University of Crete

Department of Computer Science

A framework for federated search in Digital Libraries and Learning Object Repositories providing unified formatted objects delivery (SCORM 2004)

by

Angeliki Psycharaki

Master's Thesis

Heraklion, December 2009

University of Crete
Department of Computer Science

A framework for federated search in Digital Libraries and Learning Object Repositories providing unified formatted objects delivery (SCORM 2004)

by
Angeliki Psycharaki

A thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science

Author: ____

Angeliki Psycharaki, Computer Science Department

Board of enquiry:

Supervisor: ____

Christos Nikolaou, Professor

Member: ____

Dimitrios Plexousakis, Professor

Member: ____

Georgios Georgakopoulos, Associate Professor

Approved by: ____

Panos Trahanias, Professor, Chairman of the Graduate Studies Committee

Heraklion, December 2009

A framework for federated search in Digital Libraries and Learning Object Repositories providing unified formatted objects delivery (SCORM 2004)

Angeliki Psycharaki

Master's Thesis

University of Crete
Computer Science Department

Abstract

Digital Libraries and eLearning evolution has been urged by the advent of Internet technologies. There are many international organizations, consortia and institutions that have been conducting research in these two fields, for the last fifteen years. Although these two worlds are referring to a common audience, they are evolving independently and the convergence between them is deficient. The most important aspect of this non-convergence is the interoperability problem. The purpose of this thesis is to provide an effective proposal for addressing the interoperability problem between Digital Libraries and eLearning. This thesis proposes a framework for federated searches between Digital Libraries and eLearning and supports the delivery of standardized unified formatted objects (SCORM 2004). The thesis' proposition is a bidirectional approach towards the exchange of digital and learning objects by utilizing a widely adopted standard, SCORM 2004. The proposed framework provides a unified interface between Digital Libraries and eLearning, in conjunction with the construction and delivery of SCORM 2004 objects. The framework's functionality is supported by an Application Programming Interface (API) which is based on Web services and utilizes the principles of the Service Oriented Architecture (SOA).

Σύστημα για ομόσπονδες αναζητήσεις σε ψηφιακές βιβλιοθήκες και αποθετήρια μαθησιακών αντικειμένων προσφέροντας παράδοση ενιαίας μορφής αντικειμένων (SCORM 2004)

Αγγελική Ψυχαράκη

Μεταπτυχιακή Εργασία

Πανεπιστήμιο Κρήτης
Τμήμα Επιστήμης Υπολογιστών

Περίληψη

Η εξέλιξη των Ψηφιακών Βιβλιοθηκών και της Ηλεκτρονικής Μάθησης έχει επηρεαστεί από την εμφάνιση των τεχνολογιών Διαδικτύου. Υπάρχουν πολλές διεθνείς οργανώσεις, κοινοπραξίες και όργανα που πραγματοποιούν έρευνα σε αυτούς τους δύο τομείς, για τα τελευταία δεκαπέντε έτη. Αν και αυτοί οι δύο κόσμοι αναφέρονται στο ίδιο κοινό, εξελίσσονται ανεξάρτητα και η σύγκλιση μεταξύ τους είναι ανεπαρκής. Η σημαντικότερη πτυχή αυτής της μη σύγκλισης είναι το πρόβλημα της διαλειτουργικότητας. Ο σκοπός αυτής της μεταπτυχιακής εργασίας είναι να παρέχει μια αποτελεσματική πρόταση για το πρόβλημα της διαλειτουργικότητας μεταξύ των Ψηφιακών Βιβλιοθηκών και της Ηλεκτρονικής Μάθησης. Αυτή η εργασία προτείνει ένα πλαίσιο για ομόσπονδες αναζητήσεις μεταξύ Ψηφιακών Βιβλιοθηκών και αποθετηρίων μαθησιακών αντικειμένων και υποστηρίζει την παράδοση ενιαίας μορφής αντικειμένων (SCORM 2004). Η πρόταση της εργασίας είναι μια αμφίδρομη προσέγγιση προς την ανταλλαγή των ψηφιακών αντικειμένων και αντικειμένων μάθησης με τη χρησιμοποίηση ευρέως υιοθετημένων προτύπων, SCORM 2004. Το προτεινόμενο πλαίσιο παρέχει μια ενοποιημένη διεπαφή μεταξύ Ψηφιακών Βιβλιοθηκών και Ηλεκτρονικής Μάθησης, ταυτόχρονα με την κατασκευή και την παράδοση SCORM αντικειμένων. Η λειτουργία του πλαισίου υποστηρίζεται από μια διεπαφή προγραμματισμού εφαρμογών (Application Programming Interface, API) που είναι βασισμένη στις υπηρεσίες Ιστού (Web services) και χρησιμοποιεί τις αρχές της προσανατολισμένης στις υπηρεσίες αρχιτεκτονικής (Service Oriented Architecture, SOA).

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my supervisor Christos Nikolaou for his assistance and support through this thesis work. I would also like to thank him for his belief in me and the opportunity that he offered to me to be a member of the Transformation Services Laboratory. I would also like to thank both Professors D. Plexousakis and G.Georgakoupoulos for agreeing to be members of my board of enquiry and for reading my work.

I would also like to thank my friend and colleague, Mariana Karmazi, for her invaluable support and courage in my thesis work. I would like to thank George Tesseris for our research discussions during our theses. I would also like to thank my friends that stayed by my side in the difficult times as well as my colleagues of the Transformation Services Laboratory and the Library and Information Center for their valuable help and support.

I would like to thank my mother Maria and my brother Stelios for their belief they have had in me and their encouragement during these years. I would also like to thank George Koutras for his support, encouragement and inspiration.

Finally, I would like to dedicate this work to my father, Manolis.

Chapter One: Introduction

Digital Libraries and eLearning constitute two scientific fields that have emerged due to the advent of Internet technologies. Both of these scientific fields manage valuable collections of digital resources, with the aim to inform and educate their users. Digital Libraries and eLearning are widely accepted and utilized by academia, international organizations and businesses, for the last fifteen years. However, these two worlds are evolving independently and the convergence between them is deficient, even though they address a common audience.

There are many international organizations, consortia and institutions that are conducting research in these two fields. In Digital Libraries there is increased interest in the field of resources' description, i.e. metadata and resources' discovery. Library of Congress (1), Dublin Core Metadata Initiative (2) and Open Archives Initiative (3) are some of these organizations. Widely adopted standards, specifications and protocols in Digital Libraries are MARC (4), Dublin Core Element Set (5), Open Archives Initiative Protocol for Metadata Harvesting (6), Z39.50 (7) protocol and more others. In eLearning there is increased interest in the field of learning objects' metadata, learning object content packaging, sequencing of learning and the communication between learning objects and learning platforms. Organizations that have developed standards in the field of eLearning are IEEE Learning Technology Standard Committee (IEEE LTSC), Aviation Industry Computer Based Training Committee (AICC), IMS Global Learning Consortium (IMS GLC) and Advanced Distributed Learning (ADL) Initiative. Some of the most widely adopted standards are IEEE Learning Object Metadata (IEEE LOM) (8), AICC CMI Guidelines for Interoperability (9), IMS Digital Repository Interoperability (IMS DRI) (10) specifications and ADL Sharable Content Object Reference Model (SCORM) (11).

Association between eLearning and Digital Libraries is a demanding challenge, which includes the resolution of several issues, the more important one being the interoperability problem. The Interoperability problem between Digital Libraries and Learning Object Repositories is based on the fact that the digital and eLearning resources are located in a large number of distributed and heterogeneous repositories that do not provide universal access, whilst the wide diversity of metadata schemes encumbers further progress in defining a universal schema. Moreover, localization and reusability of digital and learning objects is a major issue as it accommodates faster deployment of new systems using already known components and enables new resources' discovery.

The purpose of this thesis is to provide an effective proposal for bridging the gap between Digital Libraries and eLearning interoperability, regarding the issues that were presented above. This thesis proposes a framework for federated searches between Digital Libraries and eLearning and supports the delivery of standardized unified formatted objects. This thesis proposes a unified interface between Digital Libraries and eLearning, in conjunction with the construction and delivery of SCORM 2004 (11) objects, in order to enhance the standardized exchange of digital and learning objects. The framework's functionality is supported by an Application Programming Interface (API) which is based on Web services and utilizes the principles of the Service Oriented Architecture (SOA) (12). According to this framework, any Digital Library and Learning Object Repository, which conforms to several prerequisites, can be incorporated in a collection of repositories that are subject to federated searches, in the context of the unified interface and object delivery. The prerequisite for a Digital Library is to conform to Dublin Core Element Set (5) application profile and for a Learning Object Repository is to conform to IMS DRI (10) and deliver SCORM 2004 objects. The need for the framework presented in this thesis has emerged, as the evolution of a simpler, not generic, approach towards the interoperability between Digital Libraries and some eLearning management systems. This approach, which is included in this thesis, enables the integration of eLearning resources in a Digital Library and it has been applied in the eLearning material provided by University of Crete.

1. Thesis Contribution

This thesis deals with the interoperability problem between Digital Libraries and eLearning. The thesis' contribution to the solution of this problem concerns several aspects. More precisely, the thesis provides a framework for the integration of the whole eLearning material of University of Crete in a Digital Library, which is incorporated in the collection of target repositories of Livesearch[1]. Moreover, this thesis provides a framework for federated searches between Digital Libraries and eLearning and supports the delivery of standardized

[1] Livesearch is the federated searching portal of the Library and Information Center of University of Crete.

unified formatted objects. The thesis' proposition is a bidirectional approach towards the exchange of digital and learning objects by utilizing a widely adopted standard, SCORM 2004.

This functionality is provided through an Application Programming Interface based of Web services, according to the principles of the Service Oriented Architecture (SOA) (12). SOA enhances the integration of heterogeneous services in a distributed environment. All the functionality is implemented based on existing standards such as WSDL (13) and SOAP (14) Web services. The use of XML standards either for the communication and interaction between services (WDSL (13), SOAP (14)) or for the metadata schemes and their mappings will enable the framework to withstand technology evolution and changes and simplify the adoption of newly developed Web services that will enhance framework's functionality.

2. Thesis Organization

The structure of this thesis is designed in such a way as to first introduce the reader into the basic issues related to the areas of Digital Libraries and eLearning. Then, the author presents the interoperability gap between these two worlds and the most important approaches towards its solution. Finally, the approach of the author against the interoperability issue between Digital Library and eLearning is provided. The thesis' organization is presented below.

Chapter One presents the basic information about the thesis's scope, the basic objectives that the thesis aims to achieve and its contribution towards the solution of the interoperability gap between Digital Libraries and eLearning.

In Chapter Two, an overview of what a Digital Library is and how it works is presented. This chapter tries to introduce the reader to the overall idea of Digital Libraries, presenting definitions, standards, practices and the most important approaches addressing the issues that arise.

In Chapter Three an overview of eLearning is provided. The most important definitions regarding eLearning and Learning objects are presented along with an overview of the standards that have been suggested by organizations, which perform significant research in the area of eLearning.

In Chapter Four, the most important parameters of the interoperability gap between Digital Libraries and eLearning are presented, along with efforts made towards the solution of this issue.

Chapter Five presents the thesis' approach to the interoperability problem between Digital Libraries and eLearning. This chapter is divided into two sections, the first refers to the integration of eLearning resources in a Digital Library and the second refers to a framework for federated searches between Digital Libraries and eLearning and unified formatted object delivery.

Chapter Six presents how the implementation of the frameworks is organized and which technologies, tools and standards are used for this purpose.

Finally, in the last chapter the conclusion of the thesis and recommendations about future work upon the framework's functionality are presented.

Chapter Two: Digital Libraries

In this chapter, an overview of what a Digital Library is and how it works is presented. The basic principles and representative standards concerning metadata schemes (MARC, Dublin Core, METS) and searching mechanisms (Z39.50, SRU/SRW, OAI-PMH) are presented, since resources' description and discovery are two of the most important issues concerning Digital Libraries.

1. Digital Library

Digital library is the term that realizes an old concept, introduced as a "computerized library", long time ago. Several definitions of the term have been provided during the lifecycle of the concept. Some of them insist on narrower definitions, based on principles of traditional libraries and some of them attach to the term a service-based perspective. This section of the thesis will provide an overall idea of what a Digital Library is and how it works.

A Digital Library was initially defined as an "electronic information access system that offers the user a coherent view of an organized, selected, and managed body of information" (15). Other representative definitions that are worth quoting are presented by Greenstein and Thorin (16), "Digital Library is a library in which collections are stored in digital formats and are accessible by computers" and Arms (17), "Digital Library is a managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network", who underlines in this definition that the information is managed.

Complementary to the above definitions, DELOS[2] defines Digital Library as "a term that is currently used to refer to systems that are heterogeneous in scope and yield very different functionality. These systems range from digital object and metadata repositories, reference-linking systems, archives, and content administration systems (mainly developed by industry) to complex systems that integrate advanced Digital Library services (mainly

[2] DELOS is a Network of Excellence on Digital Libraries partially funded by the European Commission in the frame of the Information Society Technologies Programme (IST), http://www.delos.info/

developed in research environments) ". (18) This divergence of definitions is a result of lack of consensus on what Digital Libraries are and what functionality is associated with them.

The scientific life of Digital Libraries is relatively short, since they constitute a field of scientific interest for less than 15 years. Digital Libraries represent the meeting point of many disciplines and fields, including data management, information retrieval, library sciences, document management, information systems, web applications, image processing, human–computer interaction, digital curation and service science. Through these years, great effort has been conducted to the area of bridging the gaps between these disciplines. This effort has resulted to several solutions towards Digital Library functionality. (18)

A Digital Library is an evolving organisation that demands a series of development steps to combine the necessary constituents. According to Delos (19), this organisation requires three distinct notions of "systems" developed along the way forming a three-tier framework: Digital Library, Digital Library System, and Digital Library Management System. These correspond to three different levels of conceptualisation of the universe of Digital Libraries.

The keystone of a Digital Library is the efficient description of its content. As it was stated in the definitions, a Digital Library may be constituted by a great variety of collections (digital objects, metadata repositories, archives, systems etc). The key to integrated use of such a variety of collections in a Digital Library is *metadata* that efficiently represents the characteristics of each asset.

Metadata is the main tool for describing digital assets and enabling their capacity to be discovered by interested users. The word "metadata" means "data about data". Metadata articulates a context for objects of interest - "resources" such as documents, media files, library books, or satellite images - in the form of "resource descriptions". Especially, metadata is all the machine recognizable data related to data. The first appearance of resource description dates back to the earliest archives and library catalogs. A number of metadata schemes have emerged in recent years to provide frameworks to assist in the development of meaningful strategies to support the discovery of digital resources. Perhaps the most well-known metadata standard is the Dublin Core standard that emerged in mid 1990s (5). While the Dublin Core was among the first set of standards that systematically attempted to provide a system for describing digital resources, many more exist. Some of the most important metadata schemes used in Digital Library world are MARC, METS and Dublin Core, which are presented in the section 4 Metadata Standards.

Digital Libraries provide searching mechanisms for the discovery and exposure of their resources. These resources cannot typically be located by search engine crawlers, because they are deep web[3] resources. The most common strategies used for searching Digital Libraries are distributed searching, and searching previously harvested metadata. These strategies are presented in the section 5 Standards for Searching Strategies.

Digital libraries, according to Lynch (20), in the forthcoming future, will evolve to the ubiquitous and pervasive deployment of Digital Library technologies and services in the broader information services landscape. In the same direction European guidelines for the ICT Challenge 4 "Digital libraries and content", suggest that Digital Libraries should be driven to services that enhance digital preservation, personalization of data and packaging and repurposing of products.

2. Digital Object

Digital Object is a complex object which is constituted by the actual digital material and the metadata that refers to it. The digital material may be any kind of data in digital format. The metadata that describes the material except for the identifier to the digital object includes other metadata, as well. Kahn and Wilensky (21) presented a framework for distributed digital object services. According to this framework a digital object is a data structure which is constituted by two principal components that are digital material, or data, plus a unique identifier for this material. The identifier is called a handle and can be provided by an authorized handle generator. After digital objects are created, they can be deposited to repositories, from which they can be discovered and retrieved. Payette & Lagoze (22) use a metaphor to define digital object. They present digital object as a cell that has a structural kernel for encapsulating content as opaque byte stream packages and an interface layer giving contextual meaning to the data in the digital object.

---

[3] The Deep Web (also called Deepnet, the invisible Web, dark Web or the hidden Web) refers to World Wide Web content that is not part of the surface Web, which is indexed by standard search engines. (wikipedia)

The Digital Object Identifier (DOI) System (23) is a framework used for identifying content objects in the digital environment. Any digital entity may acquire a DOI name that refers to information, including where the entity (or information about it) can be found on the Internet. Information about a digital object may change over time, including where to find it, but its DOI name will not change. DOI name enables systems to achieve persistent identification and interoperable exchange of managed information on digital networks. Unique identifiers are essential for the management of information in any digital environment.

3. Digital Repository

Digital repository is a collection of resources that are network-accessible without prior knowledge of the structure of the collection. The basic services that a digital repository supports are import, identify, store and retrieve digital objects/assets. The content of repositories may be actual digital assets or the metadata that describe these digital assets. The assets and their metadata do not need to be held in the same repository. Digital repositories are the basis for knowledge discovery and object exchange.

The majority of repositories are designed and implemented according to the principles that IMS Digital Repositories Interoperability (IMS DRI) specification outlines. IMS DRI is a specification intended to provide Best practices and recommendations for the interoperability and interoperation of the basic repository services. These recommendations aim to enable services to present a common interface. The specification is based on schemas already defined (e.g., IMS Meta-Data and Content Packaging), rather than introducing a new one. IMS DRI is described in detail in chapter 3 section 4.3.5.

4. Metadata Standards

The interchange of bibliographic records brought the need for analyzing these records to their basic components and standardizing them. This effort resulted to MARC, the first standard for encoding storage and exchange of bibliographic records that relates information in machine-readable form. Today, the most common encoding language is XML, which arises from simplifying SGML[4], and most of the metadata schemes are presented in XML. All these years, there has been important research conducted to the area of metadata for Digital Libraries. In this section, some of the most important metadata standards are presented.

4.1. MARC

MARC (4) is the acronym for MAchine-Readable Cataloging. It defines a data format that emerged from Library of Congress in 1965. It provides the mechanism by which computers exchange, use and interpret bibliographic information. Its data elements make up the foundation of most library catalogs used today. MARC became USMARC in the 1980s and MARC 21 in the late 1990s.

