom content types | Speed | Securi ty |
|---|---|---|---|---|---|---|
| **Drupal** | 2265 | Very good | Exten sive | Yes | Medi um | Mediu m |
| **Joomla** | 3865 | Good | Limite d | Yes | Medi um | Low |
| **Plone** | 903 | Very good | Exten sive | No | Low | High |

Their basic difference is the server side coding language used [29]-[33]:

- Drupal and Joomla go for PHP, while Plone for Python. Python incorporates object oriented support whereas PHP only tries to emulate it (available only for PHP versions above 5).
- Plone uses the Zope environment as web server which also integrates ZODB, an object oriented database server. In any case these can be linked with Apache and MySQL which is the configuration set up of COSMOS server.
- They are all free and open source software.
- Joomla has the largest collection of add-ons, with Drupal coming second and Plone the last. However, not all of Joomla's extensions are free and open source whereas in Plone and Drupal all the offered add-ons are open source.
- Joomla has limited tagging abilities, making difficult to use metadata in order to semantically label content within the system, a problem that would have a very negative impact to the COSMOS Portal. On the other hand, Plone and Drupal offer extensive out of the box support for metadata tagging.
- Plone is slower than the others.
- Plone is the most secure system, with Drupal coming second and Joomla last.
- Multilingual support in Plone and Drupal is equally good and all the interface and content within the site can be translated in any number of languages. Joomla is relatively behind as its dedicated multilingual support package Joomfish is rather weak.
- Drupal and Joomla have considerably large communities, larger than Plone, supporting the projects, but they all offer good documentation and community support.
- Weighting all of the above and especially *speed, multilingual capabilities, security and metadata capabilities*, Drupal came in first with Plone very close and second. Because the difference was very small, we decided to experience firsthand the capabilities of these Drupal and Plone and select the proper one. To this end we set up one basic installation of Drupal and one of Plone and tested them extensively in all the aspects that interested COSMOS. Although Plone offered better tools and its code base was very well written the speed factor was against it. It took almost 25% more time for a similar page to load and render when using Plone in comparison to Drupal. The test was carried out with the same PC using the Microsoft Internet Explorer 6 and 7, Mozilla Firefox 2.X and 3.X web browsers, which are the most popular. This was considered very critical, since the users of the COSMOS Web Portal are expected to have average internet speed and middle to low end hardware capabilities (mainly schools). After this test Drupal was chosen as the best solution.

### *B. Drupal arhitecture*

Drupal treats most content types as variations on the same concept: a node. pages, blog posts and news items (some possible node types) are all stored in a common pool, and the sitemap (its Information Architecture) is an overlay that is designed separately by managing and editing navigation menus. It is a lot like the separation you find in standards-compliant page coding – xhtml provides the meaningful structure of the information.

The thing about keeping these layers separate is that it is a no-brainer to provide completely remixed sitemaps for different user types just by serving them a different navigation menu based on their login information. Pages could be grouped differently, prioritized in a different order based on user needs, and various functions and content could be shown/hidden - on a per-user-type basis.

Integral to understanding Drupal is having the right concept of how things flow within the system. Drupal is cleanly separated into different layers that keep things organized and flexible. There are five main layers in the Drupal system (Figure 1):

1. At the core of the system is the big bucket of nodes – the data pool. Before anything can be displayed on the site, it must be input as data.

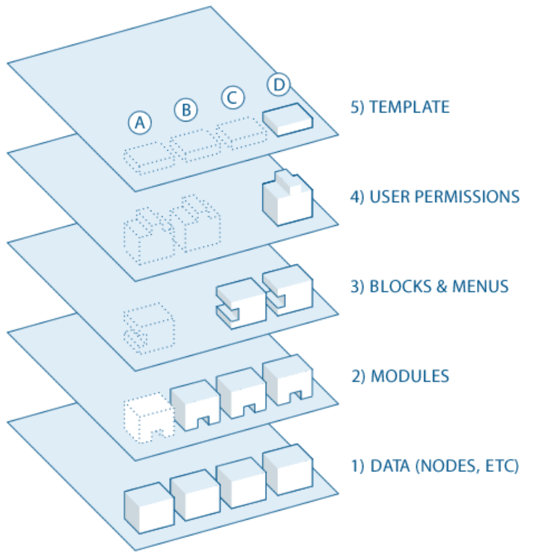


Figure 1. The Drupal flow [14].

2. The next layer out from the center is where modules live. Modules are functional plug-ins that are either part of the Drupal core (they ship with Drupal) or they are contributed items that have been created by members of the Drupal community. Modules provide various functionality to expand your site's capabilities to include things like the creation of custom data points (fields) for your nodes; event calendars; e-commerce; programmatic sorting and display of content (custom output keyed off of any number of configurable parameters that interrelate your content); and more. There are hundreds of different options within the fast growing repository of contributed Drupal modules. They represent the work of everyone from individuals to large corporations like Sony who use and rely on Drupal and are working to extend its power and usefulness.
3. At the next layer, we find blocks & menus. Blocks often provide the output from a module and can be placed in various spots in your template (theme) layout. Blocks can be configured to output in various ways, as well as only showing on certain defined pages, or only for certain defined users.
4. Next are user permissions. This is where settings are configured to determine which things different user types have access to. Permissions are assigned to various roles, and in turn, users are associated with those various roles in order to grant them the associated permissions.
5. On the surface layer is the site template. This is made up predominately of XHTML with some PHP tokens sprinkled through to insert content from the system into the correct spots. Also included with each template is a set of functions that can be used to override standard functions in the modules in order to provide complete control over how the modules generate their markup at output time. Templates can also be assigned on-the-fly based on user permissions.

