# A Learning Object Model For Electronic Learning

Shankar Mahadevan

Thesis submitted to the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University.

Master of Science
in
Electrical Engineering (Communications)

Dr. Saifur Rahman, Chair
Dr. Ing-Ray Chen
Dr. Annamalai Annamalai Jr.

June 28, 2002

Alexandria, Virginia

Keywords: E-learning, Digital library, Learning Management System, Distance Learning

# A Learning Object Model For Electronic Learning

Shankar Mahadevan

## ABSTRACT

Digital libraries are fast expanding into the role of independent educational entities that aspire not only to complementing traditional classroom teaching, but also allow open electronic learning for distance and continued education. These multifaceted roles can be realized only if the course content and the related content management system are versatile enough to be captured into any individual's learning needs. Many studies have defined a concept of "learning object" to address the issues and needs. But in attempting to solve the problem, the definitions have emphasized some aspects of the digital library while leaving the other issues to be solved later. Thus, the whole system dynamics is either weak or too cumbersome to navigate. As a part of this masters work, firstly the current model of pedagogical endowment was investigated. In order to accommodate the digital nature of education, a new modern profile of learning is proposed that allows modular yet efficient transfer of knowledge from the teacher to the pupil. The thesis then proposes a comprehensive learning object (LO) model, along with the associated system model, that will allow complete and flexible integration of content into the modern digital library profile. The process will be user-centric (both for knowledge developers and learners) as well as metadata-centric. It is scalable and interoperable with legacy and existing content databases and display systems. This thesis covers how the LO model is integrated into the core of the library's content development, discovery, and delivery process. The results of the experiment in terms of ease-of-use, flow-control, and feasibility of the model are documented. A beta-version of these concepts has been successfully tested with volunteers and implemented as a part of the Digital Library Network for Engineering and Technology (DLNET) project.

Shankar Mahadevan

# ACKNOWLEDGEMENT

I would like to use this opportunity to express my sincerely thank to Professor Saifur Rahman for his assistance and constant motivation in guiding me throughout this research. I am very grateful to him for providing me with the facilities that made this work possible. I also want to thank Dr. Annamalai Annamalai Jr. and Dr. Ing-Ray Chen for their support as my advisory committee. Their timely and constructive advice was critical in completing this thesis document.

Special thanks is also due to my fellow team members of the Digital Library Network for Engineering and Technology (DLNET) – Vara Prashanth Pushpagiri and Yonael Teklu; for providing helpful hints and continuously reviewing the software code. They are indeed true ‘band of brothers’ for this project.

I owe a sincere thank you to the many people who have both directly and indirectly made this work possible. Additionally, I would also like to thank George Hagerman of Alexandria Research Institute (ARI) for providing me with some very creative ideas during the course of this research. I gratefully acknowledge the support of National Science Foundation; grant number DUE-0085849, towards this research.

Last but not the least, thank you to my parents and friends who have always supported all my endeavors with maximum enthusiasm.

## Table of Content

# Table of Figures

# Table of Tables

Chapter 1

# Introduction

Advances in Internet technologies have changed many activities common in life. Learning is one such activity that is currently been revolutionized by access and use of this medium. Classrooms are no longer confined within the walls of a physical building, but are virtual and digitally omnipresent. By almost any measure, this form of online, distance and continuous learning, is fast gaining popularity among learners of all ages and interests. Even the parishioners of knowledge – the educators, have found digital learning medium to be a convenient and viable means of pedagogy. It is the convergence of the web and learning on all levels, whether it is elementary school, college or business [1]. The demand of online learning, in US alone, is expected to rise from just 5 percent of all students in higher education in 1998 to 15 percent by the end of 2002 [2].

An institution that allows pedagogy through electronic means via a "large" collection of information is called a ***Digital Library*** (DL). Digital libraries are fast expanding into the role of independent educational entities that aspire not only to complement traditional classroom teaching, but also allow open electronic learning for education on demand. Inherently this transforms learning management system from "teacher-centric" to "learner-centric" pedagogical endowment. Using digital libraries, education can be attained in multiple ways. Some pupils may prefer to go through a fixed set of digitized learning material for self-paced learning. Others may prefer real-time interactive education with the educator and the learner brought together in the virtual medium. Yet others may prefer asynchronous learning with educators and learners collaborating together, but not necessarily online at the same time. These forms of learning, collectively defined as electronic learning (***e-learning*** in short), are central to the activities of digital library. These multifaceted roles can be realized only if the course content and the related content management system are versatile enough to be captured into any individual's learning needs.