MARC 21 formats are standards for the representation and communication of bibliographic and related information in machine-readable form. A MARC record involves three elements: the record structure, the content designation, and the data content of the record. The structure of MARC records is an implementation of ISO 2709[5], also known as ANSI/NISO Z39.2. Content designation is the codes and conventions established to identify explicitly and characterize further the data elements within a record and to support the manipulation of this data. The content of most data elements is defined by standards outside the formats, e.g., Anglo-American Cataloguing Rules, Library of Congress Subject Headings, National Library of Medicine Classification. A MARC 21 format is a set of codes and content designators defined for encoding machine-readable records. Formats are defined for five types of data: bibliographic, holdings, authority, classification, and community information. The MARC 21 formats are maintained by the Library of Congress in consultation with various user communities.

Web revolution emerged the need for MARC to adjust to the new internet technologies. In 2002 Library of Congress with the Research Library Group (RLG) and the Online Computer Library Center (OCLC) created MARC XML. The invention of MARC XML aimed to the

---

[4] SGML is an international standard for the definition of device-independent, system-independent methods of representing texts in electronic form. (63)

[5] http://www.iso.org/iso/catalogue_detail.htm?csnumber=7675

conversion of ISO 2709 to XML without any loss of data, so as to make possible the inverse process (i.e. from MARC XML to MARC 21). MARC XML is used as a means of easy sharing of, and networked access to, bibliographic information and aims to the flexibility and scalability of MARC records. This feature enables MARC records to be converted to other metadata schemas, like Dublin Core, or to be expanded to METS.

4.2. Dublin Core Metadata Initiative

The Dublin Core Metadata Element Set (5) is a set of properties and attributes that describe any kind of information. This set of elements enhances the organisation and classification of data regardless of its academic field. The Dublin Core Metadata Element Set consists of 15 properties that organize and semantically describe any kind of data regardless of type and field, that the data belongs. The name "Dublin" is due to its origin at a 1995 invitational workshop in Dublin, Ohio; "core" because its elements are broad and generic, usable for describing a wide range of resources.

The fifteen element of Dublin Core Metadata Element set are part of metadata vocabularies and technical specifications that are provided by the Dublin Core Metadata Initiative (DCMI). The full set of vocabularies, DCMI Metadata Terms (24), also includes sets of resource classes (including the DCMI Type Vocabulary (25)); vocabulary encoding schemes, and syntax encoding schemes. The terms in DCMI vocabularies are intended to be used in combination with terms from other, compatible vocabularies in the context of application profiles and on the basis of the DCMI Abstract Model. Application profiles are schemas which consist of data elements drawn from one or more namespaces (in this case Dublin Core Metadata element Set was one of the namespaces), combined together by implementers, and optimized for a particular local application. The fifteen elements of Dublin Core Element Set are described in table 1 Dublin Core Element Set.

Term Name	**Element Definition**
Contributor	An entity responsible for making contributions to the resource.
Coverage	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.

Creator	An entity primarily responsible for making the resource.
Date	A point or period of time associated with an event in the lifecycle of the resource.
Description	An account of the resource.
Format	The file format, physical medium, or dimensions of the resource.
Identifier	An unambiguous reference to the resource within a given context.
Language	A language of the resource.
Publisher	An entity responsible for making the resource available.
Relation	A related resource.
Rights	Information about rights held in and over the resource.
Source	A related resource from which the described resource is derived.
Subject	The topic of the resource.
Title	A name given to the resource.
Type	The nature or genre of the resource.

Table 1: Dublin Core Element Set

The fifteen elements descriptions have been formally approved in several standards, such as: ISO Standard 15836-2003 of February 2003[6], ANSI/NISO Standard Z39.85-2007 of May 2007[7] and IETF RFC 5013 of August 2007[8] .

[6] http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=52142

[7]http://www.niso.org/kst/reports/standards?step=2&gid=&project_key=9b7bffcd2daeca6198b4ee5a848f9beec2f600e5

[8] www.ietf.org/rfc/rfc5013.txt

In the late nineties when Dublin Core Metadata Element Set was standardized, the evolution of Semantic Web raised the need to assign formal domains and ranges to the resources descriptions. In this case, domains and ranges indicate the kind of described resources and value resources with a given property. Therefore, DCMI, in January 2008, includes formal domains and ranges in the definition of its properties. DCMI trying not to affect the conformance of existing implementations of simple Dublin Core, where domains and ranges have not been specified (dc: namespace at http://purl.org/dc/elements/1.1/), created fifteen new properties with identical names to those of the Dublin Core Metadata Element Set Version 1.1. These new fifteen properties have been created in the dcterms: namespace (http://purl.org/dc/terms/) and have been defined as subproperties of the corresponding properties of DCMES (5). Extensive documentation of the new properties is available in the document "DCMI Metadata Terms", which replaces the first edition in 2006 (24).

4.3. METS

The Metadata Encoding and Transmission Standard (METS) (26) is an XML based standard concerning digital objects. This standard provides descriptive, administrative, and structural metadata intending to the management of digital objects within a repository and the exchange of such objects between repositories, systems and user interaction services. METS was created and designed to provide a relatively easy format and a common data transfer syntax for retaining structural metadata and enabling its sharing during the lifecycle of the digital object. METS was initially introduced in 2001 as an initiative of the Digital Library Federation (DLF), it is supported and maintained by the Library of Congress and received a NISO Registration in 2004, which was renewed in 2006.

The METS schema is constituted by seven major sections, which are the METS Header, Descriptive metadata, Administrative metadata, File section, Structural map, Structural links and Behavior (26). In brief, these sections are described in table 2 METS schema’s sections.

Section’s Name	**Section’s Description**
METS Header	The METS Header contains metadata describing the METS document itself, including such information as creator, editor, etc.
Descriptive	The descriptive metadata section may contain both internally embedded metadata and links to external metadata (e.g a MARC

Metadata	record in a library catalog). Multiple instances of both external and internal descriptive metadata may be included in the descriptive metadata section.
Administrative Metadata	The administrative metadata, respectively to the descriptive may contain either internal or external or both kinds of administrative metadata. This metadata concerns information regarding how the files were created and stored, intellectual property rights, metadata regarding the original source object from which the Digital Library object derives, and information regarding the provenance of the files comprising the Digital Library object (i.e., master/derivative file relationships, and migration/transformation information).
File Section	The file section lists all files containing content which comprise the electronic versions of the digital object. This information along with the structural ensures the integrity of digital object, even when its components are stored in different places.
Structural Map	The structural map is the heart of a METS document. It outlines a hierarchical structure for the Digital Library object, and links the elements of that structure to content files and metadata that pertain to each element.
Structural Links	The Structural Links section of METS allows METS creators to record the existence of hyperlinks between nodes in the hierarchy outlined in the Structural Map. This is of particular value in using METS to archive Websites.
Behavior	A behavior section can be used to associate executable behaviors with content in the METS object. Each behavior within a behavior section has an interface definition element that represents an abstract definition of the set of behaviors represented by a particular behavior section. Each behavior also has a mechanism element which identifies a module of executable code that implements and runs the behaviors defined abstractly by the interface definition.

Table 2 METS schema's sections

In order to ensure flexibility, but also to simplify the interoperability among METS users, a METS profile has been suggested. This profile tries to establish common practice for the profile development of institutions. Institutions are highly recommended to submit their profiles in a formal registration process that makes the profiles visible to others.

5. Standards for searching strategies

Digital Libraries support two main techniques for allowing their resources to be found. The first one is providing a searching interface and the second is exposing their metadata to other Digital libraries or search engines, using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (6). A very interesting field in the Digital Libraries research area is the federated searching service. There are two strategies for federated searching of Digital Libraries, based on the two techniques that were previously mentioned, the distributed searching and searching previously harvested metadata.

Distributed searching is based on a client sending parallel requests to the searching servers of the Digital libraries that belong in the federation. The client gathers the results, duplicates are eliminated or clustered, and the remaining items are presented to the end user. Z39.50 protocol (7) is widely used in library environments for distributed searching. An advantage of this strategy is that the resource intensive tasks of indexing of metadata and storing of the resources are left to the corresponding servers in the federation. The disadvantage of this strategy is that the result of the searching process is dependent to the indexing and ranking capabilities of each Digital Library, making it difficult to combine a result consisting of the most relevant found items.

Searching over previously harvested metadata is based on searching a locally stored index of metadata that has previously been collected from the Digital Libraries in the federation. This strategy requires the creation of a harvesting mechanism, which connects regularly to all the Digital Libraries in the federation and queries their collections for new and updated resources. Following the search, the harvested data is being indexed. OAI-PMH is the most frequent harvesting protocol used. The advantage of this strategy is that the search mechanism performs the indexing and the ranking algorithms locally, having the ability to customize them to its needs offering more consistent and relevant results. The disadvantage of this strategy, opposed to the distributed searching, is that the resource intensive tasks of indexing and storing of metadata are performed locally. In the following sections, the basic protocols for searching a Digital Library are presented.

5.1. Z39.50

Z39.50 (7) is an international protocol for communication between library and information related systems. Z39.50 is playing a significant role for the development and deployment of inter-linked and federated library systems. This standard supports search and retrieve information from different heterogeneous information systems. Z39.50 was initially developed from the library community, in order to satisfy special restrictions that earlier technologies imposed.

The core operations of Z39.50 are initialization, search and retrieve. The core functionality of the standard begins with the negotiation, between client and server, of what services can be provided from the server, continues with the creation and submission of a query and the receiving of the first results and ends with the selection and retrieval of records from the result set. This standard describes nine operation types: Init, Search, Present, Delete, Scan, Sort, Resource-report, Extended-services, and Duplicate Detection.

Z39.50 services are carried out by message exchange between the client and server. The messages are divided into two categories, requests and responses. Services are defined to be confirmed, non-confirmed, or conditionally-confirmed. A confirmed service is defined in terms of a request (from the client or server) followed by a response (from the peer). A non-confirmed service is defined in terms of a request from the client or server, with no corresponding response. A conditionally-confirmed service is a service that may be invoked as either a confirmed or nonconfirmed service (7).

Z39.50 can be incorporated into all sorts of products and systems only a few of which are currently being exploited. Z39.50 can be implemented on any computer system and so opens the way for true “interworking”. Thus a Mac Z-client can access a UNIX and a Windows NT based system simultaneously and seamlessly.

It supports a number of actions, including search, retrieval, sort, and browse. Searches are expressed using attributes, typically from the bib-1 attribute set, which defines six attributes to be used in searches of information on the server computer: use, relation, position, structure, truncation, completeness. The syntax of the Z39.50 protocol allows for very complex queries.

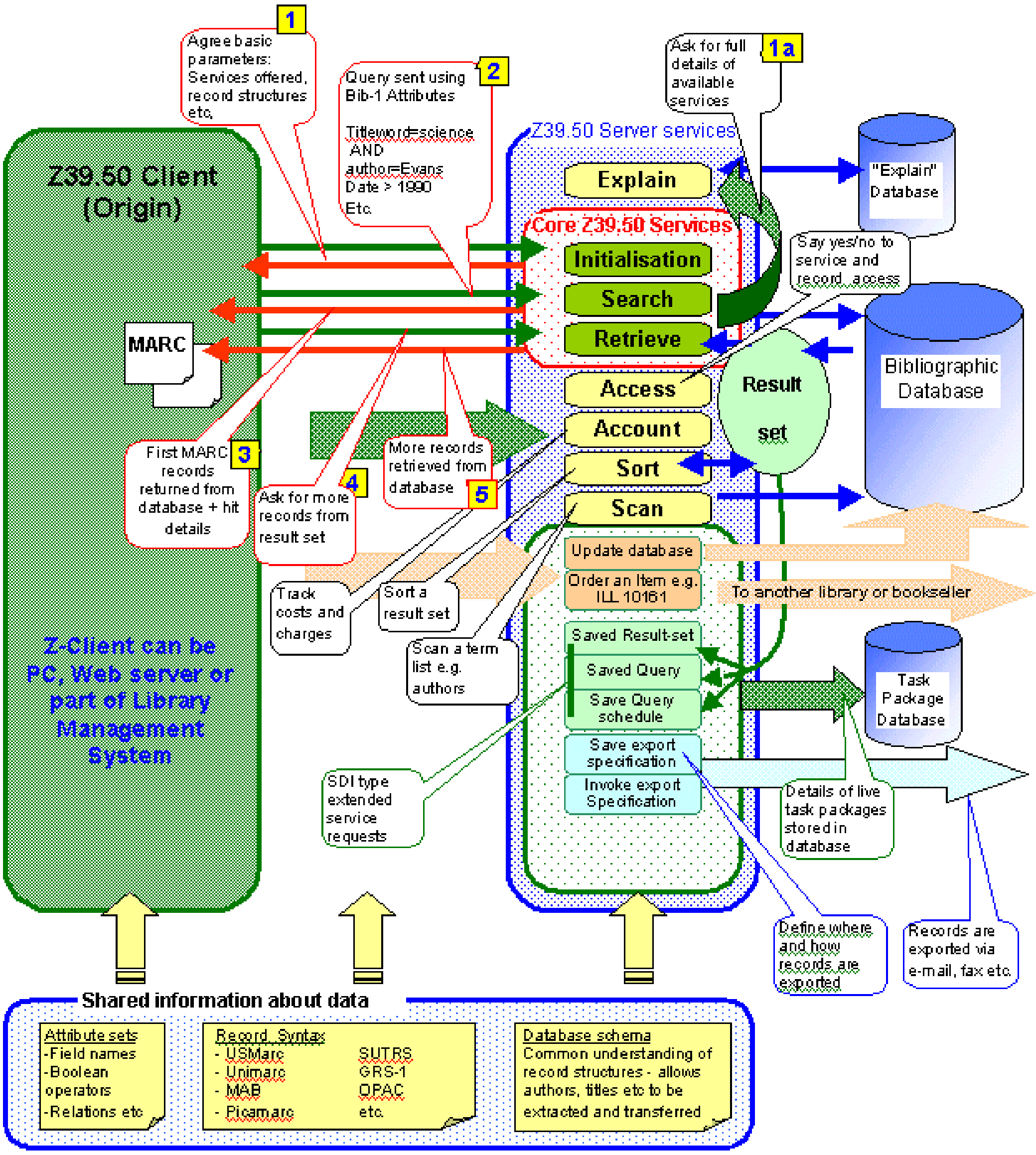

Figure 1: A graphical representation of Z39.50 (27)

5.2. SRW / SRU

Search and Retrieve Web/URL Service (28) (29) are Web services for search and retrieval based on Z39.50. SRW and SRU are intended to solve the lack of a unified format and structure of search queries and responses. SRW and SRU propose syntax for queries and responses and conduce to the creation of searching mechanisms that give efficient results even in deep web resources. SRW/U allow people and HTTP user agents to query Digital

Libraries more efficiently without the need of more expensive and complicated meta-search protocols.

SRW/U utilize CQL (Contextual Query Language), a standard syntax for representing queries. (29) SRW/U are Web services-based protocol for querying Internet indexes or databases and returning search results. SRU is based on REST protocol and SRW is based on SOAP. Even though SRW/U use different protocols, perform the same tasks and they both define a similar set of commands (known as "operations") and responses. SRW/U support three operations: explain, scan, and searchRetrieve. Each operation is qualified with one or more additional name/value pairs. Explain operations are requests sent by clients as a way of learning about the server's database/index as well as its functionality. At a minimum, responses to explain operations return the location of the database, a description of what the database contains, and what features of the protocol the server supports. Scan operations list and enumerate the terms found in the remote database's index. Clients send scan requests and servers return lists of terms. SearchRetrieve operations are the heart of the protocol. They provide the means to query the remote database and return search results (28).

5.3. OAI – PMH

Open Archives Initiative (3) develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (6) is a low barrier mechanism for repository interoperability. Data providers are repositories that expose structured metadata via OAI-PMH. Service Providers then make OAI-PMH service requests to harvest that metadata. OAI-PMH is a set of six verbs or services that are invoked within HTTP.

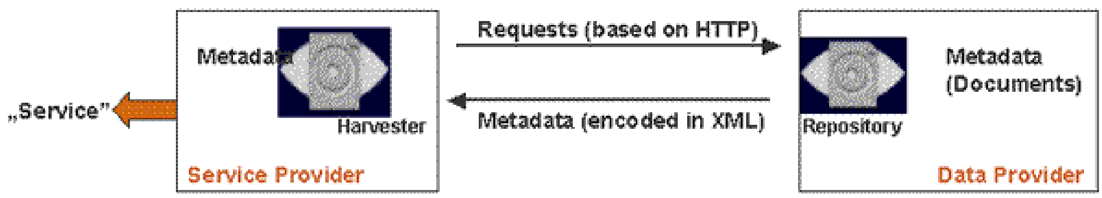

Figure 2: Basic Functioning of OAI-PMH (3)

The OAI-Protocol for Metadata Harvesting (OAI-PMH) is a harvesting protocol applied on repositories regarding records containing metadata. This mechanism enables data providers to make their metadata available to services, based on the open standards HTTP and XML. The metadata that is harvested may be in any format that is agreed by the contracting parties, meaning the data and service providers, although unqualified Dublin Core is specified to provide a basic level of interoperability. Dublin Core is the resource discovery lingua franca for metadata. Since all DC fields are optional and repeatable, most repositories should have no trouble creating at least a minimal mapping of their native metadata to unqualified DC. Repositories are not required to store their metadata in Dublin Core.

Using OAI-PMH, metadata from many sources can be gathered together in one database, and services can be provided based on this centrally harvested data. The harvested data can be indexed with the corresponding tools and used for searching. The link between this metadata and digital assets is not defined by the OAI protocol. Selective harvesting allows harvesters to limit harvest requests to portions of the metadata available from a repository. The OAI-PMH supports selective harvesting with two types of harvesting criteria that may be combined in an OAI-PMH request: datestamps and set membership. OAI-PMH defines a set of six request types (known as "verbs"): Identify, ListMetadataformats, ListSets, ListIdentifiers, ListRecords, GetRecord (6). OAI-PMH enables the surfacing of deep web content and offer low cost interoperability.

6. Summary

In this chapter were presented the basic definitions, standards and issues related to Digital Libraries. There were presented the basic metadata schemes that are used by Digital Library's community in order to describe the digital resources. Finally, there were presented the basic searching mechanisms for Digital Libraries and the protocols used for searching the deep web resources.

Chapter Three: eLearning

In this chapter an overview of eLearning is provided. The organizations that perform research in the area of eLearning are introduced, by presenting the most important of their standards. The IEEE Learning Object Metadata, SCORM standard and IMS Digital Repository Interoperability specification are presented thoroughly, since they are going to be referenced in the following chapters.