## III. COSMOS PORTAL OVERVIEW

Figure 2 presents an overview of the COSMOS Portal architecture. COSMOS Portal is comprised by three major components;

- educational material organized in two distinct data types,
- users and their individual roles in the portal,
- functionalities provided to users.

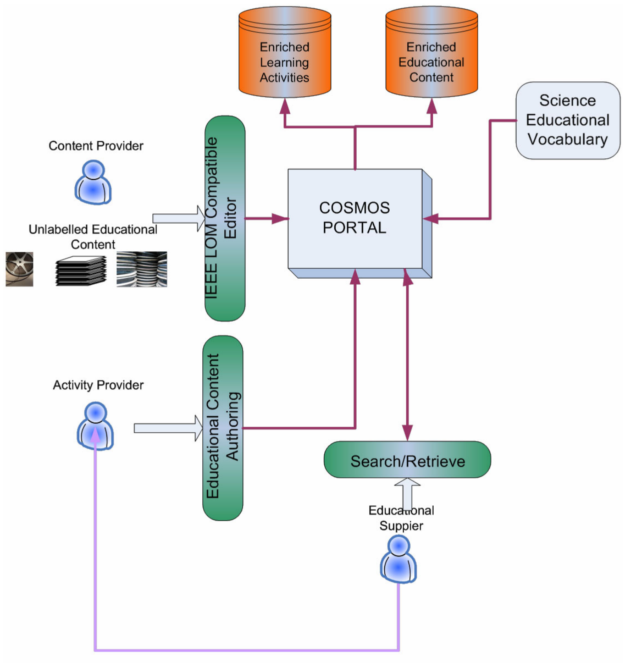


Figure 2. The COSMOS architecture.

### A. *Educational material data types*

There are two main types of educational material stored in the COSMOS repositories; *the educational content and the learning activities.*

Educational content is simple educational material enriched using learning metadata Metadata description follows the IEEE LOM standard.

Learning activities are educational materials, also enriched with metadata, structured according to specific educational models. COSMOS supports six different educational models;

- the guided research,
- the ICT supported culture awareness,
- the inquiry based teaching,

- the project-based teaching,
- the 5E instructional model and
- the learning cycle model

### *B. User roles*

As far as user roles are concerned, the portal supports six types of users; i) anonymous, ii) authenticated, iii)editor, iV0 educational content provider, v) translator and vi) administrator. Anonymous users are guests that have limited access to the portal's functionality. They are able to search and retrieve educational material. However, they can submit new only upon a positive agreement of an authenticated user. Authenticated users are able to search and download content while they can comment on material of interest. Editors are responsible for managing the submitted content and evaluating its quality. The educational content providers can submit educational material to the portal and edit only their own creations. Finally, the translators are responsible for translating the content in other languages. The COSMOS Portal supports multiple roles per each user.

### *C. Functionalities*

COSMOS portal supports three main functionalities:

- *Uploading*: This functionality allows for registered users of the COSMOS portal to upload to the repository tagged educational content and/or learning activities. COSMOS architecture supports two innovative technologies for data stored in its repositories; the IEEE Learning Objective Metadata (IEEE LOM) standard [9] and the IMS Learning Design (IMS LD), which is a specification for a metalanguage that enables modeling of learning processes [34]. These two innovative technologies of COSMOS allow the stored content to be easily exchangeable, portable and accessible.
- *Search and Retrieval*: This functionality allows simple or registered users to semantically search for educational content or learning activities and finally download the material most suited to their needs. Semantic searching exploits the educational /learning activities metadata as being interoperably described using the IEEE LOM. Multilingual technologies are also supported in the framework of COSMOS. This means that educational semantics are automatically translated in a series of European Languages through the exploitation of Science Educational Vocabularies which are related with the formal European Science Curricula [11].
- *Intellectual Property Rights functionalities*: These functionalities are responsible for enriching COSMOS educational content and learning activities with rights metadata. The structure of the ccREL (Creative Commons Right Expression Language) is used in the framework of COSMOS project [17]. ccREL is a specification that describes how license information may be described using RDF triplets [19][20]. In fact ccREL is an RDFa for HTML Web pages and resources referenced therein.

## IV. UPLOADING FUNCTIONALITY

Figure 3 presents the main operations involved with the uploading functionality. Initially, a user (educational content provider) utilizes software tools for enriching his/her content with IEEE LOM standard metadata. These metadata should include the vocabularies developed by the COSMOS consortium in order to better express the context of Educational Content and Learning Activities with respect to official educational curricula.

Then the raw material and its respective metadata, represented in an XML file, are submitted to the portal. At this point an XML parser [35] is activated with the purpose to check the validity of the metadata files and to extract the metadata from the file. In case an error occurred, a message appears with instructions on how to solve the problem. Otherwise, the material files and their respective metadata are stored in COSMOS repository.

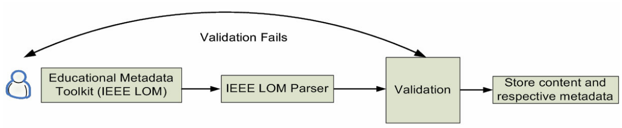


Figure 3. The main operations regarding the upload activity.

### *A. IEEE Learning Object Metadata*

IEEE LOM comprises a hierarchy of elements shown in Figure 4. At the first level there are nine categories, each of which contains sub-elements; these sub-elements may be simple elements that hold data, or may themselves be aggregate elements, which contain further sub-elements. The semantics of an element are determined by its context: they are affected by the parent or container element in the hierarchy and by other elements in the same container.