The ability to instruct from afar is hardly new. As noted in IEEE Spectrum (October 2000 issue [2]), since the mid-1800s, correspondence schools in Europe were teaching shorthand and foreign languages by mail. In the last century, radio, television, and satellite broadcasting technologies have provided new medium for delivery of distance learning. The global connectivity of the Internet and a new generation of hardware and software applications underpin the teaching of courses over the Web [2]. Since computers offer so much flexibility and variety, so does e-learning.

## 1.1 The Concept of Digital Library

The basic functionalities provided by any DL are somewhat parallel to that of traditional libraries, but it also encompasses the role of educational institutions. Books or resources, as in a traditional library, are collected and upon screening and cataloging, are finally made available to its patrons. However the similarities end there. The architecture of digital libraries to support such activities is quite different from those of traditional libraries. The brick-and-mortar buildings are replaced with databases, e-learning tools, content servers, web-portals, etc. The books, in this case are the digital resources and constructs. The common denominator among them is their ability to provide knowledge or education.

### 1.1.1 THE "BOOKS" - THE RESOURCES

Since the DL resides in the virtual medium, the learning programs are maintained in digital format. In almost all the field of arts, sciences, engineering, technology and others, the knowledge providing content come in varying sizes and formats and sometimes in mixed formats. Moreover the material can be presented as text, graphics, animated graphics, audio, video, or a combination of these. It is hence essential to follow a standardized set of rules and methods that allow collecting and processing this content. The concept of "data object" or "learning module" encapsulates these requirements. In general, it is defined as a digital building block that has pedagogy for its target audience [4]. They can be exercises, lessons, lectures, simulation codes, etc. This structure allows a high degree of configurability and reusability [4].

Since this is basically a source of information, patrons of the library should be easily able to browse, collect, reference and reuse (with permission) these data modules. But unlike a physical book, where the patron has a pre-conceived common knowledge (about how a book looks like, is it written in the language known to the patron and what is the starting point of the book) and structured guide (the table of content) to help use the information, the digital content rarely comes with such "pointers". Unless it is a popular form of content or the patron is an expert of that field and has knowledge of the tools associated with the digital construct, it is difficult even to begin viewing, let alone assimilate the information within it. It is hence essential to sculpt the learning modules with the fact in mind that the user wishing to access these objects possesses varying capabilities and characteristics. Furthermore, some users may seek varying degree of learning experience. Some may prefer a longer and in depth study as opposed to others who may be looking for introductory or refresher information. The "learning module" should be able to accommodate these different needs. Thus these "books" should be able to exist as small chunks of content as well as chunks that can be placed within instructionally functioning contexts to facilitate learning [4].

Being in digital format, users have come to expect or demand, many other electronic qualities for the book. The ability to bookmark, reference, store, use/re-use (with permission) chunks of it, pay-per-use, "play" in user's custom environment, have a supporting system that can be invoked to track and get credits for knowledge gained from the book etc should be integral part of the model. The learning module or the book model, in which educational content is stored in small chunks that are reusable in many different learning environments, has been proposed to support all these multifaceted requirements and has gained broad acceptance [4]. It should thus encapsulate the "learning" and "digitization" aspect of the physical resources.

### 1.1.2 THE CATALOGS, CUE CARDS AND LIBRARY CARDS

In a traditional library, the cue cards hold information about a book in order for easy discovery and locating. Moreover the cue cards in themselves are arranged in some intuitive order for convenience of browsing and subsequent retrieval. It's electronic counterpart – the metadata, performs a similar function in digital libraries. ***Metadata*** is defined as data about data. It holds information that will assist a user in quickly locating the learning modules of interest. A limited list of elements of interest to any audience is given in Table 1.1. The first row of metadata pertains to "ownership" or "copyright" aspect of the information. This is generic information like the author(s)'s or publisher's name, affiliations and contact information. Besides providing inherent copyright value, this kind of data can also be used to identify the content. More common identifiers of information are the title, description, classification and related items as listed in the second row of the table. Since the content in a digital library can be of various forms, the *format* metadata element is further useful to identify the relevant learning modules for the user. The accuracy and richness of this kind of data is highly valued as it is the most commonly sort identifier used by the patrons to browse or select a content for further investigation. This information hence gives credibility and provides quality assurances to the content within the DL.