1. eLearning

ELearning is a concept that arose since the beginning of the Web. Firstly, it was introduced as distant learning, where education was provided through learning resources such as articles, tutorials and supplementary media. The main characteristic of distant learning was that teacher and learner worked asynchronously, in different places and time. Following this, the term distant learning was substituted by others like “online learning”, “internet learning”, “distributed learning”, “virtual learning” and “eLearning”, where the latter seems to prevail. ELearning is a term used for learning that is enhanced by the use of digital tools and resources. It refers to the delivery of educational content through electronic means, implying especially the web but it can also include CD-ROMs, audio and video tapes, interactive TV and satellite transmissions. ELearning may involve online interaction between the learner and the teacher.

Increasingly, universities and educational institutions adopt eLearning as the main delivery method for educational content and learning experience. Though, eLearning is not used only as an academic service but is one of the main methods that organizations and companies use for the training of their personnel. Learning Management System (LMS) is a software application that automates the administration, documentation, tracking, and reporting of training events (30) and is the main system architecture for supporting the learning process. LMS are based on the idea of replicating the concept of the classroom in the context of Internet. The most representative LMS are ATutor[9], Blackboard[10], Moodle[11] and Sakai[12].

---

[9] http://www.atutor.ca/

[10] http://www.blackboard.com/

2. Learning Object

Advances in internet and web technologies have given a new shape to the eLearning mechanisms. Learning objects, as the main core of eLearning technology, have attracted the attention of academics as 'atomic' instructional entities. Universities and organizations need a generic framework for assuring the interoperability of their instructional entities. Trying to satisfy this need several organizations, such as the Learning Technology Standards Committee (LTSC) of the Institute of Electrical and Electronics Engineers (IEEE), the Alliance of Remote Instructional Authoring and Distribution Networks for Europe (ARIADNE), Advanced Distributed Learning (ADL) Initiative, the Aviation Industry CBT (Computer-Based Training) Committee, PROmoting Multimedia Access to Education and Training in EUropean Society, Dublin Core and the Instructional Management Systems (IMS) Project, developed technical standards to support the broad deployment of learning objects.

The term “Learning Object” was firstly launched by Wayne Hodgins in 1994 in the title of the CedMA working group called “Learning Architectures, API’s, and Learning Objects”. Later, the Learning Technology Standards Committee (LTSC) used the term “Learning Object” and defined it as “Any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning. Examples of technology-supported learning include computer-based training systems, interactive learning environments, intelligent computer-aided instruction systems, distance learning systems, and collaborative learning environments. Examples of Learning Objects include multimedia content, instructional content, learning objectives, instructional software and software tools, and persons, organizations, or events referenced during technology supported learning” (8). However, this definition was considered too broad, because it could not exclude any person, thing or idea since these could be “referenced during technology supported learning”. Consequently, several attempts have been made for narrowing this definition, making communication confusing and difficult. Wiley (31) defines a learning object as “any digital resource that can be reused to support learning”. Wiley proposed this supplementary definition based on the belief that it sufficiently narrows the LTSC’ definition to a reasonably homogeneous set of things: reusable digital resources and that these resources comply with the rule of

---

[11] http://moodle.org/

[12] http://sakaiproject.org/portal

compatibility and reusability as proposed by LTSC. ORMEE (32)(Observatory on Rights Management for eLearning in Europe) is a financed project in the framework of the European Commission eLearning initiative and suggests that “Learning Objects can be defined as small instructional components that can be reused a number of times in different learning contexts” based on an object oriented approach. Another definition that is worth quoting is proposed by a computer-based training (CBT) vendor NETg, Inc., which uses the term “NETg learning object” and applies a three-part definition: a learning objective, a unit of instruction that teaches the objective and a unit of assessment that measures the objective (33).

Koper (34) defines learning object as “any digital, reproducible and addressable resource used to perform learning activities or learning support activities, made available for others to use”. Koper, trying to provide a more restricted definition than IEEE LTSC and Wiley, highlights that when a learning object is aggregated to a learning activity, the aggregate is no longer a learning object, but it is now a 'unit of learning'. A common definition, as Rehak and Mason (35) suggest, is "A small chunk of learning which serves a learning objective". On the other hand, there is also the approach that learning objects do not seem to be distinguished from digital resources (36). Thus, the issue of what constitutes a learning object is very much open to debate.

The main principle behind the concept of learning object is that content designers can construct small instructional components that can be reused in different learning contexts. Additionally, the majority of learning objects that are digital entities deliverable over the Internet, enable those who use them to collaborate on and benefit immediately from new versions. According to Wiley (31) Learning Objects are characterized by several qualities, whose rate of adoption differentiate one object to another. Wiley proposes five learning object types based on the qualities taxonomy. These types are described in Table 3 Learning Object Types (31).

Types names	**Types Definitions**
Fundamental	A fundamental learning object is a digital resource that is not combined with any other and serves an individual purpose

Combined-closed	A Combined-closed learning object is a small number of digital resources combined at design time by the learning object's creator, whose constituent learning objects are not individually accessible for reuse (recoverable) from the Combined-closed learning object itself. An example of a combined-close learning object is a video clip
Combined-open	A Combined-open learning object is a larger number of digital resources combined by a computer in real-time when a request for the object is made, whose constituent learning objects are directly accessible for reuse (recoverable) from the Combined-open object
Generative-presentation	Logic and structure for combining or generating and combining lower-level learning objects (Fundamental and Combined-closed types). Generative-presentation learning objects can either draw on network-accessible objects and combine them, or generate (e.g., draw) objects and combine them to create presentations for use in reference, instruction, practice, and testing
Generative-instructional	Logic and structure for combining learning objects (Fundamental, Combined-closed types, and Generative-presentation) and evaluating student interactions with those combinations, created to support the instantiation of abstract instructional strategies (such as "remember and perform a series of steps").

Table 3: Learning Objects' types (31)

Most researchers (37), (38), (39) in the field of eLearning suggest that the smaller and more specific the learning objects are, the more reusable they can be found in new instructional contexts.

3. Learning Object Repository

Learning Object Repositories (LORs) are digital repositories, as defined in the previous chapter in the section 3 Digital Repository, whose content is learning objects. LORs are the basic constituent for knowledge discovery and learning object exchange. The need for easy-accessible and exchangeable Learning Objects leads to the use of systems that organize, index, search and deliver its content that are Learning Objects Repositories.

Through these years, several organizations have created Learning Object Repositories like Merlot, SMETE and CAREO. The basic disadvantage characterizing most of the efforts is that each repository tries to satisfy its organization's needs, uninterested in interoperability. This finding resembles the field of Digital Libraries. The eLearning community has seen fruitful initiatives in the standardization of learning field. An appropriate effort has been launched by IMS GLC, through the IMS Digital Library Interoperability standard. This standard was initially launched for the eLearning field, but it has gained popularity in the digital community, as well, since it is based on schemas already defined (e.g., IMS Meta-Data and Content Packaging), rather than introducing a new one. IMS DRI is being presented thoroughly in the section 4.3.5.

4. Standards and Specifications

Standards are essential for the reusability and interoperability of resources either digital or learning. Without standardization users, such as content assemblers, would face severe difficulties with reusing resources in their content aggregations (learning objects or courses), sharing these resources or even use their own resources in different learning environments. Standards in eLearning focus on three main directions, metadata, creation and deployment of resources and learning interactions. Standards for metadata denote description based on specific properties, depending on the standard, of the resource. These standards enhance interoperability of resources across different platforms and allow efficient results in resource discovery. A number of organizations are currently developing international standards for metadata such as IEEE LOM (8), Dublin Core (2), IMS metadata etc. The second approach related to learning objects creation and deployment, includes standards that provide guidelines to unified and sufficient leaning object creation and standards that enhance the collaboration between learning environments and distributed repositories. This approach aims to the technical interoperability of resources and the efficient communication between learning objects and learning environment. Some of these standards are IEEE Content Package (40), SCORM (11), AICC Computer Managed Instruction (CMI) Guidelines for Interoperability (9), IMS Digital Repository Interoperability (10) and others. Finally, as far as it concerns the Learning Interaction approach, there have been a serious number of standards presented, like IMS Learning Design (41), Simple Sequencing (42), and Question & Test Interoperability (43). These standards aim to enhance instructional view of eLearning process and support the broadest range of pedagogic approaches. In this subsection, the main organizations that deal with eLearning standards are presented, along with the most important of their standards.

4.1. IEEE Learning Technology Standards Committee (LTSC)

IEEE Learning Technology Standards Committee (LTSC) is an organization that was founded in order to develop and promote standards, guidelines and best practices for Learning. IEEE LTSC participates in consortia and cooperates with other international organizations on developing widely accepted specifications and standards. Since 1996 IEEE LTSC, when it was formed, has published several widely accepted and adopted technology standards. These standards are described in the following subsections. There is a more thorough description of IEEE Learning Object Metadata Standard, since it is one of the most important and it is used in this thesis.

4.1.1 IEEE Standard for Learning Technology - Learning Technology Systems Architecture (LTSA)

Learning Technology System Architecture (LTSA) (44) is a standard suggesting a high-level architecture for designing learning-oriented systems, such as Learning Management Systems, Computer-based Training systems etc. The LTSA specification does not provide any pedagogical, content or platform prerequisites. According to the editors of the standards, it aims to accomplish three main goals, firstly to provide a framework for understanding existing and future systems, secondly to promote interoperability and portability by identifying critical system interfaces and finally incorporates a technical horizon (applicability) of a least 5-10 years while remaining adaptable to new technologies and learning technology systems.

4.1.2 IEEE Standard for Learning Technology - Data Model for Content to Learning Management System Communication

Data Model for Content to Learning Management System Communication standard (45) presents a data model that enables Learning Objects to interact with the runtime service of a Learning Management Systems. The standard does not specify either the communication protocols or the behavior that the two parts adopt during their communication. It is based on the data model presented in "Computer Managed Instruction (CMI) Guidelines for Interoperability" (9). The standard adopts the data model of CMI with some modifications in order to balance the need to support existing implementations and to make technical corrections. The standard aims to build consensus around, resolve ambiguities, and correct defects in the CMI data model for the interaction between learning objects and a runtime service used by Learning Management Systems.

4.1.3 IEEE Standard for Learning Technology - ECMAScript Application Programming Interface for Content to Runtime Services Communication

ECMAScript Application Programming Interface for Content to Runtime Services Communication (46) is an API based on the corresponding API defined in "Computer Managed Instruction (CMI) Guidelines for Interoperability" (9). The API services in the ECMAScript language enable the communication between learning objects and the runtime service used by learning management systems. The standard does not deal with the data structure, data security or communication between the runtime service and the learning management system. The standard aims to build consensus around, resolve ambiguities, and correct defects in the CMI ECMAScript API for the communication between learning objects and a runtime service used by Learning Management Systems.

4.1.4 IEEE Standard for Learning Object Metadata

Learning Object Metadata (8) is a standard that describes “learning objects”. The IEEE LOM is the first part of a multipart standard for learning objects, focusing on a conceptual data model referring to the structure of a Learning Object. This data model deals with which properties of the learning object should be described, using the appropriate vocabulary. These properties may be grouped in nine categories. These categories are general, lifecycle, meta-metadata, educational, technical, rights, relation, annotation, and classification. The purpose of this standard is to facilitate search, evaluation, acquisition, and use of learning objects, for instance by learners or instructors or automated software processes. It also aims to enhance sharing and exchange of learning objects, by enabling the development of catalogs and inventories while taking into account the diversity of cultural and lingual contexts in which the learning objects and their metadata are reused.

The LOM comprises a hierarchy of elements where the first level are the nine categories, which were referenced above and 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. The data model also specifies the value space and data type 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. None of the elements of LOM are obligatory in instantiating the LOM scheme. Briefly, the nine categories, as described in the standard, are presented in the Table 4 Categories of LOM Elements.

Category	**Description**
General	Groups the general information that describes the learning object as a whole.
Lifecycle	Groups the features related to the history and current state of this learning object and those who have affected this learning object during its evolution.
Meta-Metadata	Groups information about the metadata instance itself (rather than the learning object that the metadata instance describes).
Technical	Groups the technical requirements and technical characteristics of the learning object.
Educational	Groups the educational and pedagogic characteristics of the learning object.
Rights	Groups the intellectual property rights and conditions of use for the learning object.
Relation	Groups features that define the relationship between the learning object and other related learning objects.
Annotation	Provides comments on the educational use of the learning object and provides information on when and by whom the comments were created.
Classification	Describes this learning object in relation to a particular classification system.

Table 4: Categories of LOM elements

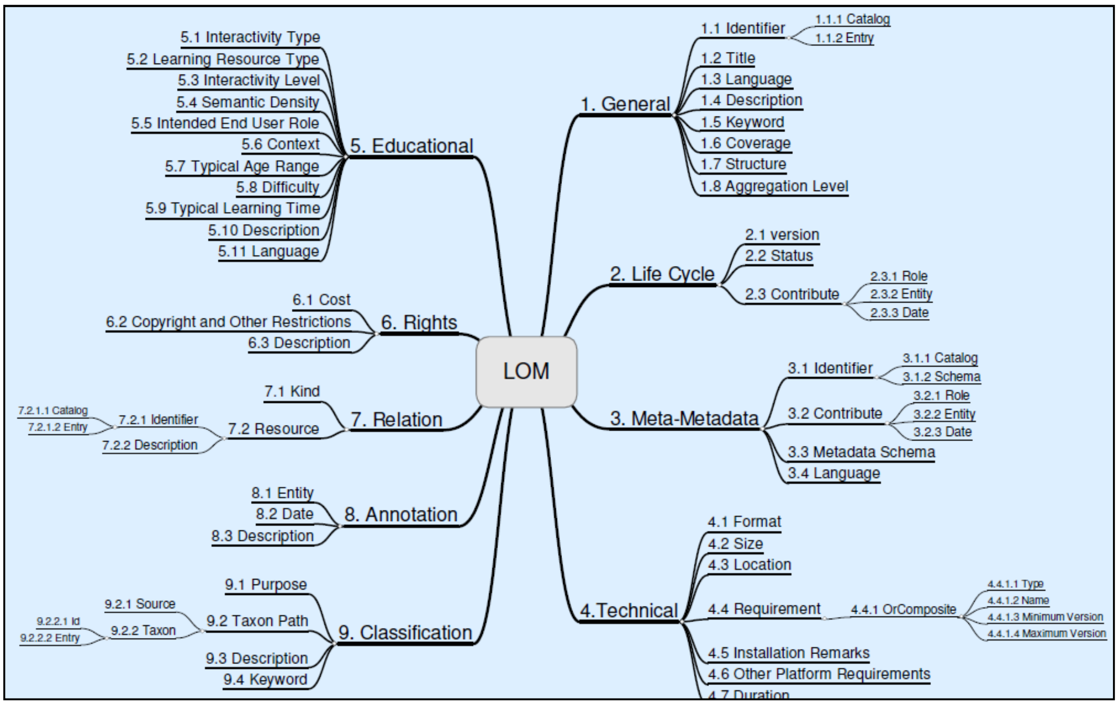

Figure 3: A schematic representation of the hierarchy of elements in the LOM data model (47)

LOM can be used in an application profile where users can specify which elements and vocabularies suit best in their needs. In an application profile elements from the LOM may be excluded and elements from other metadata schemas (i.e. Dublin Core) may be included. Equally, the vocabularies in the LOM may be supplemented with values appropriate to the users' needs (47).

4.1.5 IEEE Standard for Learning Technology-Extensible Markup Language (XML) Schema Definition Language Binding for Learning Object Metadata

Extensible Markup Language (XML) Schema Definition Language Binding for Learning Object Metadata (48) is a standard that provides an XML Schema definition language binding of the learning object metadata (LOM) data model defined in the previous section. An implementation that conforms to this standard shall conform to the IEEE LOM, as well. The purpose of this standard is to allow the creation of LOM instances in XML. This standard enhances interoperability among exchange of LOM instances between various systems.

4.2. Aviation Industry CBT (Computer-Based Training) Committee

The Aviation Industry CBT (Computer-Based Training) Committee (AICC) is an international committee, founded in 1988, in order to provide hardware standardization of Computer-Based Training delivery platforms and since then develops guidelines for CBT to professionals. The AICC's guidelines focus on the development, delivery and evaluation of technology- based training for aviation industry. The main goals of AICC are to support the implementation of CBT media through guidelines and to enhance interoperability of these guidelines. Some of the most important guidelines concerning CBT are described above.

4.2.1. CMI Guidelines for Interoperability

The CMI Guidelines for Interoperability (9) defines a number of Computer Managed Instruction (CMI) principles and terms in order to enhance interoperability among CMI systems. These guidelines enable a CMI system to manage any Computer-Based Training course. Especially, they refer to the interaction between a course and a CMI, the ability of moving a course between CMI systems and the storage of evaluation data of a course. The first version of these guidelines was released in 1993 and it was LAN-Based. These guidelines were revised in 1998, when they were updated in order to enable CMI integration with web technology and in 1999 when, they included a Javascript API and finally in 2001, when the main purpose of the revision was to make the AICC API match the ADL SCORM version 1.1.

4.2.2. CBT Package Exchange Notification Services (PENS)

The CBT Package Exchange Notification Services (PENS) (49) guidelines were designed in order to simplify the content package exchange between learning management systems. These guidelines allow Authoring and Content Management systems to integrate publishing of content with LMSs. A notification service is suggested and announces the location of the content packages available for transfer. The intent is to automate the notification, transfer and delivery confirmation of content packages between authoring tools and learning management systems.

4.3. Instructional Management Systems (IMS) Global Learning Consortium

Instructional Management System (IMS) started as a project in 1995 within the National Learning Infrastructure Initiative of EDUCAUSE. In the beginning, the IMS Project focused on specifications for higher education with effect on K-12 schools, corporate and government training, as well. Over time, it was renamed to IMS Global Learning Consortium and concerned with the interoperability for learning systems and learning content and the enterprise integration of these capabilities. Now, IMS GLC is a non-profit association including the world's leading educational technology representatives, content providers, academia and government organizations focusing on improving technology-based learning.

IMS collaborates with other organizations on providing standards, innovation, best practice and recognition of superior learning impact and guidelines. IMS specifications refer to both online and off-line processes, taking place in synchronous or asynchronous learning. IMS GLC is now enabling the next generation of Digital Learning Services, combining digital content, assessment, applications, and administrative services. The main goal of IMS standards is to improve the educational experience and educational attainment. The most representative standards, suggested by IMS, are presented in the following sections.

4.3.1. IMS GLC Content Packaging

The IMS GLC Content Packaging Specification (40) defines a data format for the exchange of learning objects among digital repositories and learning management systems. The goal of this specification is to enhance delivery, reusability and sharing of learning material. The specification enables users to import, export, aggregate, and disaggregate IMS Packages, retaining information describing the media in the IMS Package, and how it is structured, such as a table of contents or which web page to show first. The IMS Content Packaging specification focuses on the packaging and transport of resources, but doesn't determine the nature of those resources. This is because the specification allows adopters to gather, structure and aggregate content in an unlimited variety of formats.