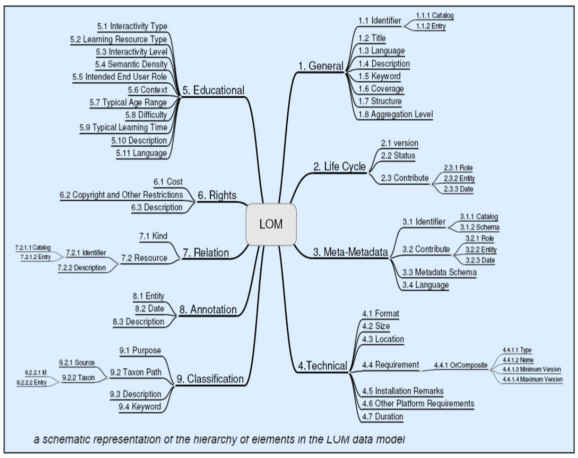


Figure 4. The COSMOS toolbox and the COSMOS ASK-LOM AT.

The data model specifies that some elements may be repeated either individually or as a group: for example the elements 9.3 (*Description*) and 9.1 (*Purpose*) can only occur once within each instance of the *Classification* container element. However the *Classification* element may be repeated, thus allowing many descriptions for different purposes.

The data model also specifies the *value space* and *datatype* for each of the simple data elements. The value space defines the restrictions, if any, on the data that can be entered for that element. For many elements the value space allows any string of Unicode character to be entered, for other elements entries must be drawn from a declared list (i.e., a controlled vocabulary) or must be in a specified format (e.g., date and language codes). Some element datatypes simply allow a string of characters to be entered, others comprise two parts as described below.

- LangString items contain Language and String parts, allowing the same information to be recorded in multiple languages.
- Vocabulary items are constrained in such a way that their entries have to be chosen from a controlled list of terms. Such elements are composed of Source-Value pairs; the Source should contain the name of the list of terms being used and the Value should contain the chosen term.
- DateTime and Duration items contain one part that allows the date or duration to be given in a machine readable format, and a second that allows a description of the date or duration (for example “mid summer, 1968”).

### *B. XML parsing technology of the COSMOS Web Portal*

As COSMOS Web Portal was built in Drupal which uses the PHP 5.X scripting language, we have decided to create the XML parser in PHP in order to integrate seamlessly in the Drupal platform. There are four different ways to parse an XML file in PHP using a) Regular PHP expressions, b) Simple API XML (SAX) parser, c) Document Object Model (DOM) library and d) XML Path Language (XPATH) on top of the DOM library. Among them, the most popular techniques are the SAX and DOM parsers, because they are of more versatile and effective.

#### *1) Simple API XML - SAX*

The SAX parser is a serial access API which runs on a callback model. Every time a tag is opened or closed, or any time the parser sees some text, it makes callbacks to some user-defined functions with the node or text information.

SAX parsers are lightweight because they do not keep anything in memory for very long, so it can be used for extremely large files. The disadvantage is that writing SAX parser callbacks can be complicated and the code written is difficult to read. For that reason SAX parsers are very popular in applications dealing with very simple structures no matter how large the actual file sizes are.

#### *2) Document Object Model -DOM*

Instead, DOM is a platform- and language-independent standard object model for representing HTML or XML and related formats, a specification of W3C [35]. In the case of XML files the DOM library reads the entire document into memory and represents it as a tree of nodes.

The user is able to traverse through the nodes using a set of commands and retrieve the values needed. Although the tree structure makes DOM’s use more intuitive it requires more memory since it first stores the whole document there. This is not a problem if the XML files are small, like the case of COSMOS, what poses a problem though is the complexity of the document to be parsed. If there are many nesting nodes, if the document tree has many levels and even worse when some nodes in different branches share the same name, then using the DOM commands to traverse though the nodes can be rather difficult and slow.

#### *3) XML Path Language - XPATH*

In order to solve the last problem XPATH was introduced. XPATH is a language for selecting nodes from an XML document. In addition, XPATH may be used to compute values (strings, numbers, or boolean values) from the content of an XML document. The XPATH language is based on the same tree representation of the XML document as DOM is and provides the ability to navigate around the tree, selecting nodes by a variety of criteria and with a more user friendly syntax. This syntax allows users to access node values and attributes as if navigating though a command line operating system. Unfortunately it lacks speed when compared to DOM.

Carefully weighting the advantages and disadvantages of each approach the DOM parser was selected to be implemented since it combines good speed with clear syntax. All the developing work was completed using the Eclipse PHP Development Tools (PDT).

#### *4) Embedding the parser to publishing workflow*

With the parser ready and operational the next task was to embed it to the actual code of the portal. There were two options to do that; to run the parsing script periodically or to run it real time while the object was uploaded. The first approach saves processing power as parsing can be scheduled during the time that the server has low stress, but in any case it slows the publishing workflow down. On the other hand, by doing it online at the time the file is submitted places more stress on the server, but the item is ready to be indexed and published right away. Furthermore if there are any problems in the xml the system can automatically detect them and propose right away a possible solution to the user. These last two advantages led to the adoption of the online approach. This way, users submitting their work can see it online almost immediately after submission, a feature that contributes to user satisfaction.

### *C. An Example of Uploading Content*

At this point an overview of the whole publishing procedure is presented to give the reader a better understanding on how content is published in the COSMOS Portal.