**TABLE 1.1: Metadata Elements of Interest for Cataloging**

| Category | Metadata Element |
|---|---|
| Information about the Contributor | • Author Name<br>• Email Address<br>• Department<br>• Organization |
| Basic Metadata | • Title<br>• Language<br>• Description<br>• Keywords<br>• Date of Creation<br>• Format<br>• Classification |

Since each learning module has a "cue card" associated with it, one of the simplest ways to arrange and catalog the cards is through the classification element. Classification element utilizes the fact that different disciplines of study can be divided into many subject areas, which can be further branched into associated sub-topics. This in essence uses taxonomical arrangement of the learning modules. In a digital library users may browse through their specific areas of interest, to see a comprehensive set of learning material available in that sub-topics. Cataloging (using cue cards) bridges the learning modules with the underlying support architecture in the digital library. Additionally, keywords and mixed phrases can also be used to seek and isolate the content of interest. The information within the metadata reflects on the quality of the resources and hence the credibility of the library. It is hence a very essential component within a digital library.

Besides cataloging and “cue cards”, the profile of the library subscriber is another essential component of indexing. In a traditional library, this helps in tasks of checking the books issued, renewed or overdue with the patron. But digital library are more than just a collection of learning modules, they seek to impart knowledge. Hence, the user profile has to store additional information like the correlation between the various types of “books” downloaded, test and track knowledge (if any) gained via these modules, inform or initiate actions based percepts either from a teacher or history of user’s action, maintain security and privacy, etc. In a commercial setting, they can be further used for billing such as pay-per-use or to track royalty for authors etc. Thus, in digital libraries, the user’s profile or “library card” is far more critical and central in providing a “learner-centric” pedagogy.

### 1.1.3 THE LIBRARY

Digital libraries similar to traditional libraries, undertake the tasks of collection, screening, packaging and archiving of educational content for easy development, discovery, delivery, use and reuse. Collections are developed by acquiring contents through active harvesting and soliciting authors for submissions. Authors are the academic or professional experts in their respective field of study. Since, as mentioned before, the learning modules are available in heterogeneous formats, they have to be packaged before submission. There are many content packaging methodologies available for bundling resource for e-learning. The packaging has to bind the resources to the metadata and also perform validity and integrity check of the resource files. This saves costly error checking and negotiation time between the author and the library. Packaging will convert the submission to pre-defined and standardized package. This allows smooth flow-control and accessibility across the library network. After submission, contents are screened for their quality and pedagogical value using peer reviewers. Archiving will involve the process of indexing/cataloging the metadata about the content and placing it in the search repository. This is now ready for discovery and use by patrons (such as the practicing professionals, other educators/course-builders, students). Most of these processes can be automated by smart coding techniques. Limited manual intervention may be required from time to time.

To realize the above different components of a digital library in the virtual medium, the hardware and software technologies need to be robust, inter-operable and scalable at all times.

## 1.2 Technology Requirements of Digital Library

People often interchange e-learning with the terms computer-based training, web-based training, computer-based instruction, and technology-based instruction. Actually, each of these is a form of e-learning, and the terms have a specific meaning [5].

- Computer-based training (also called computer-based instruction) refers to courses presented on a computer. The course does not provide links to learning resources outside of the course. Often, learners take a computer-based training course on a computer that is not connected to a network.
- Web-based training is a form of computer-based training; it refers to courses available on an intranet, extranet, or Internet and that are linked to learning resources outside of the course, such as references, electronic mail and discussions, and videoconferencing.
- Technology-based instruction has a broader meaning; it refers to training through media other than the classroom. That includes computers, but also refers to television, audiotape, videotape, and print.

Terminology and concepts become confused as people equate delivery mechanisms with program structure and focus on technical issues, sometimes to the oversight of content. Currently many technologies attempt to address one or the other need of a digital library administrator. These technologies can be broadly classified under the following two categories [5]:

- Technologies for developing learning programs (the “books” and “cue cards”)
- Technologies for handling and delivering learning programs (the “library” architecture)

The technologies for developing the learning program are “many-to-one” type – implying that the user of the technology should have the ability to form a single coherent collection out of many different chunks of resources having heterogeneous educational value. They encompass the hardware and software needed to create the learning materials, prepare them for "publication," and administer the use of the materials once they are published. The knowledge or the content developers, teachers and subject experts are primary users of these technologies. They use it to create and package the raw content into "bundles" that allow efficient and streamlined use of the educational material between different parts of the library.