The content of an IMS Content Package may be common web pages, documents and images or specialized objects, like Java Applets, IMS Question and Test Interoperability items or IMS Learner Information Packaging fragments. The basic component of an IMS Content Package is the manifest file. This XML-based file describes the logical package and the relationships among all of its components. Manifest contains a list of all the files included in the interchange package and all references to resources that are located elsewhere. It may also contain specialized structural content in the resources section of the manifest document itself. Finally, the organization section of the manifests arranges all the content package’s components constructing an educational object.

4.3.2. IMS Question and Test Interoperability

IMS Question and Test Interoperability Specification (43) is designed in order to describe a data model for the representation of question and test data and their corresponding results reports. The specification supports the interoperability and interaction of an item (question or test) among authoring tools, item banks, learning management systems and assessment delivery systems. The IMS QTI specification enhances interoperability and extendibility utilizing well-defined innovative extensions.

4.3.3. IMS Simple Sequencing

The IMS Simple Sequencing Specification (42) is designed in order to provide a method for describing paths of an authored learning experience, based on the intended behavior, so as to enable any learning management system (LMS) to consistently sequence discrete learning. Utilizing this specification a learning content designer can specify the sequence according which elements of electronic content are going to be presented to the learner and the conditions under which a piece of content is selected, delivered, or skipped during presentation.

4.3.4. IMS Learning Design

The Learning Design Specification (41) provides a generic and flexible language in order to utilize pedagogical methods in eLearning. This language is designed to enable different pedagogies to be expressed. The approach has the advantage over alternatives in that only one set of learning design and runtime tools need to be implemented in order to support the desired wide range of pedagogies. The language was originally developed at the Open University of the Netherlands (OUNL).

4.3.5. IMS Digital Repository Interoperability

The IMS Digital Repository Interoperability Group provided a functional architecture and reference model for repository interoperability, the IMS Digital Repository Interoperability specification (10), for the resolution of interoperability issues among repositories. The scope of this specification was to be applied in a wide range of repositories, thus the specification focuses basically on recommendations to certain level and avoids recommendations in the resolution of operational and implementing issues.

Five basic functions defined by IMS DRI are: search/expose, gather/expose, submit/store, request/deliver, and alert/expose. For the *search* function, the specification recommends using either XQuery (www.w3c.org/XML/Query) with SOAP protocol or Z39.50. For the *gather* function, the OAI's harvesting protocol is recommended. No recommendation is made for the other three functions in the current version of the specification.

Although the IMS DRI was designed to deal with heterogeneous repository, it does not explicitly faces it, since it just provides recommendation in a certain level. So, it is up to the implementers to ensure format compatibility. The DRI Group suggest to the repository implementers to provide "search intermediaries", which are going to support multiple formats.

4.4. Advanced Distributed Learning (ADL) Initiative

The Advanced Distributed Learning (ADL) Initiative (50) realizes a strategy developed by the US Department of Defense (DoD) as response to the need for on-demand training. ADL was founded in 1997 in order to serve three main purposes, these purposes are:

- To identify and recommend standards for training software and associated services purchased by Federal agencies and contractors.
- To facilitate and accelerate the development of key technical training standards in industry and in standards-development organizations.
- To establish guidelines on the use of standards and provide a mechanism to assist DoD and other Federal agencies in the large-scale development, implementation, and assessment of interoperable and reusable learning systems.

In order to fulfill efficiently these purposes ADL collaborates with government, industry and academia. Nowadays, the main goal of ADL becomes the need to promote international specifications and standards for designing and delivering learning content. In this field, ADL collaborates with four organizations that are Aviation Industry CBT Committee (AICC), Alliance for Remote Instructional Authoring and Distribution Networks for Europe (ADRIANE), Institute of Electrical and Electronics Engineers (IEEE) Learning Technology Standards Committee (LTSC) and IMS Global Learning Consortium, Inc. (IMS GLC). The center of interest in this collaboration is the Sharable Content Object Reference Model (SCORM) (11), which is a standard introduced by ADL in 1999.

4.4.1. Sharable Content Object Reference Model (SCORM)

The Sharable Content Object Reference Model (SCORM) (11) is a collection of technical standards, specifications, and guidelines related to online learning and communication protocols, aiming to the accessibility, interoperability, durability and reusability of content and systems. SCORM is built upon the work of the Aviation Industry Computer-Based Training (CBT) Committee (AICC), the IMS Global Learning Consortium, Inc., the Institute of Electrical and Electronics Engineers (IEEE), the Alliance for Remote Instructional Authoring and Distribution Networks for Europe (ARIADNE), and others to create one unified "reference model" that meet US DoD high-level requirements for Web-based learning content and systems. Some of these standards are IEEE Learning Object Metadata Standard (8), IEEE ECMAScript Application Programming Interface for Content to Runtime Services Communication standard (46), IEEE Standard for Extensible Markup Language (XML) Schema Binding for Learning Object Metadata (48), IEEE Standard for Learning Technology – Data Model for Content Object Communication (45), IMS Content Packaging Specification (40), IMS Simple Sequencing Behavior and Information Model (42), Aviation Industry CBT Committee (AICC) Computer Managed Instruction Guidelines for Interoperability (9) .The latest release of SCORM is SCORM 2004 4th Edition Version 1.1 (11), which was announced from ADL on 14th August 2009.

The SCORM 4th Edition Version 1.1 is constituted by three books, the Content Aggregation Model (CAM), the Run-Time Environment (RTE) and the Sequencing and Navigation (SN) (11). While these three books are intended to stand alone, there are areas where they relate with common issues and their content overlaps .The SCORM Content Aggregation Model (CAM) book (51) describes the components used in a learning experience, how to package those components for exchange from system to system, how to describe those components to enable search and discovery and how to define sequencing information for the components. The scope of SCORM CAM is to support consistent storage, labeling, packaging, exchange and discovery of learning content. The SCORM Run-Time Environment book (52) describes the learning management system (LMS) requirements in managing the run-time environment (i.e., content launch process, standardized communication between content and LMSs and standardized data model elements used for passing information relevant to the learner’s experience with the content). The RTE book also covers the requirements of Sharable Content Objects (SCOs) and their use of a common application programming interface (API) and the SCORM Run-Time Environment Data Model. The SCORM Sequencing and Navigation book (53) describes how SCORM-compliant content may be delivered to learners through a set of learner- or system-initiated navigation events. The branching and flow of that content may be described by a predefined set of activities. It covers the essential learning management system (LMS) responsibilities for sequencing content objects (Sharable Content Objects (SCOs) or assets) during run-time and allowing those SCOs to indicate navigation requests. The SN book also offers guidance for providing navigation controls to learners.

The SCORM components used to make a learning experience are assets, sharable content objects (SCOs), activities, a content organization and content aggregations (51) and they are briefly described in Table 5 SCORM components.

SCORM Component	**Description**
Asset	The basic building block of a learning resource. It can be any electronic representation of media.
Sharable Content Object (SCO)	A collection of one or more assets that represent a single launchable learning resource that uses the SCORM RTE to communicate with an LMS.
Learning activity	A meaningful unit of instruction, it is conceptually something the learner does while progressing through instruction. A learning activity may provide a learning resource (SCO or asset) to the learner or it may be composed of several sub-activities.
Content Organization	A representation or map that defines the intended use of the content through structured units of instruction (activities). The map shows how activities relate to one another.
Content Aggregation	It can be used as both an action and as a way of describing a conceptual entity. Content aggregation can be used to describe the action or process of composing a set of functionally related content objects so that the set can be applied in a learning experience.

Table 5: SCORM components

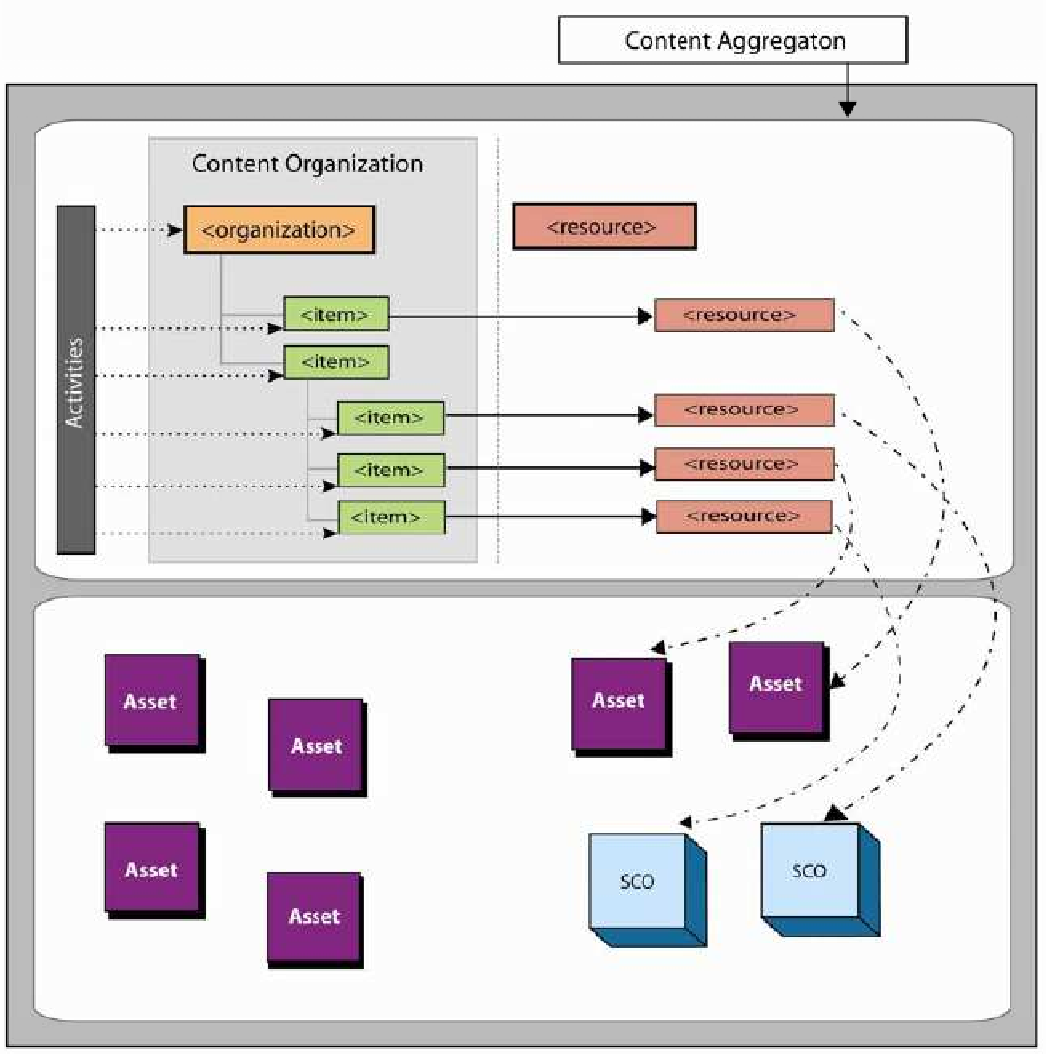

Figure 4: Conceptual Illustration of a Content Aggregation (51)

SCORM Content Packaging is a set of specific requirements and guidance, or application profiles, of the IMS Content Packaging Specification. SCORM Content Packages adheres strictly to the IMS Content Packaging Specification and provides additional explicit requirements and implementation guidance for packaging assets, SCOs and content organization. A content package contains two major components:

- A special XML document describing the content structure and associated resources of the package called the manifest file (imsmanifest.xml). A manifest is required to be present at the root of the content package.
- The content (i.e., physical files) making up the content package.

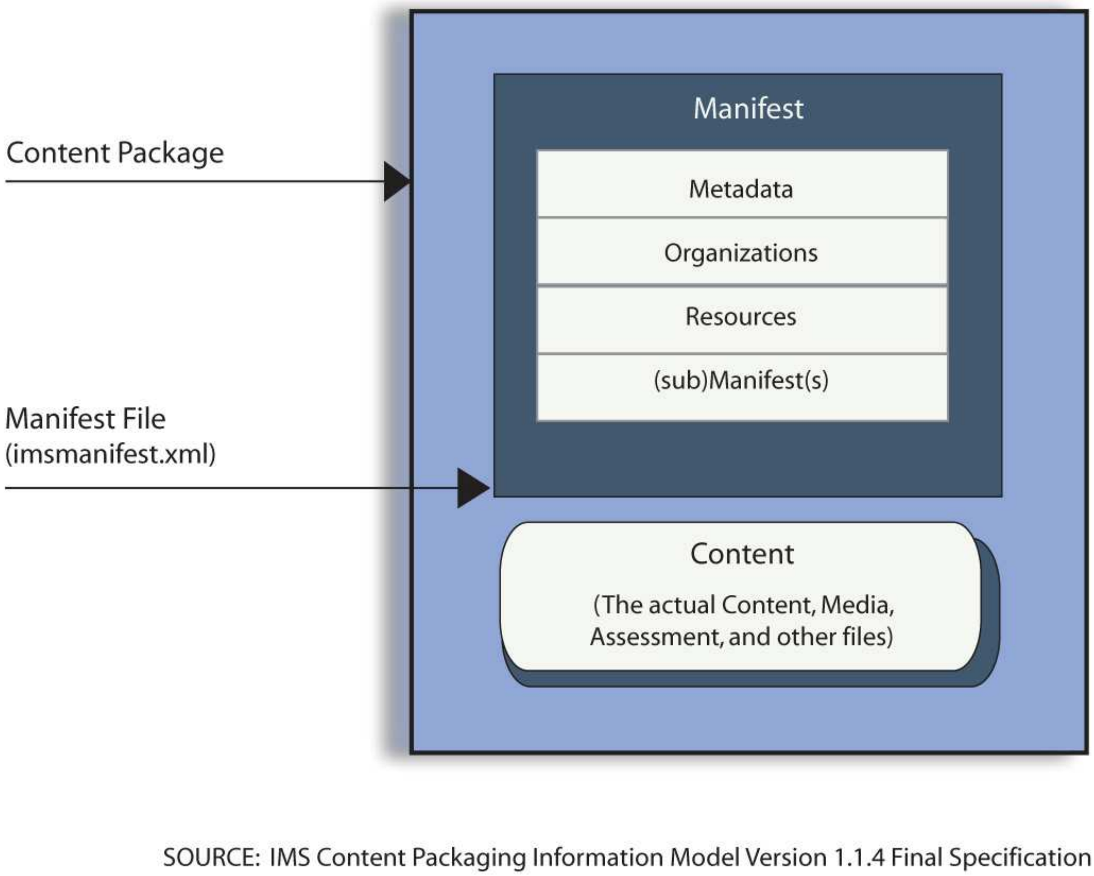

Figure 5: Content Package Conceptual Diagram (51)

5. Summary

In the beginning of this chapter, the basic definitions related to eLearning were presented. In the following sections, the author presents the international organizations that are dealing with eLearning and the most important of their suggested standards. The organizations presented are IEEE Learning Technology Standard Committee (IEEE LTSC), Aviation Industry Computer Based Training Committee (AICC), IMS Global Learning Consortium (IMS GLC) and Advanced Distributed Learning (ADL) Initiative.

Chapter Four: Bridging the two Worlds

Regarding difficulties in the association between eLearning and Digital Libraries, the interoperability problem seems to prevail. The Interoperability problem between Digital Libraries and Learning Object Repositories is the most important, thus it requires immediate actions. First of all, the fact that the learning and digital resources are located in a large number of distributed and heterogeneous repositories raises the need for universal access, whilst the wide diversity of Metadata Models encumbers further progress in defining a universal schema. Moreover, localization and reusability of digital and learning objects is a major issue as it accommodates faster deployment of new systems using already known components and enables new resources' discovery.

Apart from the technical difficulties concerning the interoperability issue there are also some semantic issues. It is not efficiently specified how educational information is represented in metadata fields. The lack of expertise in writing metadata is a serious problem in educational metadata, since its quality and content depend on the instructor's view. The management and quality assurance of learning resources, is a complicated procedure that has until now been carried out by librarians. In the case of learning object and digital repositories interoperability, arises the demand for educators to acquire some of the skills that are currently the preserve of the librarians (37). Xavier Ochoa and Erik Duval (53) conducted research in the area of metadata quality control. They propose metadata quality metrics to implement an automatic evaluator of metadata that can flag low-quality instances. Exploiting these metrics frameworks that perform federated searches could choose to return only resources that have an efficient description. Hughes (54) provides a metadata quality assessment scheme within a specialized Open Archives Initiative (OAI) sub-domain, the Open Language Archives Community (OLAC), based on quality metrics.

1. Defining the Interoperability problem

The interoperability problem could be decomposed to less complex sub-problems whose addressing facilitates the delivery of an overall solution. The first sub-problem is the definition of a uniform approach towards providing in the common repository functions. One well-established approach is IMS-DRI specification provided by IMS Global Learning Consortium which provides recommendations for the interoperation of the most common repository functions while these recommendations should be implementable across services to enable a common interface.

Another important sub-problem is the selection of the model that will store learning and digital object's metadata. The diversity of the available models, underlines the incompatibility issues that needs to be addressed. The main barrier that has to be overcome is the fact that each model was designed to meet specific requirements and effectively solve a particular problem. Many of the models overlap with each other but no one can be considered as the most generic without making any concessions. Thus, there are two possible solutions. Either adopt the most common and suitable model and try to overcome any imposed constraints, or create a more generic unified model that will meet our exact requirements.

Another important issue is the repository's connectivity with external world (source, agents, other repositories etc.). The diversity of connectivity options is summarized in three categories: Incoming, Outgoing and Bidirectional. The incoming connectivity is the ability to search and retrieve learning objects from different (maybe heterogeneous) repositories. This could also be considered as a centralized repository that offers the ability to search through it other repositories. The outgoing connectivity is the ability to provide learning objects to external repositories / applications in a specific format. This could also be considered as the source of the Incoming category. Finally, the bidirectional connectivity is the ability to both search and retrieve data from other sources but also export your own data in a specific format. Each repository's decision is highly influenced by its distribution policy. For instance, many commercial repositories only offer incoming connectivity in order to harvest data and distribute them under their own policy as profit is a very important factor. The drawback though is that both their contents are quite poor and their user base is limited. Bidirectional connectivity is the widest and most efficient, as it both provides richer repositories and preserves commercial profit (through indirection policy).