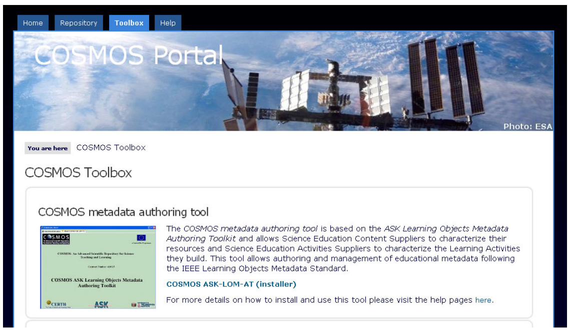


Figure 5. The COSMOS toolbox and the COSMOS ASK-LOM AT.

To give a more complete view of publishing, it is supposed that the user has just created the material and has not yet authored the XML file with the object’s metadata. The first action needed is to visit the portal, log in or create a new account and then visit the toolbox section to find the IEEE LOM compatible editor for metadata authoring. In our implementation, the ASK-LOM AT was used (Figure 5).

When metadata tagging is completed one should visit the portal again, log in and in case he/she wishes to upload Educational Content he/she should follow the link Submit Educational Content. This link takes you to the material submission form (Figure 6). There, you select a title for your work and state if the submitted object is a file or a URL address. In case you have a file you wish to upload then select "file" in the drop down menu and then in the "Educational Content file" field press the browse button to find your file in your hard disk. If you wish to upload a URL address, then select "URL address" from the first drop down menu and then write the address you like under the "URL address of your educational content". In any case you also have to attach the XML file you authored, in the "Metadata XML file". The object is uploaded and published by pressing the "Submit" button at the bottom of the page. Respectively the same procedure should be followed for submitting a Learning Activity. At this point the parser is executed and the metadata are written to the database. If an error occurs during this process (e.g., an error in the XML structure) the system automatically notifies you what is wrong and what you should do to repair it.


Figure 6. Educational content submission form.

The major difference between submitting Educational Content and Learning Activities is the fact that Learning Activities should go through an additional authoring step in order to for the author to create the activity's structure. To author a Learning Activity the user has to visit the Toolbox section of the portal, download and install the IMS LD compatible tool and proceed with authoring. In COSMOS the ASK-LDT tool is being used [36].

## V. SEARCH AND RETRIEVAL

Figure 7 presents the main operations regarding the search and retrieval functionality of the COSMOS interface. A set of different search interfaces is used in COSMOS; i) free text interface that enables user to look for content of interest using keywords, ii) tag cloud which visually depicts the most used semantic tags, iii) simple filter search which contains the most basic educational and learning activity metadata and finally iv) advanced filter search which expands simple filter search with additional metadata. All the four supported interfaces allow for the users to perform their search using either educational and/or learning activity semantics. In addition, COSMOS supports the use of curriculum based searching which is implemented through Science curriculum vocabularies. All the search functionalities of COSMOS are developed in a multilingual framework.

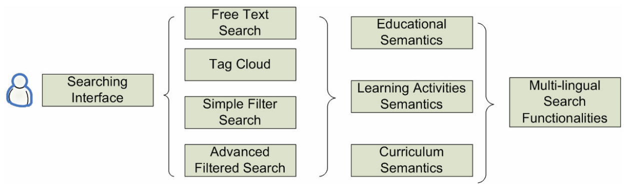


Figure 7. The main operations regarding the search activity.

### A. *Text Search*

The text search interface provided by the core Drupal module *Search* was tested during the first development stages of the portal and it was relatively slow, thus it was replaced by *Faceted Search* which offers a faster and more robust search engine together with an extra feature that proved to be very useful, the "continue searching" block. This block can be inserted in every page of search results giving users the ability to search within the already returned results. This way you can slowly filter out non relevant objects and discover what you are looking for. In every case, query results are returned ordered from the most relative to the least relative.

For the search engine to work properly, content of the site must be indexed regularly. If an object is not indexed it cannot be found because the search engine ignores its existence. It is evident that the time delay between two indexing sweeps affects greatly the searching quality. Unfortunately, indexing cannot be set to occur in short time periods as it places considerable stress on the server. Trying to balance between the two, indexing takes place automatically every 3 hours.

### B. *Semantic search*

The second interface implemented was a view able to search and retrieve results based on their metadata tags, green rectangles). A view is a Drupal native way of constructing queries through an interface of user exposed filters.

One view was created for searching the repository of Educational Content and one for the repository of Learning Activities. The first one has three filters; for Classification, Language and Age Range terms. The second has the same filters and one additional for Educational Model terms. These filters seemed appropriate because they give users the ability to search with only a few clicks selecting terms from the most interesting categories. The results from a query of this type are returned in alphabetical order.

There are two great advantages of this type of search, besides the fact that it is semantic based; it does not need indexing and it can be multilingual. The view queries directly the database and since an object is published it can be discovered right away. As for the multilingual aspect, if all the terms of the vocabularies are translated then switching between languages does not cause any problems in searching, meaning that one can search with the German terms and get exactly the same results as one who is searching with the English.

### C. *Tag cloud*

The tag cloud is a set of weighted Classification terms. The bigger the font size a term is depicted with, the more popular it is in tagging content of the Portal. This way the community of the COSMOS portal actively builds the tag cloud which is dynamically changing over time.

Although the tag cloud by itself cannot be considered a search interface, the terms in the COSMOS Portal tag cloud are links leading to a list of all the educational items tagged with the respective term. This list is accompanied by the "continue searching" block which enable user to search within the results. This way the user sets out to discover something that might be of interest without knowing a priori exactly what this could be. The only thing that he/she should know is just the general area to search or maybe it's just curiosity about a very popular term.