The technologies for handling and delivering the learning programs are "one-to-many" type – implying that the users of these technologies should have the ability to dissipate knowledge possessing modules in a manner that enable anyone requesting them to gain pedagogy via the material. They can be further grouped as follows:

(a) The technologies that include the computer hardware (personal computers, personal digital assistant, etc) and the software installed on it (such as the operating system that controls the computer, client side of the learner management system) that lets learners manage and use the learning materials and

(b) The technologies with specialized hardware and software for discovery, storage and archiving of the learning material. They may include specialized databases, high-end multimedia servers (for playing video and audio), and related technologies for handling complex interactions between the student and the library. These technologies may also include the cables and software that connect the computers in a network, the protocols and standards that help users accurately read information transmitted on the network, and specialized software, such as client-server-database programming languages. These technologies are of special interest to the administrator the digital library.

Above all, the success of a digital library is directly proportional to the ability of its patrons and administrator to successfully use and navigate these technologies. Hence it is critical that they be user-friendly and efficient at all points of access.

## 1.3 Research Objective

Currently, due to the growth of available bandwidth and efficiencies introduced by Internet activities, the technologies for handling and delivering learning programs has sufficiently matured. Search engines such as Google [6] are very efficient in discovering learning materials.

The high-speed servers and databases are adequately serving web-based contents. But there's a big piece missing, according to Steven Gnagni [7]: "While it's getting easier and easier to find learning material, it's still fairly tricky to get them to run on whatever course platform you're using, or to have a consistent methodology to track (or transfer) the learning usefulness (results) via existing learning management system (LMS). There may be plenty of valuable ready-to-use parts out there, but they're still too hard to find and to adapt and assemble into new, useful products."

It is becoming clear that, technologies for developing, i.e. "bundling" resources with metadata, learning programs are lagging behind and are hence inhabiting the realization of the full potential of digital libraries. The standards and specification of these tools are currently in their infancy. Even standardizing bodies, which aim to bring about consistency in the implementation and the use of these technologies are many in number, restrictive (in some cases proprietary) and non-interoperable [2,3,7].

The research effort presented in this thesis aims to address this issue of standardization of development of learning material and other relevant issues. More specifically it will target the following:

> *Develop a cohesive and comprehensive "learning object" model, alternately called the "data module" or "learning module". This refers to the chunks of resources that have educational value. Important issues that need solution here are how to meta-tag (identify and link) and "package" these chunks; how to ensure the "packaging" is scalable yet malleable to be easily handled across various parts of the digital library; and how to make use/reuse of these chucks efficient while maintain copyright considerations.*

It has become clear to the author that while addressing the above issue, many known standards do exist that attempt to answer some of these issues. To avoid "re-inventing the wheel", part of the effort is directed to understanding the motivation and methodologies behind these solutions. A fair comparison is drawn and the closest and most efficient standard will be chosen towards finding the optimum solution for the learning object. Extensions will be made as needed.

It is also hypothesized that a "tool of sorts" would be needed to demonstrate how these ideas would work in "real life". Currently it is envisioned that such a tool will provide functionality to parse the resources, seek metadata and finally assist authors in "bundling" the resources into a learning object. Additionally a minimalist implementation of the digital library itself would be required to show how this definition of the learning object would pass through many processes from the time it is created by the author to the time it is archived and finally searched, retrieved and used by users. In this process the current and required model of digital libraries is also investigated.

In the longer run, it is hoped that, these ideas will help to demonstrate that standardizing the development of learning programs is essential for easy across-the-board acceptance of digital libraries as a credible and qualitative source of education.

## 1.4 Thesis Outline

The thesis begins first by investigating the current structure of pedagogical endowment (Chapter 2). Educational process has traditionally been done via school and university environment. However in the digital world these structures are no longer needed. Teaching can be done virtually. A model that allows such education – called the modern profile of pedagogical endowment, is thus proposed. This model represents the digital library. A practical scenario is then considered where the functionalities provided by this model are illustrated.

Once a clear model representing education is available, it is now useful to see how learning programs would interface with such a system. Though many concepts do exists that propose integrating learning programs into digital libraries, this document (in Chapter 3) proposes a new and updated concept of "learning object" to tackle some of the known and outstanding issues of learning program integration. The major concerns here include metadata standards selection and compilation, resources checking and invocation, file structure implementations, etc. Additionally, special effort is made to study and improve the LO definition for allowing scalability and intelligence to the model.

Following this point (Chapter 4) discuss a LO development tool. This tool creates LO from raw resources. It provides core functionalities of metadata collection, resource compilation and packaging. Architecture and implementation of the tool is done with the view to maximize user experience for minimum effort, especially for first time users. The GUI layout and usage for the tool is also discussed.