2. Existing Frameworks

In this section an overview of alternative approaches trying to solve some aspects of the interoperability issue is presented. It presents a brief description of the approaches and comparative results, trying to bring the gap out.

ARIADNE[13], a European knowledge pool system, (54)is an application profile fully compatible with IEEE LOM that accommodates metadata sharing between repositories. ARIADNE is able to export its contents as LOM objects in order to be used from external systems. The ARIADNE approach does not conforms to IMS DRI Specification and support only outgoing connectivity.

ASIDE (55) provides a concrete framework and architecture that tries to address the identified interoperability problems and offer a framework reusable audiovisual learning objects. It is service-oriented, conforms to the IMS DRI Specification. Though, ASIDE combines METS, MPEG7 and LOM in order to construct specific audiovisual learning objects.

The EduSource project implements an open network for learning services. The project is focused on the creation of a network of linked and interoperable Learning Object Repositories across Canada. (56) To achieve its goals, the EduSource project is implementing IMS DRI specification as closely as possible and uses SOAP as a communication layer. The variety of communities that EduSource will serve contains: server-type repositories, peer-to-peer repositories, repositories of harvested metadata and outside repositories / networks.

The ELENA project works on solutions to provide personalization, openness, and interoperability in the context of smart spaces for learning. It (57) investigates how to integrate the advantages of open learning repositories with strategies and techniques successfully employed in web-based educational systems. ELENA's central component integrates and uses other services to find learning resources, courses or complete learning paths suitable for a user and provides personalized support for learners.

LEBONED project (58) focuses on the integration of Digital Libraries and their contents into web-based learning environments. They describe in general how the architecture of a standard learning management system has to be modified to enable the integration of Digital Libraries. In practice, the eVerlage Digital Library is integrated into the learning management system Blackboard.

---

[13] http://www.ariadne-eu.org/index.php

Moodle (59) is working on repository interoperability as well. It is going to provide in Version 2.0 customized plugins for a series of known repositories. Moreover, the latest Version of Moodle conforms to SCORM 1.2.

It is obvious that most of these efforts focus on specific needs and they provide customized solutions. It is important to create an ‘umbrella’ system that will connect both Digital Libraries and Learning Object Repositories based on international standards and provide a unified format of objects delivery. Moreover, there is a serious lack of systems that deliver SCORM 2004 objects.

3. Summary

In this chapter, the most important parameters of the interoperability gap between Digital Libraries and eLearning are presented, along with the several approached made towards the resolution of this issue.

Chapter Five: System Architecture

ELearning and Digital Libraries are two worlds addressing a common audience. Both of them were urged by the evolution of internet and serious research has been conducted in the area of their standardizing. However, these two scientific fields are being independently developed, although they seem to have several similarities. The scope of this thesis is to provide an effective proposal for bridging the gap between Digital Libraries and eLearning interoperability.

This chapter is separated into two sections, one for each approach used in order to achieve the goal of the thesis. The first approach presented deals with the integration of eLearning resources in a Digital Library. This work is focused on the eLearning material of University of Crete and it was funded within the Istros Project (60). This approach enhances the interoperability between eLearning and Digital Libraries; however the eLearning material of University of Crete is not created according to standards and consistent procedures, having as a consequence to result to a customized solution. This experience brought out the lack of a framework that connects Digital Libraries with eLearning in a bidirectional way, based on international standards. As a result of these findings, a second approach, introduced by the author provides a framework for federated searches between Digital Libraries and eLearning and supports the delivery of standardized unified formatted objects. This approach supports a wider range of Digital Libraries and Learning Management systems than the first one and is based on international standards, such as Dublin Core (5), IEEE LOM (8), IMS DRI (10), Z39.50 (7), SOAP (14), WSDL (13) and SCORM 2004 (11), which is the format of the object delivered. The framework is implemented utilizing the principles of the Service Oriented Architecture (12).

1. Integrating eLearning Resources in a Digital Library

The first approach was organized within Istros (60), a project run in the context of the 3rd Community Support Framework. The initial duration of the project was 7 years (01/01/2000 to 31/12/2006) and it was extended until 31/12/2008. Its ultimate goal was the further improvement of library standards and resources. This project ran from the Library and Information Center of University of Crete (UoC). The project’s objectives were divided in 5 work packages. The current approach constitutes the extension of the fifth work package and specifically the Work Package 5.2: Integrating Digital Libraries Content, which started on 01/01/2007 and in which participated Mariana Karmazi, George Koutras and the author as software engineers.

Scope of this work package was to integrate learning material of University of Crete in the Digital Collections of the institution's Library & Information Center. The work package extension's goal was the creation of a Digital Library, containing the digital learning material of the University of Crete. This Digital Library contains course lectures, video and audio streams (products of course webcasting) and related links to bibliography and the web. The content of the Digital Library is subject to searching through well-established international searching protocols and standards and the Digital Library constitutes one of the distributed Digital Repositories supported by Livesearch[14] (figure 6 and 7). Livesearch is the federated searching portal of the Library and Information Center of University of Crete.

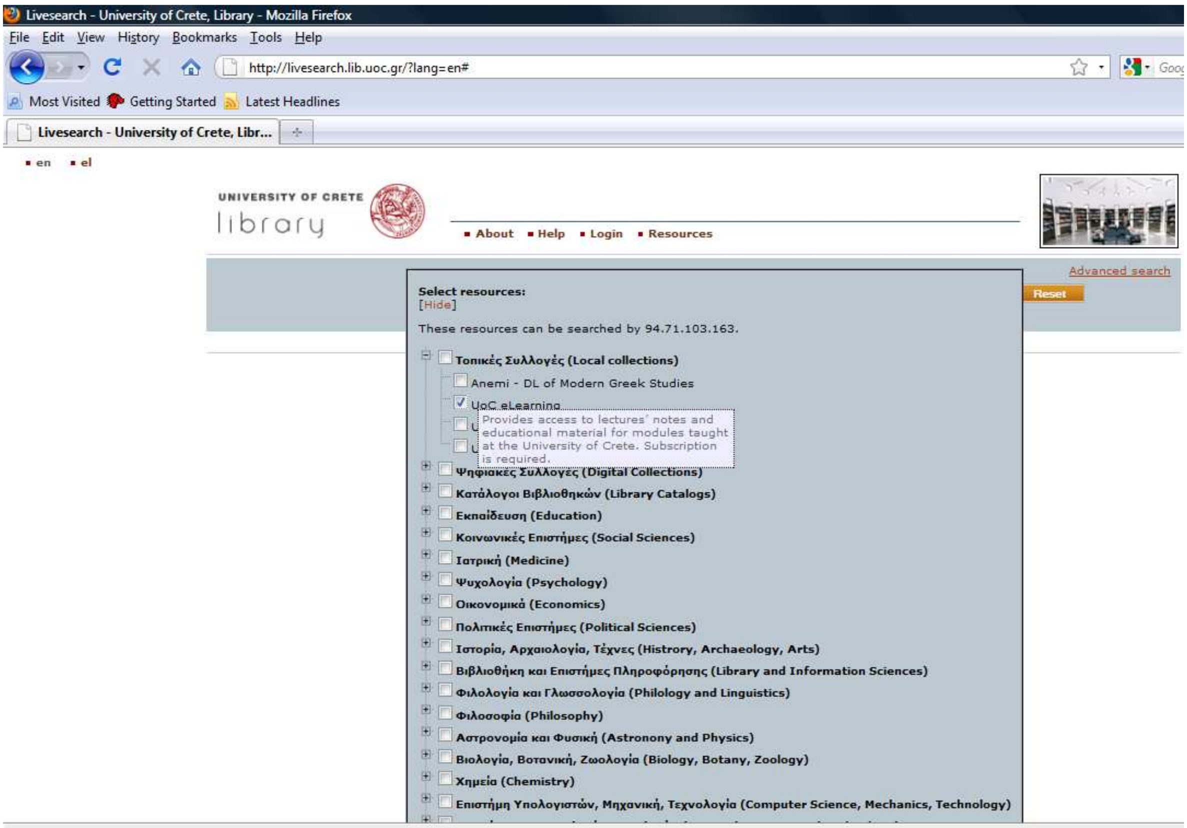

Figure 6 The Digital Library as a Livesearch target repository

[14] http://livesearch.lib.uoc.gr

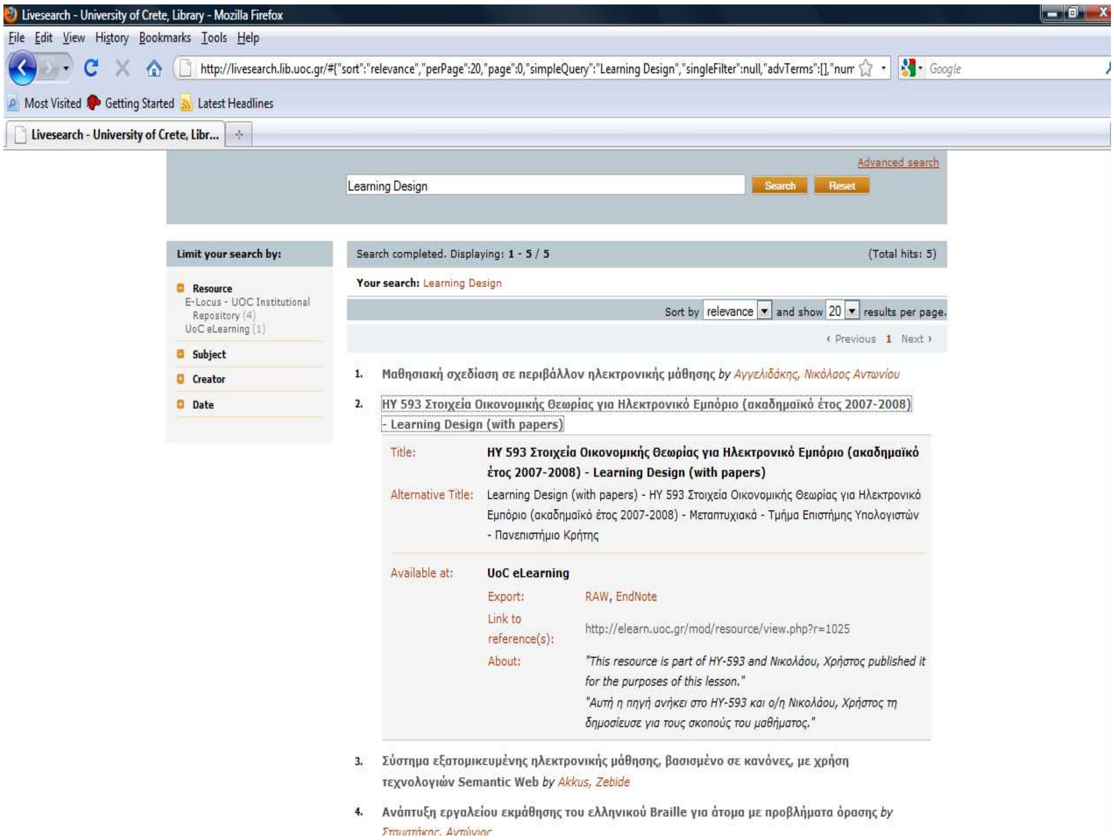

Figure 7 Searching a resource in Livesearch

The first challenge of this effort was to gather the whole eLearning material used from the instructors of UoC. The University of Crete is characterized by a lack of consensus in managing eLearning material. The eLearning material is dispersed on the instructors' personal websites, courses' websites and several platforms including Learning Management Systems, and Teaching Assisting Systems. The diversity and dissemination of eLearning material brought out a great variety of practices used for storing it.

Another important issue, regarding this approach, is that the eLearning material should have been described with metadata in order to be included in the Digital Repository; otherwise this effort would be in vain. Even in cases where metadata for the resources existed, there still were some problems in their integration in the Digital Repository. It is doubtful how the instructional information is mapped in metadata fields, so the usefulness of this metadata is dependent on its quality. The instructors have to create educational metadata both when they create their learning object and when they reuse it for a different purpose. However, it is not always effective since they lack expertise in this field. Hence there might be a wide variability in the vocabulary and terminology in educational metadata descriptions, causing serious malfunctions in searching for the appropriate learning objects and their reuse.

Taking into consideration, the above difficulties, it was decided to divide the eLearning resources into two basic categories, the static resources and the Moodle resources (figure 8). The static resources include the eLearning material coming from individual web pages, either belonging to an instructor or to a specific course. The Moodle resources include the eLearning material which is stored in the Learning Management System, Moodle (59). In the following subsections the architecture of the system is described along with the different approaches used for the two different categories.

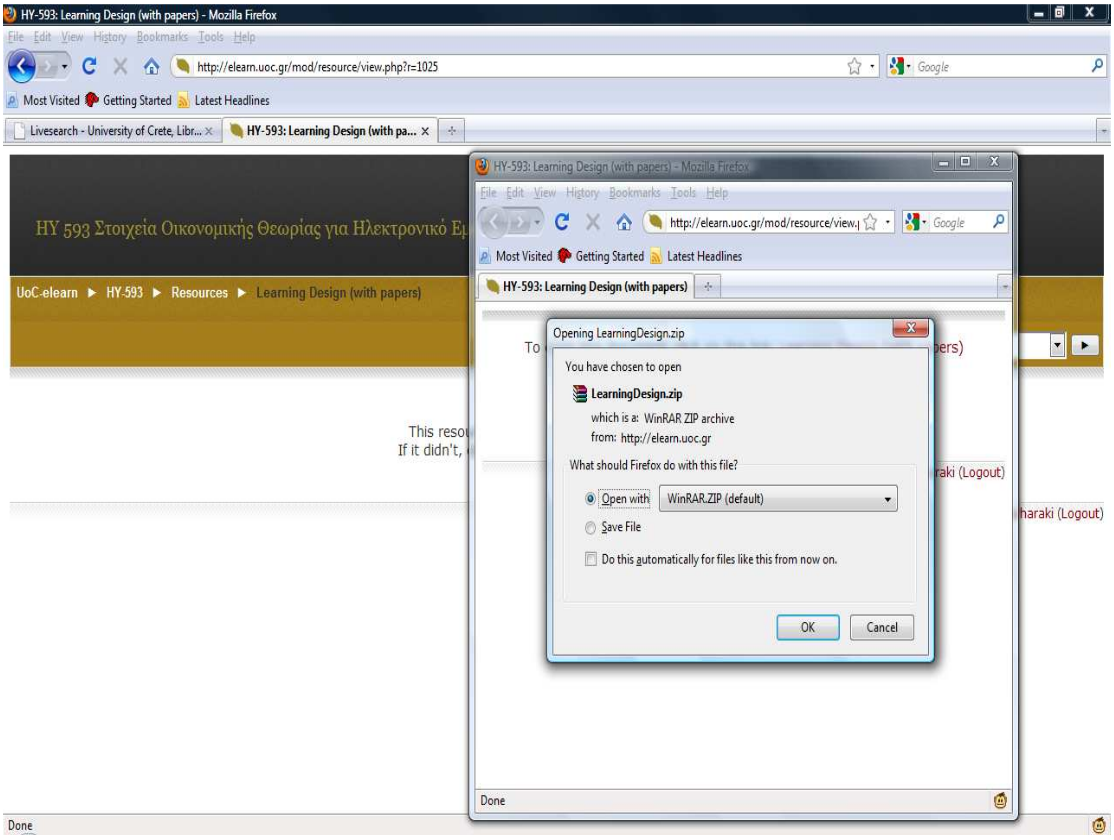

Figure 8 Moodle resource

In the following subsections, there are first described the two approaches related to the corresponding resources' categories, the static and the Moodle resources. Then, the framework's architecture is described and finally a short evaluation of this approach is provided.

1.1. Static Resources Case

The eLearning resources coming from individual web pages are integrated in the Digital Library, as well. However, these web pages are not implemented on a certain platform, which means that their resources could not be retrieved in an automatic way. With the intention to include these resources in the Digital Library, it was decided to gather all the course material that is not hosted in a learning management system and to create a metadata document for each one that describes them. The metadata documents conform to the same metadata schema that is adopted for all the resources of the Digital Library.

These documents were created manually and only once, since there is a growing trend of the instructors to leave behind the outdated practices and adopt more students- and instructors- friendly means, such as LMS. Furthermore, this procedure is very difficult to repeat due to the diversity and dissemination of the resources. The material gathered in the context of this case came from 429 different courses.

1.2. Moodle Case

The dominant Learning Management System used in University of Crete is Moodle (59). Moodle[15] is a Learning Management System i.e. a software package for producing Internet-based courses and effective online learning sites. It is a global development project designed to support a social constructionist framework of education, which means an education framework that is created in social contexts by a particular group (i.e. Moodle Community). Moodle is provided freely as Open Source software (under the GNU Public License). It can be installed on any computer running PHP and supports an SQL type database.

Moodle is widely used in the University of Crete as a Learning Management System. By this time, the institution hosts three different distributions of Moodle and these three

---

[15] The word Moodle was originally an acronym for Modular Object-Oriented Dynamic Learning Environment, which is mostly useful to programmers and education theorists. It's also a verb that describes the process of lazily meandering through something, doing things as it occurs to you to do them, an enjoyable tinkering that often leads to insight and creativity. As such it applies both to the way Moodle was developed, and to the way a student or teacher might approach studying or teaching an online course

distributions host approximately 175 courses. Moodle manages its SCORM and IMS Content Packages with a certain Learning Object Repository (Repository Api). The Moodle 1.9.5 version announced on 12/10/2009 has been certified as SCORM 1.2 compliant. However, Moodle distributions in UoC are not SCORM compliant, because they are earlier versions.

In the case of Moodle, as eLearning resources are considered the whole content uploaded or created at the LMS. There has been constructed a metadata document for any resource created or activity delivered in the context of a course, as well as for the course. These metadata documents are constructed on a regular basis and they are updated, in case of changes in the content. The metadata documents conform to the same metadata schema that is adopted for all the resources of the Digital Library and is provided in figure 9.