### D. *Linked vocabulary terms*

In the bottom of each page presenting a repository item, a full list of its metadata tags can be found. Each of these terms is a link which, as in the case of tag cloud terms, leads to a list of all the items marked with it. This list also includes the "continue searching" block giving additional searching capabilities.

## VI. INTELLECTUAL PROPERTY RIGHTS

Digital Right Management (DRM) needs technologies to protect and securely deliver digital content. To achieve this, it is also needed to have a Rights Expression Language (REL), which is a formal language used to specify this protection and secure delivery. A REL is a formal language, designed to express rights and conditions for digital content access. These languages can be used for example to control the number of times that a right is exercised over a certain digital content, express the copyright associated to a given digital content, describe an agreement between a content provider and a distributor, or between a distributor and an end user, etc.

Right expression languages (RELs) are languages devised to express conditions of use of digital content. Several RELs have been proposed to describe licenses governing the terms and conditions of content access. The most common rights expression language is that proposed in the framework of the MPEG-21 standard, which is actually based on the eXtensible rights Markup Language (XrML). The MPEG-21 REL specifies a machine-readable language that can declare rights and permissions using the terms as defined in the Rights Data Dictionary (RDD). RDD specifies a dictionary of key terms required to describe users' rights [16].

Figure 8 presents the main operations as far as the activity of the intellectual right property is concerned. In this activity, the educational content providers are able to introduce licenses to their submitted content. In the framework of COSMOS the Creative Commons family of licenses is supported. The licenses are encoded using the ccREL specification. Initially, a rights interface appears to the user over which he/she can define *work/license properties*. Example of the work properties are the attribution name, content description and title. License properties include work permissions, jurisdiction and the conditions under which the permissions are granted.

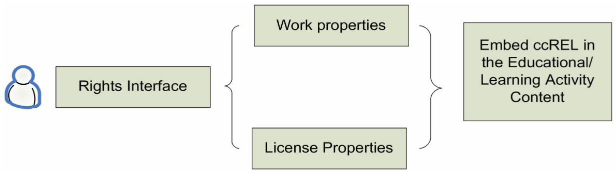


Figure 8. The main operations regarding the intellectual rights property activity.

### A. *Creative Commons*

The Free Software Foundation, which was created by Richard Stallman in 1983, supports the General Public License (GPL) and Lesser General Public License (LGPL) as far as the software code is concerned. These type of licences, in contrast to the traditional "All Rights Reserved", significantly affect the evolution of the software engineering, since they found the legal status for the development of the open source software. In the same direction, the Creative Commons (CC) defines the spectrum of possibilities between full copyright "all rights reserved" and the public domain "no rights reserved". Our licenses help you keep your copyright while inviting certain uses of your work — a "some rights reserved" copyright.

CC licenses let people easily change their copyright terms from the default of "all rights reserved" to "some rights reserved". Creative Commons licenses are not an alternative to copyright. They apply on top of copyright, so you can modify your copyright terms to best suit your needs.

The above described advantages of the Creative Commons family of licenses make it ideal for publishing educational content.

### B. *Right Expression Language*

After selecting the proper license it is time to adopt the proper REL standard. In the COSMOS project the Creative Commons Right Expression Language has been adopted and implemented for the description of the rights metadata. ccREL aims in introducing a machine readable interface that allows content creators to describe rights in their work. ccREL is based on the World-Wide Web Consortium's Resource Description Framework (RDF).

The ccREL consists of a small, but extensible, set of RDF properties that should be provided with each licensed object. ccREL information can be embedded both in web pages and in free-floating files (images, pdfs). The abstract model of ccREL distinguishes two classes of properties:

- Work properties, describe aspects of specific works including under which license the work is published,
- License properties, that describe aspects of the selected license.

Publishers will normally be concerned only with Work properties, since this is the only information they provide to describe a Work's licensing terms. License properties are used by Creative Commons itself to define the authoritative specifications of the licenses they offer.

The publisher of a work must, at a minimum, provide one RDF triple that specifies the work's license property by targeting the Creative Commons URL with the text of the selected license. However it is proposed to include additional triples for giving extra information about the work. The additional work properties proposed are: a) dc:title – the document's title, b) cc:attributionName – the name to cite when giving attribution when the original work is adapted and redistributed, c) cc:attributionURL – the provided URL to link when attributing, d) dc:type – the type of the licensed document, e) dc:source – indicates the original source of the modified work, specified as an URI.

If the above are not considered enough it is possible to add new properties not defined within the standard, but not in the cc: namespace since it is only controlled by Creative Commons.

*1) License properties*

The ccREL recommendation leaves to Creative Commons the task to provide the License properties thus the publisher does not deal with them at all. Typically, the License properties governing a work will be found by URL-based discovery. The license properties defined in ccREL will be briefly presented below mainly because potential tool builders should take them into account.

- *cc:permits* – permits a particular use of the work above what default copyright law predicts. The possible values for this field are, a) cc:Reproduction – copying the work in various forms, b) cc:Distribution – redistributing the work, c) cc:DerivativeWorks – creating derivatives of the work.
- *cc:prohibits* – prohibits a particular use of the work. The possible values for this field is cc:CommercialUse – using the work for commercial purposes.
- *cc:requires* – requires certain actions of the user who employs the permissions given by cc:permits. The possible values for this field are a) cc:Notice – providing an indication of the license that governs the work, b) cc:Attribution – giving credit to the appropriate creator, c) cc:ShareAlike – when redistributing derivatives of the work, using the same license, d) cc:SourceCode – when redistributing this work, used only for software.
- *cc:jurisdiction* – associates the license with a particular legal jurisdiction.
- *cc:deprecatedOn* – indicates that the license has been deprecated on the given date.
- *cc:legalCode* – references the corresponding legal text of the license.