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
        targetNamespace="http://tsl3.tsl.gr/digital_library/"
                xmlns="http://tsl3.tsl.gr/digital_library/"
                xmlns:dc="http://purl.org/dc/elements/1.1/"
                xmlns:dcterms="http://purl.org/dc/terms/"
                elementFormDefault="qualified"
                attributeFormDefault="qualified">
</xs:schema>
```

Figure 9: Metadata schema of Digital Libary

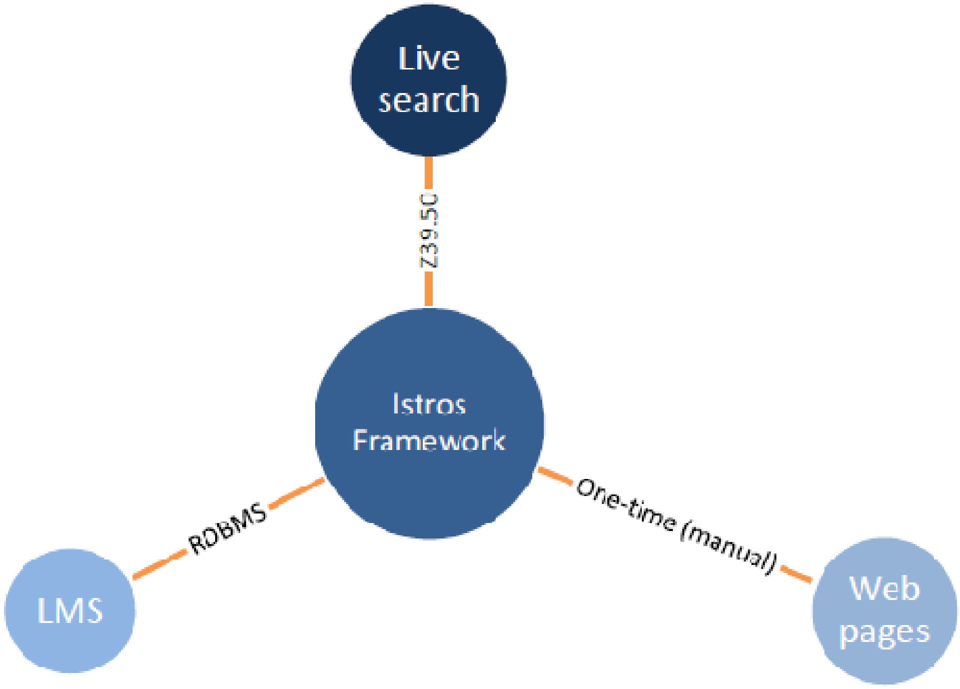

Figure 10: Integrating eLearning to a Digital Library - Framework's Architecture

1.3. System Architecture

The Moodle distributions in UoC keep their resources in databases, along with their metadata. Unfortunately, this metadata is not well documented, since while constructing a resource, filling the metadata is optional. However, the system keeps relatively enough information in order to describe resources and make them discoverable from a specialized tool (i.e. a metasearching tool such as Livesearch).

The first step in designing this framework was to connect to the Moodle databases and identify the information which was possible to add value describing these resources. According to the information needed for the resources description; an application profile is designed in order to describe the structure of the metadata documents. This application profile is a subset of the Dublin Core Metadata Element Set (5) and the Dublin Core Metadata Terms (24).

For the purpose of the Digital Library, a digital repository was created in accordance with IMS DRI Specification (10). In this repository, the metadata of the Moodle and static resources are stored, containing links for the initial resources. The reason of keeping only the metadata of the resources in the digital repository was basically the rights. Learning resources may have restricted access to a specific group of students, instructors or wider groups. The issue of rights is a crucial matter in the field of Digital Libraries and eLearning, as well. The intention of this framework is that the instructors of UoC should be able to retain the user access management of their resources. Therefore, this has been achieved, by utilizing an element of metadata for keeping the URI of the resource in the Digital Library. In this way, the user who tries to retrieve the actual resource is forced to enter the learning management system, which in its turn controls the access rights of the user in accordance the instructor's preferences.

The content of the Digital Repository is being updated using PHP scripts, which connect to the Moodle databases, collect all the updated resources and construct the updated metadata documents. Then, the content is provided to the indexing and retrieval server (Zebra server[16]) and is being indexed. These operations are controlled by Linux cron daemon[17] and are triggered on a regular basis. Zebra works as a Z39.50 server for the records and documents search, presentation, insert, update and delete operations. The UoC Livesearch exploits this Zebra's utility in order to send queries to the Digital Library and retrieve the results.

The functionality of the system, described above, can be divided in five modules. These modules are being thoroughly described in chapter six, which covers the implementation. Briefly the five modules are:

- Construct a digital repository
- Provide Z39.50 server
- Collect learning resources' information
- Construct Dublin Core Metadata for the resources
- Index Resources

---

[16] http://www.indexdata.com/zebra

[17] Cron is a unix, solaris and Linux utility that allows tasks to be automatically run in the background at regular intervals by the cron daemon. Crontab (CRON TABle) is a file which contains the schedule of cron entries to be run at specified times.

1.4. System Evaluation

The outcome of this approach is a Digital Library that aggregates and provides the eLearning resources of University of Crete. This is the only effort made in order to aggregate this valuable material in one digital collection. This Digital Library has been incorporated in the collection of Digital Libraries, Databases and collections which are subject to federated searches performed by Livesearch. The system presented in this section enhances interoperability between Digital Library and eLearning, by providing a mechanism for automatic integration of eLearning resources in a Digital Library. However, this approach is not bidirectional.

2. Framework for federated searches in Digital Libraries and Learning Object Repositories, providing unified formatted object delivery (SCORM 2004)

Influenced by the first approach, it became obvious, that a standardized framework supporting bidirectional interoperability is needed. This resulted to the second approach, which provides a more general framework that can be applied in any organization that intents to combine digital and learning objects, with the purpose of retrieving unified formatted objects. The scope of this effort is to achieve unified interface between eLearning and Digital Libraries. This framework provides federated search in both Digital Libraries and Learning Object Repositories and support the delivery of objects, based on international standards and protocols adopted by eLearning and Digital Libraries, as well. This framework is based on the principles of the Service Oriented Architecture (SOA) (12).

2.1. System Architecture

The scope of this thesis is to provide a framework for federated search in both Digital Libraries and Learning Object Repositories and support the delivery of unified formatted objects (SCORM 2004). This framework (figure 11 and 12) aims to create a bridge between Digital Libraries and eLearning, these two worlds in conjunction can provide a very rich knowledge base. The ultimate goal of this framework is to enable any user either an instructor or librarian or a student or just a "curious" internet user to have access with a unified interface to both digital and learning objects regardless where and in which format the resources are stored. All the resources provided by this framework are conformed to

SCORM 2004 specifications and can be imported, used, edited and presented in any system that supports SCORM 2004 (61).

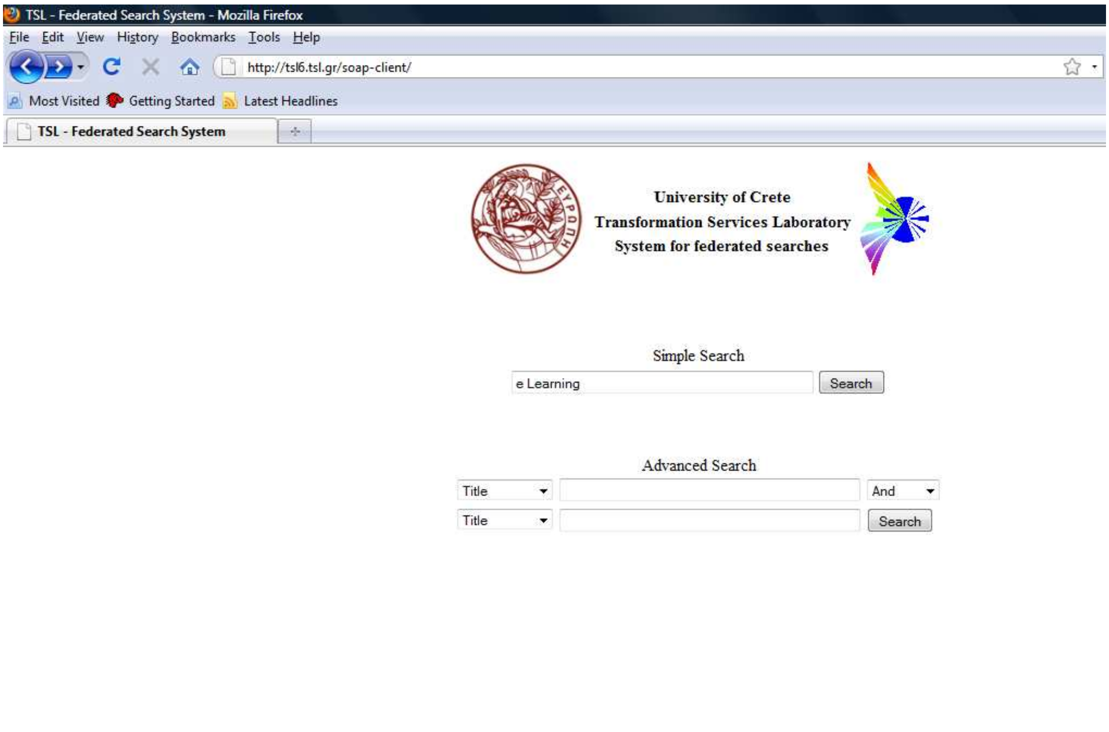

Figure 11 Interface for federated searches

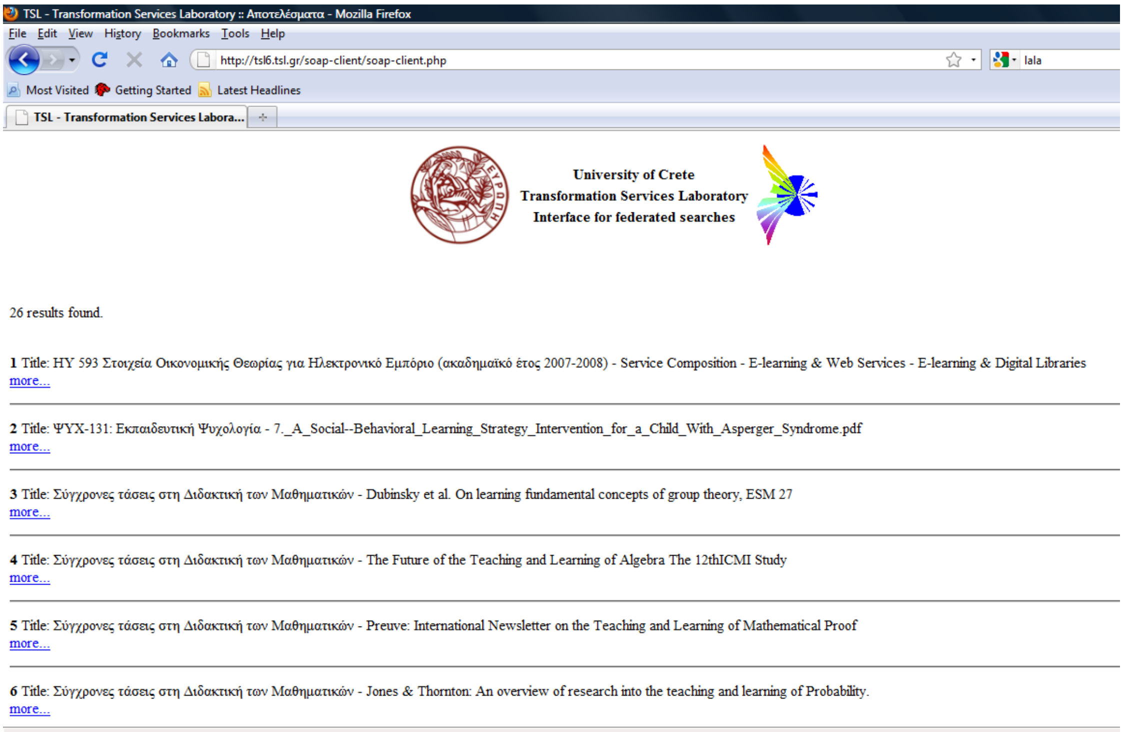

Figure 12 Search results

The first crucial issue to solve is which would be the unified format of the objects delivered. There are several standards (40), (49) that provide a reference model for content packaging, although SCORM is the only content package standard that works as an umbrella based on already established international standards and provides interoperability among platforms. SCORM combines packaging standards, metadata specifications, communication protocols, sequencing specification etc., introduced by the major organizations dealing with eLearning ADL (50), AICC (9), IMS (40), IEEE LTSC (8). The latest version of SCORM, which is the one used in this framework is SCORM 2004 version 4 which was launched in 2009. An example of a SCORM download is provided in figure 13.

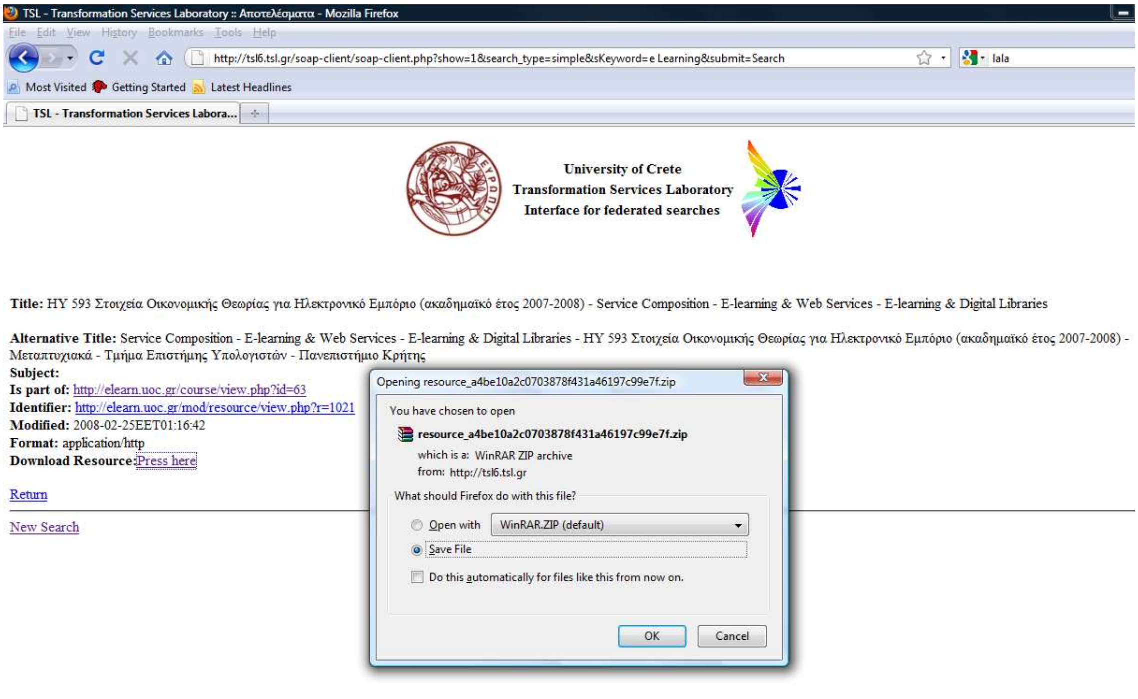

Figure 13 SCORM resource

The second issue that should be resolved from the beginning of the framework design is the storage of the resources. There are three different approaches about this issue. In the first case, all the resources are stored with their metadata in a local digital repository. In the second case, a local digital repository is created in order to store the harvested metadata from the distributed repositories. Finally, in the third case no information of the resources is kept in a local digital repository and in case of a query there are distributed queries performed in the distributed repositories that belong to the target collection. For the purpose of the specific framework, the third proposition was selected. The reason for rejecting the first approach is the preservation problem. The cost of keeping all the resources locally is growing with the enrichment of the repository and moreover mechanisms for repository updates and synchronizing should be developed. The reason for rejecting the second approach is the repository updates. In the second case, the harvested metadata has to be updated regularly, although any changes made in the distributed repositories would not be visible to the framework until the next regular update. The third approach enables the discovery of a resource by the time the new resource is inserted in its repository and the indexed records of the repository are updated. Furthermore, by choosing the third approach, the intensive tasks of indexing the metadata and storing the resources are left to the corresponding servers of the distributed repositories. So, the framework is not going to keep locally any resources, excluding these that are inserted by the respective service (Insert a learning object).

With the intention to describe the architecture of this approach, a brief overview of the framework's functionality is provided. The services provided to the users are divided in three categories; there are the services, which enable the user to add a Digital Library, a Learning Object Repository and a Learning Object into the collection of the target repositories, there are the services that support the searching interface and finally, there are the services that constitute the core mechanism of the framework. All the described services are SOAP Web services; the WSDL documents describing them are available in the chapter Implementation.

In the first category of services, the user can add his object or resources' collection, as far as they conform to the framework's specifications. In order to insert a Learning Object, the user should ensure that the object conforms to the SCORM 2004 specifications and includes metadata, since metadata is optional in SCORM but necessary for the resource's discovery in the framework. In the case of inserting a Digital Library, its metadata schema should conform to the Dublin Core Metadata Initiative Metadata Terms (24) . The Digital Library has to support a Z39.50 (7) server for managing searching queries and the user, who calls the service, should provide the authentication's information for connecting to the server. Finally, in case of inserting a Learning Object Repository, the content of the repository should be SCORM 2004 (11) objects and the repository should conform to the IMS DRI (10) specification. Respectively, the user should send the authentication's information of connecting to the Z39.50 server.

In the second category, the framework provides the users with two services; the searching and the delivery service. Calling the searching service, the user submits his searching request to the framework, filling just the desirable arguments of the service's argument set. The service responds to the client with the result set, as created by the searchingCore service. The delivery service is called by the client, which requests a specific resource. The service responds to the client with the resource, in SCORM 2004 format, as it was created by the SCORMpackageBuilder service.

The searchingCore and SCORMpackageBuilder services along with the indexing service are the core mechanism of the framework. The searchingCore service orchestrates the searching process. It is called by the searching service, which sends the searching request in the appropriate format to the searchingCore service. The searchingCore service broadcasts the searching query in the distributed repositories via Z39.50 protocol, collects the results and sends back the result set. The SCORMpackageBuilder service is called by the delivery service, which sends the request id of the resource. The SCORMpackageBuilder checks if the resource is already a SCORM 2004 object, in the opposite case it constructs the SCORM 2004 object from a digital resource and sends to the calling service the URI of the new resource. Finally, the indexing service is responsible for indexing regularly the resources of the framework's local repository, which stores the Learning Objects added to the framework. This local repository provides a Z39.50 server and belongs to the collection of the searching repositories.

The architecture of the framework presented in figure 14 is constituted by eight major modules. The modules are:

- Insert a Digital Library
- Insert a Learning Object Repository
- Insert a Learning Object
- Searching interface
- Searching mechanism
- Indexing mechanism
- SCORM package builder
- Delivery of SCORM objects

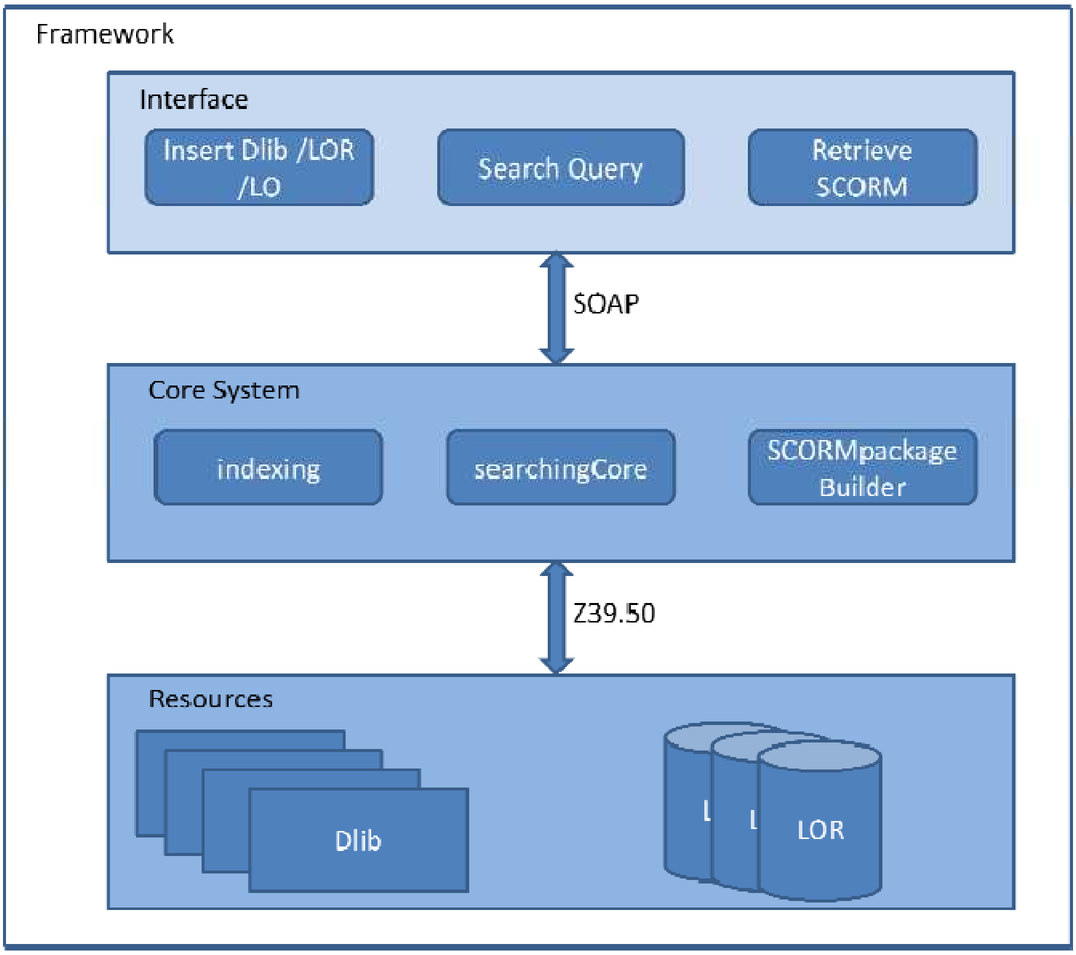

Figure 14: Framework's Architecture

2.2. System Evaluation

This system provides a standardized procedure for incorporating Digital Libraries, Learning Object Repositories and Learning Objects in a collection of resources, which are subject to unified searching and object delivery. This system achieves a bidirectional approach regarding the issue of interoperability. Digital and eLearning objects are combined in a framework that supports federated searches and enhances the exchange of objects among Digital Libraries and Learning Management systems. The delivered objects conform to SCORM 2004 and ensure the standardized exchange of the objects, enhancing interoperability between Digital Libraries and eLearning.

3. Summary

In this chapter, the architecture of two frameworks is presented. These two frameworks aim to bridge the gap between Digital Libraries and eLearning community. The first approach presented a framework for enhancing the integration of eLearning resources to the world of Digital Libraries. It presents an approach, according to which eLearning resources that are hosted in eLearning platforms, based on Moodle, can be integrated to a digital collection and incorporated to the collection of Digital Libraries supported by the Library and Information Center of University of Crete. The second approach that is presented, describes a framework that supports bidirectional interoperability between Digital Libraries and eLearning. This framework enables the incorporation of Digital Libraries and Learning Object Repositories in a federated search context and enables the standardized delivery of unified formatted objects (SCORM 2004), with intention to enhance the exchange of objects among Digital Libraries and Learning Managements systems.

Chapter Six: Implementation

In this chapter, the implementation of the two presented approaches is described. Information about the standards, protocols, tools and techniques that are used in order to realize the frameworks' architecture are presented. In the first section, the modules that constitute the "Integration eLearning Resources in a Digital Library" approach are described and in the second section the modules of the "Framework for federated searches in Digital Libraries and Learning Object Repositories, providing unified formatted object delivery (SCORM 2004)".

1. Integrating eLearning Resources in a Digital Library

1.1. Construct a digital repository

For the purpose of Digital Library, a digital repository was constructed. In this digital repository, as it was argued in the chapter of the Architecture, are stored just metadata related to the eLearning resources of UoC. The repository is constructed according to the IMS DRI (10) specification. It supports the search/expose, gather/expose, submit/store, request/deliver, and alert/expose operations. The search function is implemented with Z39.50 protocol and the gather function is implemented with the OAI PMH (6) protocol. All the metadata documents stored in this digital repository conform to a specific XML schema that is provided in figure 15.

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
        targetNamespace="http://tsl3.tsl.gr/digital_library/"
                xmlns="http://tsl3.tsl.gr/digital_library/"
                xmlns:dc="http://purl.org/dc/elements/1.1/"
                xmlns:dcterms="http://purl.org/dc/terms/"
                elementFormDefault="qualified"
                attributeFormDefault="qualified">
</xs:schema>
```

Figure 15: Metadata Schema

1.2. Provide Z39.50 server

In order to make the eLearning resources in the Digital Library discoverable, a Z39.50 server is provided. In this way, any Z39.50 client could send searching requests to the Digital Library and retrieve the appropriate results set. In the UoC case, Livesearch supports a Z39.50 client that communicates with the Z39.50 server of the Digital Library.

For the purpose of the Z39.50 server, Zebra server was selected, which is based on an open-source protocol implementation. Zebra server is software provided by Indexdata[18] and is the most widely used in the industry of indexing and resource retrieval. Zebra is a high-performance, general-purpose structured text indexing and retrieval engine. It reads structured records in a variety of input formats (eg. Email, XML, MARC) and allows access to them through exact Lifecycle search expressions and relevance-ranked free-text queries. Access to data, which is stored in Zebra, is achieved by any Z39.50 client. The bib-1 attribute set included in the Z39.50 protocol, defines six attribute types (i.e., use, relation, position, structure, truncation, and completeness) and assigns values specific to each type. The bib-1 attribute set is used for expressing queries for searching bibliographic databases and other information resources. A representative part of bib-1 attribute set, which contains the Dublin Core Use Attributes , needed for the configuration of our Z39.50 server is provided in the figure 16.

att 1097	DC-Title
att 1098	DC-Creator
att 1099	DC-Subject
att 1100	DC-Description
att 1101	DC-Publisher
att 1102	DC-Date
att 1103	DC-ResourceType
att 1104	DC-ResourceIdentifier
att 1105	DC-Language
att 1106	DC-OtherContributor
att 1107	DC-Format
att 1108	DC-Source
att 1109	DC-Relation
att 1110	DC-Coverage
att 1111	DC-RightsManagement

Figure 16: Bib-1 Attribute Set

[18] http://www.indexdata.com/

1.3. Collect learning resources' information

In UoC, there are three distributions of Moodle. The first[19] and most widely used is the one provided by Ucnet, which is a center in UoC responsible for networking in the University. The second distribution[20] is provided by the Department of Economics and concerns only courses performed in the Department. The third distribution[21] is provided by the Department of Primary Education concerns only courses of this department, as well.

In order to collect the information regarding the resources of the three distributed systems, three PHP scripts have been implemented, which are controlled by Linux cron daemon[22] and are triggered on a regular basis. Each one of the daemons connects to the database of each Moodle distribution and collects the information of all the resources that have been altered or inserted since the last time the daemon ran. In the Moodle database, apart from the actual resources, information is kept such as title, contributor, date modified, description of the resource etc that is used later in the metadata construction module.

1.4. Construct Dublin Core Metadata for the resources

The metadata that describes the eLearning resources added in the digital repository is a subset of DCMI (Dublin Core Metadata Initiative) Metadata Terms (24) and DCMI Element Set (5) . An application profile is designed and the metadata documents should conform to it, so as to be validated. The elements used from the two Dublin Core Specifications are described in Table 6 Metadata elements of Application Profile.

---

[19] http://elearn.uoc.gr

[20] http://wolverine.soc.uoc.gr/moodle/

[21] http://pegasus.clab.edc.uoc.gr/

[22] Cron is a unix, solaris and Linux utility that allows tasks to be automatically run in the background at regular intervals by the cron daemon. Crontab (CRON TABle) is a file which contains the schedule of cron entries to be run at specified times.

Element name	Element Description
dc:title	The title of the resource
dcterms:alternative	The title of the resource in conjunction with the name of the course, the Department where the course belongs and the name of the Institution. For example “Lecture 2 – Introduction to Python – CS100 Introduction to Computer Science – Computer Science Department – University of Crete”
dc:contributor	The name of the instructor
dcterms:abstract	The summary of the resource’s content
dc:subject	This element is used only in metadata documents describing courses and contains the subjects related to each course according to the classification of Library of Congress
dc:description	Information about the instructor and the course. For example “This resource is part of CS100 Introduction to Computer Science and Christos, Nikolaou published it for the purposes of this lesson.”
dcterms:isPartOf	The URI of the course, where the resource belongs
dc:identifier	The URI of the resource
dcterms:modified	The date and time when the resource was last modified
dc:format	The MIME Type of the resource

Table 6: Metadata elements of Application Profile

The schema that the metadata documents conform to is presented in the figure 15. Metadata documents are constructed not only for the eLearning resources of every course hosted in Moodle but also for the courses themselves. The construction of the metadata documents is performed by a PHP script that is called by the daemons presented in the previous section. The xml files that encode the metadata are implemented using the Document Object Model (DOM) library of PHP 5.

1.5. Index resources

One of the most significant modules in this framework is the process of indexing the resources. The indexing method affects the quality of the searching results. For this purpose, Zebra server is selected, which is an indexing and resource retrieval engine. Zebra server was presented in a previous section (1.1). The indexing of the resources is performed on a regular basis, after the creation of the metadata documents of the resources.

The index configuration of the server for the metadata documents of the Digital Library is presented in figure 17. This document represents the rules, which are used by the indexer to parse the metadata documents and create the index. For example, on attribute DC-Title the indexer will index the content of the metadata documents as word (:w), phrase (:p) and string (:s).

```
attset bib1.att
xpath enable

encoding UTF-8
#esetname F @
#esetname M link-medium.est
#esetname S link-short.est

xelm /*/dc:title                 Body-of-text:w,DC-Title:w,DC-Title:p,DC-Title:s,
                                 DC-Description:w, Title:w,Any:w
xelm /*/dcterms:alternative      Body-of-text:w,DC-Title:w,DC-Title:p,DC-Title:s,
                                 DC-Description:w,Any:w
xelm /*/dc:subject               Body-of-text:w,DC-Subject:p,DC-Subject:w,
                                 Subject-heading:w,Any:w
xelm /*/dc:description           Body-of-text:w,DC-Description:w,DC-Description:p,
                                 Any:w
xelm /*/dcterms:modified         Date/time-last-modified:s
```

Figure 17: Index configuration for the metadata documents

2. Framework for federated searches in Digital Libraries and Learning Object Repositories, providing unified formatted object delivery (SCORM 2004)

The second approach that is based on the principles of the Service Oriented Architecture provides a Web services' API for inserting Digital and Learning content, for performing federated searches and for delivering SCORM object results. This API is constituted by eight modules that are implemented by utilizing Web services standards (eg. WSDL, SOAP). The services have published interface, they are described using WSDL 1.1 (13) and they communicate with each other, by exchanging SOAP 1.2 (14) messages, in order to collectively support a specific process.

In this section a short overview of the Service Oriented Architecture is provided, so as to introduce the reader into the basic principles on which the implementation of the framework is based. After this overview, the constituent modules of the framework are described, providing detailed information about their implementation.

2.1. Service Oriented Architecture

Service Oriented Architecture (SOA) (12) provides a set of principles and requirements of loosely coupled, standards-based, and protocol-independent distributed computing, releasing organizations of their system and subsystems strict interrelationships. SOA is focused on creating a design style, technology, and process framework that will allow organizations to develop, interconnect, and maintain enterprise applications and services efficiently and cost effectively.

In a SOA, software resources are described as “services”, which are well defined, self-contained modules that provide standard functionality and are independent of the state or context of other services (12). The description of the services is achieved by a published interface, utilizing a standard definition language. The services communicate with each other requesting an operation to be performed in order to collectively support a common task or process. The most common services today utilize Web services standards, e.g.,Web Services Description Language (WSDL) (13), Simple Object Access Protocol (SOAP) (14), and Universal Description, Discovery and Integration registry (UDDI) (62).

2.2. Insert a Digital Library

In this module, a service is provided which incorporates a digital repository in the collection of the repositories that are being used in the federated searches. The user, who consumes this service, provides the necessary information in order to enable the system to connect and send Z39.50 (7) queries. The digital repository intending to be added in the collection should support Z39.50 searches and its metadata should conform to the Dublin Core

Metadata Initiative Metadata Terms (24). The messages and PortType of the WSDL file is presented in figure 18.

```
<message name='serverLocation'>
  <part name='Keyword' type='xsd:string'/>
</message>
<message name='serverPort'>
  <part name='Result' type='xsd:string'/>
</message>
<message name='serverUsername'>
  <part name='Keyword' type='xsd:string'/>
</message>
<message name='serverUsername'>
  <part name='Result' type='xsd:string'/>
</message>
<portType name='InsertDlibPortType'>
  <operation name='insert'>
    <input message='tns:serverLocation'/>
            <input message='tns:serverPort'/>
            <input message='tns:serverUsername'/>
            <input message='tns:serverUsername'/>
  </operation>
</portType>
```

Figure 18: Messages and PortTypes of Insert a Digital Library WSDL

2.3. Insert a Learning Object Repository

This module respectively to the previous one enables a Learning Object Repository to be added in the collection of the repositories so as to be referenced in the federated searches. The Learning Object Repository should follow some certain prerequisites. The basic requirement is to be IMS DRI (10) conformed. The content of the repository should also conform to SCORM 2004 standard, enriched with metadata, since metadata are optional in SCORM 2004. This is important for the searching mechanism (searchingCore service); in order to provide efficient results and exploit the whole content of the repository. Finally, if its resources are kept locally, they should be free of access, since the framework does not perform any identity checking. The user who consumes the service should send authorization information about the repository and its Z39.50 server. The messages and PortType of the WSDL file is presented in figure 19.

```
<message name='serverLocation'>
  <part name='Keyword' type='xsd:string'/>
</message>
<message name='serverPort'>
  <part name='Result' type='xsd:string'/>
</message>
<message name='serverUsername'>
  <part name='Keyword' type='xsd:string'/>
</message>
<message name='serverUsername'>
  <part name='Result' type='xsd:string'/>
</message>
<portType name='InsertLORPortType'>
  <operation name='insert'>
    <input message='tns:serverLocation'/>
          <input message='tns:serverPort'/>
          <input message='tns:serverUsername'/>
          <input message='tns:serverUsername'/>
  </operation>
</portType>
```

Figure 19: Messages and PortTypes of Insert a Learning Object Repository WSDL

2.4. Insert a Learning Object

In this module, the user of the service can upload a learning object to the local repository of the framework. The learning object should be a SCORM 2004 object and contain metadata, since without metadata; it cannot be indexed and later discovered from the searching mechanism. As it was mentioned in the previous module, metadata in a SCORM object is not mandatory, although for this module it is.

2.5. Searching interface

This module is constituted by a service which performs the role of an intermediate between the user and the searching mechanism. This service is called by the client, which sends the searching query. The query does not need to be expressed in PQF (Prefix Query Format) since the service converts the query to the appropriate format. The client service that calls the searching interface service should just fill the appropriate arguments of the service set. The arguments of the service in the case of advance search are title, subject, description etc. and in the case of simple search is just one, covering all the metadata. The messages and portType of WSDL file of the searching interface service is presented in figure 20.

```
<message name='SearchRequest'>
  <part name='Keyword' type='xsd:string'/>
</message>
<message name='SearchResponse'>
  <part name='Result' type='xsd:string'/>
</message>
<message name='AdvancedSearchRequest'>
  <part name='Keyword' type='xsd:string'/>
</message>
<message name='AdvancedSearchResponse'>
  <part name='Result' type='xsd:string'/>
</message>
<portType name='SearchPortType'>
  <operation name='getSimpleSearch'>
    <input message='tns:SearchRequest'/>
    <output message='tns:SearchResponse'/>
  </operation>
  <operation name='getAdvancedSearch'>
    <input message='tns:AdvancedSearchRequest'/>
    <output message='tns:AdvancedSearchResponse'/>
  </operation>
</portType>
```

Figure 20: Messages and PortTypes of the Searching Interface WSDL

2.6. Searching mechanism

The searching mechanism is a service that performs the management of the searching procedure in digital and learning object repositories. This service broadcasts the searching query in the distributed repositories, collects the results and sends back the result set. The broadcasted queries are performed by Z39.50 clients that connect to the distributed digital and learning object repositories. These clients are implemented in PHP/YAZ. PHP/YAZ is a PHP extension that implements Z39.50 client functionality. A part of the PHP/YAZ client code is represented in figure 21. The results are collected and sent back to the service client in XML.

```
$y = yaz_connect($att['target'], array('persistent' => FALSE, 'charset' => 'UTF-8'));
yaz_range($y, $start, $number);
yaz_syntax($y, $syntax);
…
yaz_search($y, 'rpn', $query);
yaz_wait();
if (yaz_errno($y))
        return "<error code=\"" . yaz_errno($y) . "\">" . yaz_error($y) . " at target " .
        $att['target'] . "</error>";
  $hits = yaz_hits($y);
```

Figure 21: Z39.50 client

2.7. Indexing mechanism

The indexing mechanism is a daemon used only for the local repository. In this version of the framework, local is only the learning object repository that stores the learning objects inserted directly in the system. For the purpose of this service, a Zebra server[23] has been installed. This daemon (figure 22) runs on regular basis and indexes the content of the local Learning Object Repository.

```
/usr/bin/zebraidx-2.0 init
/usr/bin/zebraidx-2.0 update content
/usr/bin/zebraidx-2.0 commit
```

Figure 22: Indexing mechanism script

2.8. SCORM package builder

This module is the most important part of the framework. This is the service responsible for constructing the SCORM 2004 object from a digital resource. The SCORM object is a SCORM 2004 version 4. This service is responsible for constructing the imsmanifest.xml and for the object's content packaging. This service is consumed by the Delivery of SCORM objects service, when the user sends a download request to the system.