## *C. Rights Metadata in COSMOS*

*1) License selection interface*

When uploading content in the portal the user is asked first to select the actual content file or URL, then to upload the XML file which contains the IEEE LOM encoded metadata and finally to decide the licensing conditions of the work. Because the users of COSMOS are not expected to be experienced in selecting licensing policies, the implemented interface is in the form of a simple questionnaire (Figure 9).


Figure 9. The questionnaire for selecting the proper Creative Commons license.

The first question lets the publisher decide if he/she wishes his work to be used for commercial purposes and the second if the original work can be adapted. Publishers wishing to permit adaptation of their work can also select if the adapted work should be published under the same license as the original. By answering these questions and keeping in mind that all the Creative Commons licenses permit distribution and reproduction, the full range of the six Creative Commons licenses can be automatically generated. Selecting a license is by no way imposed to the users. That is why the Creative Commons license was selected as a default license (in particular in COSMOS portal the license Creative Commons, Attribution, non-Commercial, Share Alike, i.e., BY-NC-SA) in case the user does not select a license himself.

*2) Expressing rights with ccREL*

After answering the questions illustrated above the proper license is automatically generated and the suitable ccREL properties are embedded in the HTML code of the page. From the whole range of available properties the title, attributionURL and attributionName were selected from the Work properties and from the license properties all the permits, prohibits and requires properties were selected.

An example of HTML code generated dynamically by this procedure is illustrated in Figure 10. This format is ideal for machines-software (like browsers) to read. Developers can exploit these tags for their software to use. In fact this is the case for browser plug-ins used to parse these tags and present to humans a user-friendly notation of the current license.

```
   The license of the educational content
   <span property="dc:title">
     <b>
        Quasars      </b>
   </span>
   by
   <a property="cc:attributionName"
      rel="cc:attributionURL"
      href="1277">
        admin
   </a>
   permits:
     <span property="cc:permits" rel="cc:Reproduct

     <span property="cc:requires" rel="cc:Notice">

     <span property="cc:requires" rel="cc:Attribut
</div>
```

Figure 10. ccREL embedded in HMTL code of an educational object.

In order to provide a human readable form of the license without the need of additional software from the user side, license specific icons are used for each licence. These icons are placed in the bottom left of the screen presenting an educational object (Figure 11).

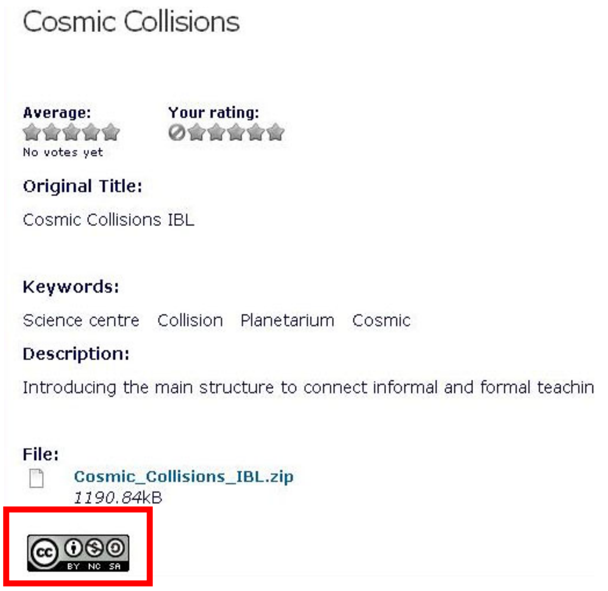


Figure 11. An educational object page. On the bottom left, inside the red rectangle, the license icon is displayed.

These icons serve also as links to the respective Creative Commons web pages which gives further information about the current license (Figure 12).

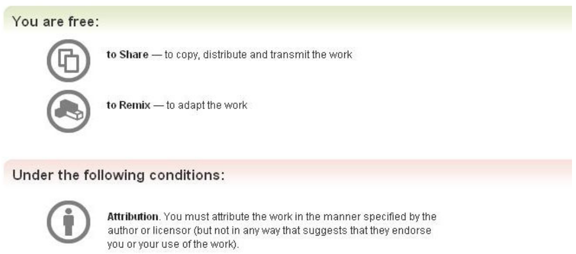


Figure 12. The Creative Commons web page of the license.

## VII. COSMOS USE CASES

Using the COSMOS application students and teachers are able to directly apply the theories learned and taught in the classroom to real, interactive research. They personally experience the procedures involved in an authentic research project and thereby gain a far better understanding of science and engineering. The COSMOS initiative contributes in changing the present situation in science teaching and learning by implementing the following benefits: (i) teaching science through the use of a network of advanced scientific instruments, (ii) reinforcing interdisciplinary approaches and (iii) promoting inquiry based learning.

In the following, an application scenario is described so that COSMOS capabilities are better clarified. The scenario focuses on measuring the asteroids rotation periods. Asteroids present n excellent case for short time observations. Some of them are rotating very quickly (one full turn in less than 24h) giving unique opportunities for observations. The specific activity is introduced in the science curriculum in the framework of the study of periodic motions (at high school or university level). The students select the suitable for the season asteroid and the request from the telescope to conduct the continuous observations for a certain period. After the completion of the observations, series of images of the asteroid will be available in the COSMOS repository for further use (see Figures 13 and 14).

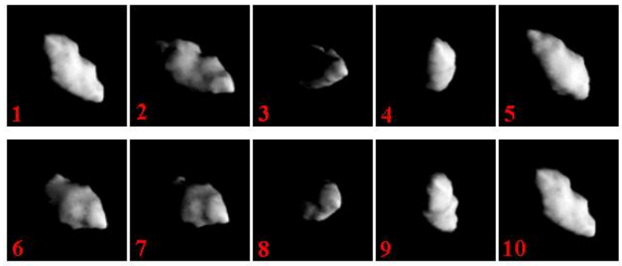


Figure 13. The Sun's light is reflected on the asteroid's surface as it turns and captured by the CCD camera of the robotic telescope which is following the asteroid for a specific period of time requested by the user.

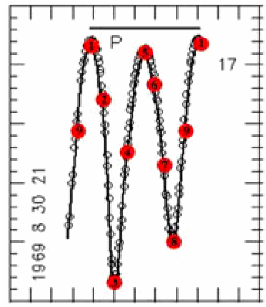


Figure 14. By measuring the light that it is reflected on the asteroid's surface as it turns the students are creating a graph indicating the minimum and maximum values of brightness (picture in the right). The numbers on the graph represent the 10 frames shown in Figure 13.

## VIII. CONCLUSIONS – FUTURE PLANS

In this paper, we present an overview of the technical solution implemented in the framework of COSMOS project. The main purpose of the COSMOS proposal is to create an experimental laboratory for students and teachers in order to improve science instruction by expanding the resources for teaching and learning in schools and universities and by providing more challenging and authentic learning experiences for students. In particular, the project will create a Web repository of educational content using multi-lingual vocabularies that will facilitate end-users' search, retrieval, access and use of both scientific and educational resources. Digital right description strategies are incorporated in the COSMOS portal to guarantee intellectual rights properties issues for the submitted multimedia content.

The future goal of COSMOS is to generate a structured set of recommendations that will form a pan European roadmap for Science learning services to sustain the development/deployment of science educational content services. To develop the roadmap we need to apply the followings actions: (a) to define quality methods and criteria for Science Education Content, (b) to define

Content Enrichment Strategy and (c) to define the Quality Certification Method.

As far as the first issue is concerned, we define *quality methods* as well as the criteria which are used to design, monitor and finally implement the use of the content. This is very important especially for Science Education. To achieve this, we need the development of concepts and strategies for quality development, quality assurance and mutual recognition of the quality of educational contents and systems between countries, as well as the development of internationally recognised quality marks specifically for the field of technology enhanced learning where the use and reuse of digital content is included.

For content enrichment, it is needed to develop a strategy that will act in pan European level. To this aim COSMOS defines and recommends a specific strategy that the EU countries could follow in order to continuously enrich their content in respect with the Science Education. Finally for the quality certification methods, COSMOS includes experts able to analyze the input and then specify certain recommendations for quality certification. It also includes educational authorities able to certify that the COSMOS outcome will be recognized.

Thus, the final outcome for COSMOS will expect to be a pan-European Science Learning Service that will contain: (i) *Educational specifications* (methods criteria and practices for selecting content), (ii) *Technological specifications* (methods and practices for presenting and access the content) and (iii) *Specifications for the certification of the integrated system* (methods and practices).

## REFERENCES


[1] Pui Mun Lee, W.G. Sullivan, "Developing and implementing interactive multimedia in education," *IEEE Transactions on Education*, Vol. 39, No. 3, pp. 430-435, Aug. 1996. (doi:10.1109/13.538769)

[2] A.A *Read*, "Computers and computer graphics in the teaching of field phenomena," *IEEE Transactions on Education*, Vol. 33, No. 1, pp. 95-103, Feb. 1990. (doi:10.1109/13.53632)

[3] R. Datta, W. Ge, J. Li, and J.Z. Wang,"Toward bridging the annotation-retrieval gap in image search," *IEEE Multimedia*, Vol. 14, No. 3, pp. 24–35, 2007. (doi:10.1109/MMUL.2007.67)

[4] Jan Nesvadba, "From push-based passive content consumption to pull-based content experiences," *IEEE International Workshop on Image Analysis for Interactive Multimedia Services*, Santorini, Greece, 2007.

[5] i2010, Europe's Information Society, a European Information Society for growth and employment [Online], Available: http://ec.europa.eu/information_society/eeurope/i2010/

[6] Discovery Space, [Online], Available: http://www.discoveryspace.net.

[7] Colorado Simulation Web Repository, [Online], Available: http://phet.colorado.edu/simulations/

[8] C. E Wieman, and K. K. Perkins "Powerful Tool For Teaching Science," *Nature Physics*, pp. 290-292, 2006. (doi:10.1038/nphys283)

[9] IEEE Learning Technology Standards Committee (2001) 'IEEE LOM Working Draft 6.1', [Online], Available: http://ltsc.ieee.org/wg12/index.html.

[10] Merlot, Multimedia Educational Resource for Learning and Online Teaching, [Online], Available: www.merlot.org.

[11] S. Sotiriou, 'COSMOS: An Advanced Scientific Repository for Science Teaching and Learning,' *The 8th IEEE International Conference on Advanced Learning Technologies*, Santander, Cantabria, Spain, July 1st- July 5th, 2008 http://www.ask4research.info/icalt/2008/ .

[12] COSMOS, An advanced Repository for Science Teaching and Learning [Online], Available: www.cosmos-project.eu.