2.8.1. Resource Content Package versus Content Aggregation Content Package

The purpose of the content package is to provide a standardized way to exchange learning content between different systems or tools. SCORM standard supports two SCORM Content

[23] http://www.indexdata.com/zebra

Package Application Profiles. These application profiles adhere to XML specifications and maintain the structure of IMS Content Packaging Information Model. The first application profile is the Resource Content Package Application Profile. This is a content package that contains only resources (i.e., no organization). This type of content package can be used for bundling a set of learning resources with no defined organization or content structure. These learning resources may or may not have relationships between each other. The second application profile is the Content Aggregation Content Package Application Profile. This is a content package for bundling a set of learning resources and their intended static structure and sequencing requirements (i.e., the manifest contains 1 or more organizations of the learning resources) (51).

For the purpose of this framework the Resource Content Package Application profile is chosen, since the lack of organization and hierarchy enhances the object's reusability. These content packages contain only the resources and their metadata, enabling their move among courses and systems regardless hierarchical restrictions. Resource Content Packages adhere to the suggestion (37), (38), (39) presented in the Chapter of eLearning that the smaller and more specific the learning objects are, the more reusable they can be found in new instructional contexts.

2.8.2. SCORM construction

The imsmanifest.xml is the keystone of a SCORM object. The manifest (imsmanifest.xml) file is an XML document providing a structured inventory of the content of the package. This file contains information about the content's organization, if there is any. In this document, metadata information is kept, either by providing it inline, or by providing a link to a local metadata document. Metadata is used to describe content package as a whole and can also be used at various locations within the manifest to describe the different aspects of the content package. For the purpose of this framework, metadata is provided inline and only once in each package. The SCORM standard supports any metadata schema, although ADL highly recommends the use of the IEEE LOM metadata schema, which is followed by this framework.

In order to convert digital objects to SCORM 2004 objects, it is necessary to map the metadata of the digital objects that is Dublin Core to IEEE LOM. The semantics of the Dublin Core elements are intentionally rather broad. It should be noted that the mapping is a one-way mapping, from the LOM to Dublin Core and that transforming meta-data from the IEEE LOM to Dublin Core will result in the loss of information. In table 7 Mapping Dublin Core Element Set to IEEE Learning Object Metadata are the mapping rules from Dublin Core to IEEE LOM presented.

Dublin Core Metadata Elements	**IEEE Learning Object Metadata**
dc:identifier	1.1.2: /lom/general/identifier/entry
dc:title	1.2: /lom/general/title
dc:language	1.3: /lom/general/language
dc:description	1.4: /lom/general/description
dc:subject	1.5: /lom/general/keyword or 9: /lom/classification with 9.1: /lom/classification/purpose equals “discipline” or “idea”.
dc:coverage	1.6: /lom/general/coverage
dc:type	5.2: /lom/educational/learningResourceType
dc:date	2.3.3: /lom/lifecycle/contribute/date when 2.3.1: /lom/lifecycle/contribute/role has a value of “publisher”.
dc:creator	2.3.2: /lom/lifecycle/contribute/entity when 2.3.1: /lom/lifecycle/contribute/role has a value of “author”.
dc:otherContributor	2.3.2: /lom/lifecycle/contribute/entity with the type of contribution specified in 2.3.1: /lom/lifecycle/contribute/role
dc: publisher	2.3.2: /lom/lifecycle/contribute/entity when 2.3.1: /lom/lifecycle/contribute/role has a value of “publisher”.
dc:format	4.1: /lom/technical/format

dc:rights	6.3: /lom/rights/description
dc:relation	7.2.2: /lom/relation/resource/description
dc:source	7.2: /lom/relation/resource when the value of 7.1: /lom/relation/kind is “isBasedOn”.

Table 7: Mapping Dublin Core Element Set to IEEE Learning Object Metadata

Apart from the imsmanifest.xml SCORM package contains the actual resource. In the case of this framework, the SCORM packages do not keep the eLearning resource locally, since there is the access rights issue, as it has been mentioned before. The resource kept in the SCORM package is a web document that contains a link to the actual resource, leaving the access rights management to the distributed collections.

2.9. Delivery of SCORM objects

This module is a service which is called in case the user selects one resource of the result set to download. This service calls the SCORM package Builder module for the certain resource and delivers the SCORM object to the user via URL.

3. Summary

In this section an overview of the implementation of the two suggested approaches provided in this thesis is presented. In the first section, the five modules which constitute the Integration of eLearning resources to a Digital Library were presented. Respectively, in the second section referring to the framework for federated searches and unified formatted object delivery, were presented the eight modules that constitute the approach architecture along with an overview of what is the Service Oriented Architecture.

Chapter Seven: Conclusion & Future Work

This thesis deals with the interoperability issue between Digital Libraries and eLearning and focuses on the resources' exchange and representation. In this thesis, are presented two approaches, which aim in the solution of interoperability issue. The first approach represents a framework that enables the eLearning resources provided by University of Crete, to be integrated in a Digital Library. This approach achieves the effective integration of Learning Objects to a Digital Library and enables users to have access to both eLearning and Digital objects through the Livesearch, which is the federated searching portal of the Library and Information Center of University of Crete. However, this approach is customized to the needs of University of Crete and it brings out the need for a standardized framework to enhance the interoperability issue between Digital Libraries and eLearning. The second approach presented in this thesis, provides a framework for federated searches between Digital and Learning Object repositories and supports the delivery of unified formatted objects that are SCORM 2004 objects. This framework provides a Web services API for inserting content to the targeted repositories, for performing federated searches and finally for constructing and delivery of SCORM 2004 objects.

There are some missing features in the second approach, regarding the mechanism for the federated searches. It is highly recommended for an efficient mechanism for federated searches to apply a ranking and a de-duplicate algorithm upon the results. A relevance ranking algorithm compares results from all distributed sources against one another and displays the results in order. A de-duplicate algorithm identifies duplicate results in the federated searches' response and provides only one result for each resource. This process is a real challenge, since two resources may have the same title and author but might actually be different revisions of one document. Another feature that is important to be added in this framework is a Learning Object authoring tool that is going to enable users to compose a Learning Object, by combining SCORM objects delivered by the framework and enrich them with sequencing rules.

Bibliography

1. **LOC.** Library of Congress - Resources for Librarians and Archivists. *Library of Congress.* [Electronic] 8 November 2009. http://www.loc.gov/library/.

2. **DCMI.** DCMI Specifications. *Dublin Core Metadata Initiative.* [Electronic] 5 October 2009. http://dublincore.org/specifications/.

3. **OAI.** Open Archives Initiative - Specifications. *Open Archives Initiative.* [Electronic] http://www.openarchives.org/.

4. MARC Standards. *Library of Congress.* [Electronic] http://www.loc.gov/marc/.

5. Dublin Core Metadata Element Set, Version 1.1. *Dublin Core Metadata Initiative.* [Electronic] 14 1 2008. http://www.dublincore.org/documents/dces/.

6. **Lagoze, Carl, et al.** The Open Archives Initiative Protocol for Metadata Harvesting. *OAI-PMH Version 2.0 Specification.* [Electronic] 7 December 2008. http://www.openarchives.org/OAI/openarchivesprotocol.html.

7. *Information Retrieval (Z39.50): Application Service Definition and Protocol Specification.* Bethesda, Maryland : National Information Standards Organization (NISO), 2003.

8. **LTSC, IEEE.** IEEE Standard for Learning Object Metadata. *IEEE Learning Technology Standards Committee (LTSC).* [Electronic] 2002. http://ieeexplore.ieee.org/servlet/opac?punumber=8032.

9. **AICC.** AICC Guidelines and Reccommendations. *AICC - Aviation Industry CBT Committee.* [Electronic] 02 April 2001. http://www.aicc.org/docs/tech/cmi001v3-5.pdf.

10. IMS GLC: Digital Repositories Specification. *IMS Global Learning Consortium.* [Electronic] 30 January 2003. http://www.imsproject.org/digitalrepositories/.

11. SCORM 2004 4th Edition Version 1.1 Documentation Suite. *SCORM.* [Electronic] 14 8 2009. http://www.adlnet.gov/Technologies/scorm/SCORMSDocuments/2004%204th%20Edition/Documentation.aspx.

12. *Service oriented architectures: approaches, technologies and research issues.* **Papazoglou, Mike P. and van den Heuvel, Willem-Jan.** 2007, The VLDB Journal, p. 389-415.

13. **Christensen, Erik, et al.** Web Services Description Language (WSDL) 1.1. *World Wide Web Consortium.* [Electronic] 15 March 2001. http://www.w3.org/TR/wsdl.

14. **Mitra, Nilo and Lafon, Yves.** SOAP Version 1.2 Part 0: Primer (Second Edition). *World Wide Web Consortium.* [Electronic] 27 April 2007. http://www.w3.org/TR/2007/REC-soap12-part0-20070427/.

15. **Lynch, Clifford, et al.** *CNI white paper on networked information discovery and retrieval.* Washington, DC : Coalition for Networked Information, 1995.

16. **Greenstein, Daniel and Thorin, Suzanne.** *The Digital Library: A Biography.* Washington : Digital Library Federation, 2002.

17. **Arms, William.** *Digital Libraries.* s.l. : MIT Press, 2000.

18. *The Digital Library Manifesto.* **Candela, Leonardo, et al.**

19. **Candela, Leonardo, et al.** *The DELOS Digital Library Reference Model Foundations for Digital Libraries.* 2007.

20. *Where Do We Go From Here? The Next Decade for Digital Libraries .* **Lynch, Clifford.** July/August 2005, D-Lib Magazine.

21. *A framework for distributed digital object services.* **Kahn, Robert and Wilensky, Robert.** 2006, International Journal on Digital Libraries, p. 115-123.

22. *Flexible and Extensible Digital Object and Repository Architecture (FEDORA).* **Payette, Sandra and Lagoze, Carl.** Berlin / Heidelberg : Springer, 1998, ECDL'98, LNCS 1513, p. 41-59.

23. DOI® Handbook. *The DOI® System.* [Electronic] 8 June 2009. http://www.doi.org/hb.html.

24. DCMI Metadata Terms. *Dublin Core Metadata Initiative.* [Electronic] 14 1 2008. http://dublincore.org/documents/dcmi-terms/.

25. DCMI Type Vocabulary. *Dublin Core Metadata Initiative.* [Electronic] 14 1 2008. http://dublincore.org/documents/dcmi-type-vocabulary/.

26. *Metadara Encoding and Transmission Standard: Primer and Reference manual.* s.l. : Digital Library Federation, 2007.

27. *An Introduction to the Search/Retrieve URL Service (SRU).* **Morgan, Eric Lease.** 2004, Ariadne Issue 40 .

28. SRU Search/Retrive via URL . *SRU Search/Retrive via URL .* [Electronic] 5 August 2009. http://www.loc.gov/standards/sru/.

29. **Ellis, Ryann K.** Field Guide to Learning Management Systems. *ASTD Learning Circuits.* [Electronic] 2009. http://www.astd.org/NR/rdonlyres/12ECDB99-3B91-403E-9B15-7E597444645D/23395/LMS_fieldguide_20091.pdf.

30. *Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy.* **Wiley, David A.** 2000, The Instructional Use of Learning Objects: Online Version.

31. **Europe, ORMEE - Observatory on Rights Management for eLearning in.** Best practices in dealing with Learning Objects. *elearningeuropa.info Promoting innovation in lifelong learning.* [Electronic] 2003. http://www.elearningeuropa.info/files/media/media8541.pdf.

32. **L'Allier.** NETg's precision skilling: The linking of occupational skills descriptors to training interventions. *SkillSoft e-learning.* [Electronic] 1998. http://www.netg.com/research/pskillpaper.htm.

33. **Koper, Rob.** Combining re-Usable Learning Resources to Pedagogical Purposeful Units of Learning. [book author] A. Littlejohn και S. Buckingham Shum. *Reusing Online Resources: A Sustainable Approach to eLearning.* London : Kogan Page, 2003.

34. **Rehak, Dan and Mason, Robin.** Keeping the Learning in Learning Objects. [book author] A. Littlejohn και S. Buckingham Shum. *Reusing Online Resources: A Sustainable Approach to eLearning.* London : Kogan Page, 2003.

35. **Laurillard, Diana and McAndrew, Patrick.** Reusable Educational Software: A Basis for Generic e-learning Tasks. [book author] A. Littlejohn και S. Buckingham Shum. *Reusing Online Resources: A Sustainable Approach to eLearning.* London : Kogan Page, 2003.

36. **Duncan, Charles.** Conceptions of Learning Objects: Social and Educational Issues. [book author] A. Littlejohn και S. Buckingham Shum. *Reusing Online Resources: A Sustainable Approach to eLearning.* London : Kogan Page, 2003.

37. **Liber, Oleg και Olivier, Bill.** Learning Technology Interoperability Standards. [book author] A. Littlejohn και S Buckingham Shum. *Reusing Online Resources: A Sustainable Approach to eLearning.* London : Kogan Page, 2003.

38. **Quinn, Clark.** Learning Objects and Instruction Components. *IFETS - Discussions.* [Electronic] February 2000. http://ifets.ieee.org/discussions/discuss_feb2000.html.

39. IMS GLC Content Packaging Specification. *IMS Global Learning Consortium.* [Electronic] http://www.imsglobal.org/content/packaging/index.html.

40. IMS Learning Design Specification. *IMS Gloabal Learning Consortium.* [Electronic] http://www.imsglobal.org/learningdesign/.

41. IMS GLC Simple Sequensing Specification. *IMS Global Learning Consortium.* [Electronic] http://www.imsglobal.org/simplesequencing/index.html.

42. IMS GLC Question and Test Interoperability Specification. *IMS Global Learning Consortium.* [Electronic] http://www.imsglobal.org/question/.

43. **LTSC, IEEE.** IEEE standard for learning technology-learning technology systems architecture (LTSA). *IEEE Learning Technology Standards Committee (LTSC) .* [Electronic] 2003. http://ieeexplore.ieee.org/servlet/opac?punumber=8897.

44. IEEE Standard for Learning Technology - Data Model for Content to Learning Management System Communication. *IEEE Learning Technology Standards Committee (LTSC) .* [Electronic] 2005. http://ieeexplore.ieee.org/servlet/opac?punumber=9661.

45. IEEE Standard for Learning Technology - ECMAScript application programming interface for content to runtime services communication. *IEEE Learning Technology Standards Committee (LTSC) .* [Electronic] 2003. http://ieeexplore.ieee.org/servlet/opac?punumber=8972.

46. **Barker, Phil.** *What is IEEE Learning Object Metadata / IMS Learning Resource Metadata?* s.l. : Centre For Educational Technology Interoperability Standards, JISC, 2005.

47. **LTSC, IEEE.** IEEE Standard for Learning Technology-Extensible Markup Language (XML) Schema Definition Language Binding for Learning Object Metadata. *IEEE Learning Technology Standards Committee (LTSC) .* [Electronic] 2005. http://ieeexplore.ieee.org/servlet/opac?punumber=10263.

48. **AICC.** CBT Package Exchange Notification Service (PENS). *Aviation Industry CBT Committee.* [Electronic] 28 June 2005. http://www.aicc.org/docs/AGRs/agr011v1.pdf.

49. ADLnet - Advanced Distributed Learning. [Electronic] 1997. http://www.adlnet.gov.

50. **Jesukiewicz, Paul.** *Sharable Content Object Reference Model (SCORM®) 2004 4th Edition Content Aggregation Model (CAM) Version 1.1.* s.l. : Advanced Distributed Learning (ADL), 2009.

51. *Sharable Content Object Reference Model (SCORM®) 2004 4th Edition Run-Time Environment.* s.l. : Advanced Distributed Learning (ADL), 2009.

52. *Sharable Content Object Reference Model (SCORM®) 2004 4th Edition Sequencing and Navigation.* s.l. : Advanced Distributed Learning (ADL), 2009.

53. *Automatic evaluation of metadata quality in digital repositories.* **Ochoa, Xavier and Duval, Erik.** 2-3, Berlin / Heidelberg : Springer, 2009, International Journal on Digital Libraries, Vol. 10, p. 67-91. 1432-5012 (Print) 1432-1300 (Online).

54. **Hughes, Baden.** Metadata Quality Evaluation: Experience from the Open Language Archives Community. *Digital Libraries: International Collaboration and Cross-Fertilization.* Berlin / Heidelberg : Springer, 2005.

55. **Vidal, Philippe, et al.** Learning Objects Interoperability: The Ariadne Experience. *Building the Information Society.* Boston : Springer, 2004, p. 551-556.

56. **Arapi, Polyxeni, et al.** A Framework and an Architecture for Supporting Interoperability Between Digital Libraries and eLearning Applications. *Digital Libraries: Research and Development .* Berlin / Heidelberg : Springer, 2007, p. 137-146.

57. *The EduSource Communication Language: Implementing Open Network for Learning Repositories and Services.* **Hatala, Marek, et al.** s.l. : ACM, 2004. ACM Symposium on Applied Computing. p. 957-962.

58. *Smart Space for Learning: A Mediation Infrastructure for Learning Services.* **Simon, B., et al.** Budapest : s.n., 2003. Twelfth International Conference on World Wide Web.

59. *Integrating Digital Libraries into Learning Environments: The LEBONED Approach.* **Oldenettel, Frank, Malachinski, Michael και Reil, Dennis.** Los Alamitos, CA, USA : IEEE Computer Society, 2003, Vol. 0. 0-7695-1939-3.

60. Moodle.org: open-source community-based tools for learning . [Electronic] Moodle community. http://moodle.org/.

61. 3rd CSF - Istros. *Library & Information Center - Univeristy of Crete.* [Electronic] 2007. http://www.lib.uoc.gr/info/projects/3istros.

62. **Karmazi, Mariana.** *A SCORM conformant Sequencing Engine based on WS-Resource Framework and Principles of Service Oriented Architecture.* Heraklion, Crete : University of Crete, 2009.

63. UDDI: Online community for the Universal Description, Discovery, and Integration. [Electronic] OASIS. http://uddi.xml.org/.

64. **Connolly, Dan.** Overview of SGML Resources . *World Wide Web Consortium (W3C).* [Electronic] November 1995. http://www.w3.org/MarkUp/SGML/.

65. **Evans, Peter.** Z39.50: Part 2 - Technical Details. *Biblio Tech - Information Technology for Libraries - Technical Briefings.* [Electronic] 12 April 2001. http://www.biblio-tech.com/html/z39_50_part_2.html.