[13] N. Doulamis, C. Psaltis, A. Georgopoulos, M. Sotiriou S., Sotiriou and I. Doxaras, "The COSMOS Approach for Teaching Science: An Architecture that Combines IEEE LOM compatible Content with Multi-lingual Science Educational Vocabularies and Rights," *7th European Conference on e-Learning* Grecian Bay Hotel, Agia Napa, Cyprus

[14] Drupal, [Online], Available: http://www.drupal.org (accessed 30 Sep. 2008)

[15] Creative Commons, [Online], Available: http://www.creativecommons.org, 2008.

[16] J. Bormans, J. Gelissen, and A. Perkis, 'MPEG-21: The 21st century multimedia framework,' *IEEE Signal Processing Magazine*, Vol. 20, No.2, pp. 53-62, 2003. (doi:10.1109/MSP.2003.1184339)

[17] H. Abelson, B. Adida, M. Linksvayer, N. Yergler, "ccREL: The Creative Commons Rights Expression Language," available on http://wiki.creativecommons.org/images/d/d6/Ccrel-1.0.pdf Version 1.0, March 2008.

[18] N. Doulamis, C.Psaltis and A. Doulamis, "The COSMOS Experience in Managing Educational Rights Metadata and Licences" *International Conference on Educational Research and innovation*, Madrid Spain, 17-19 Nov. 2008.

[19] S. Decker, P. Mitra, and S. Melnik, "Framework for the Semantic Web: An RDF Tutorial," *IEEE Internet Computing*, Vol. 4, No. 6, pp. 68-73, Nov-Dec. 2000. (doi:10.1109/4236.895018)

[20] D. Brickley and R. Guha, "Resource Description Framework (RDF) Schema Specification," *W3C Candidate Recommendation*, Mar. 2000; available at http://www.w3.org/TR/2000/CR-RDF-schema-20000327.

[21] The Apache Software Foundation, [Online], Available: http://www.apache.org/ (accessed 2 Oct. 2008).

[22] Netcraft, "Web Server Survey," [Online], Available: http://news.netcraft.com/archives/2008/06/22/june_2008_web_server_survey.html (accessed 30 Sep. 2008).

[23] MySQL, [Online], Available: http://www.mysql.com/ (accessed 30 Sep. 2008).

[24] J. Martin, "Rapid Application Development," Macmillan Coll Div, May, 1991.

[25] R. C. Martin, "Agile Software Development, Principles, Patterns, and Practices," Prentice Hall, Oct., 2002.

[26] Joomla, [Online], Available: http://www.joomla.org (accessed 30 Sep. 2008).

[27] Plone, [Online], Available: http://www.plone.org (accessed 30 Sep. 2008).

[28] Packt Publishing, "Packt Awards," [Online], Available: http://www.packtpub.com/open-source-cms-award-previous-winners (accessed 30 Sep. 2008).

[29] "Quick Trivia - How many Drupal modules," [Online], Available: http://www.agileapproach.com/blog-entry/quick-trivia-how-many-drupal-modules (accessed 30 Sep. 2008).

[30] M. Bowen, "Plone vs. Drupal, Take One," [Online], Available: http://m.odul.us/2007/09/02/plone-vs-drupal-take-one (accessed 30 Sep. 2008).

[31] Alledia, "Joomla and Drupal - Which One is Right for You," [Online], Available: http://www.alledia.com/blog/general-cms-issues/joomla-and-drupal-%11-which-one-is-right-for-you?/ (accessed 2 Oct. 2008).

[32] Idealware, "Comparing Open Source CMSes: Joomla, Drupal and Plone," [Online], Available: http://www.idealware.org/articles/joomla_drupal_plone.php (accessed 30 Sep. 2008).

[33] Drupal tutorial, "Why Drupal," [Online], Available: http://drupaltutorials.com/books/why-drupal (accessed 2 Oct. 2008).

[34] C. Tattersall, "Using IMS Learning Design to model collaborative learning activities," *6th IEEE International Conference on Advanced Learning Technologies*, July 2006.

[35] L. Wood, "Programming the Web: the W3C DOM specification," *IEEE Internet Computing*, Vol. 3, No. 1, pp. 48 – 54, Jan.-Feb. 1999. (doi:10.1109/4236.747321)

[36] D. G. Sampson, P. Karampiperis and P. Zervas, “ASK-LDT: A Web-based Learning Scenarios Authoring Environment based on IMS Learning Design,” *Advanced Technology for Learning*, Vol. 2, 2005.

## AUTHORS

**N. D. Doulamis** is with National Technical University of Athens, 9, Heroon Polytechniou street, 157 73 Zografou-Athens, Greece (e-mail: ndoulam@cs.ntua.gr).

**C. Psaltis** is with the National Technical University of Athens, 9, Heroon Polytechniou street, 157 73 Zografou-Athens, Greece (e-mail: chpsaltis@yahoo.com).

**A. D. Doulamis** is with the Technical University of Crete, Polytechnioupolis, 731 00, Chania, Crete, (e-mail: adoulam@cs.ntua.gr).

**S. Sotiriou** is with the Research and Development Division of Ellinogermaniki Agogi School, Pallini-Attikis, 153 51, Athens, Greece, (e-mail: sotiriou@ea.gr)

**M. Sotiriou** is with the iKnowHow, 3 Lefkados street, Glyka Nera, 153 54, Athens, Greece, (e-mail: m.sotiriou@gmail.com)

This work is supported by eContent Plus Initiative of the European Commission (Contract No. 2006 – 410025). This article was modified from a presentation at the International Conference of Interactive Computer Aided Learning ICL2008, September 24 - 26, 2008 in Villach, Austria. Manuscript received 05 December 2008. Published as submitted by the author(s).