Chapter 5 deals with the handling and distribution of LO at the server-side of the digital library. Many process such as content checking, review and packaging are integral functions provided by the library server. Apart from LO services, it also handles user services like registration of patrons, profile maintenance, etc. These are discussed in detail in this chapter. The LO and the associated server represents the microcosm of any digital library.

Many government and private efforts are currently underway that what to merge these digital libraries into one big entity via the Internet. These initiatives are discussed in Chapter 6. The most notable among them is one by NSDL which aims to provide enhanced educational experience via the digital medium to all Americans. This could possibly the future of digital library – an interconnected network of educational resources available at will. The final chapter (Chapter 7) summarized the over all task accomplished under this thesis work.

## 1.5 Digital Library Network for Engineering and Technology (DLNET)

There have been many attempts to develop successful digital libraries, some led by government funded agencies and others by private/commercial efforts. These libraries can be though of as a large collection of distributed, heterogeneous information managed by autonomous sources [10]. The Digital Library Network for Engineering and Technology (DLNET) is one such web-based library portal intended to complement engineering and technology education. The objective of DLNET is to develop a specialized collection of engineering and technology-related content targeted at the practicing engineer and technologist so as to facilitate “lifelong learning,” i.e., education beyond the classroom using digital libraries. It is part of the National Science Foundation (NSF) under its Science, Technology, Engineering, and Mathematics (STEM) Education [11] initiative of establishing a digital library at the national level that will constitute an online network of learning environments and resources. As a result of such initiatives and other efforts, digital libraries are fast becoming reality.

This thesis is derived from the work on DLNET. Many of the solutions have been tested and are publicized via the DLNET. DLNET is hosted at Alexandria Research Institute (ARI), Alexandria, VA – a part of Virginia Tech. The virtual web address is http://www.dlnet.vt.edu/. It is a three-member team effort.

## 1.6 Thesis Contribution

The author's (the author of this thesis) contribution has been in many areas towards realization of DLNET. Specifically, it has been in areas relating to the concept of learning object definition, development of related tool and LO management in the DLNET server. Though there exist many concepts for learning modules, they are limited in one or many ways from being fully implemented in a generic digital library. This thesis has come up with a unique model that not only encompasses functionalities provided by previous models but has added features that make it unique, flexible and yet scalable for use in wide ranging scenarios. This model facilitates versatile use and re-use of learning programs in any setting by both the teachers and the learners. Special aspects include the LO file structure and the concept of LO preamble (to be discussed in Section 3.4) – these are unique contributions of this thesis work and it's author. It conforms to known standards of learning management and is yet easily "playable" using popular web browsers. These aspects are neither application specific nor are they operating system specific, hence maximizing interoperability among different systems across the digital library and the Internet. Additionally some special aspects such as nested LO and adding intelligence to the LO model have been devised by the author. These ideas are new to LO model and do not exist in other third party LO concepts. The overall concept is designed as a basic building block for universal education. Thus the proposed learning object model can be implemented as a part of DLNET or any other digital library or as a stand-alone object.

Besides the LO definition and model, a JAVA based tool, called the LO Development Tool, was created by the author. The tool allows physical realization of the LO model into a zip file. The architecture and coding of the tool is author's contribution for client-side system for DLNET. The tool incorporates a unique and innovative algorithm to perform resource validation and bundling. Volunteers have successfully tested the tool for its robustness and user-friendly attributes. The tool is available via the DLNET website.

For the DLNET server-side architecture, the author's efforts have assisted in creating a roadmap, including implementation strategies. Though the author did not perform any server coding, he performed test and integration duties.

Chapter 2

# Evolution of Digital Libraries

*"GOOD MORNING AND WELCOME to the classroom of the future. It's the second day of Calculus 101 at a four-year public university in the South, and the professor is teaching the students how to differentiate. She explains the concept and then runs through a few examples on the whiteboard: $3x^2$ becomes 6x, $4x^4$ turns into $16x^3$, and so on. After she's finished, a quiz developed by a CalTech professor appears on a screen that's embedded in each student's desk. Five minutes later, the results are in: All but a single student, Johnny, aced the quiz.*

*While the rest of the class moves on, Johnny gets a little bit of extra help. According to his student profile—stored on a central university computer—he learns best visually. Armed with this information, the class computer searches the Internet for an animated lesson about differentiation and delivers it to Johnny's desk screen. He watches the animation, takes another quiz, and aces it."*

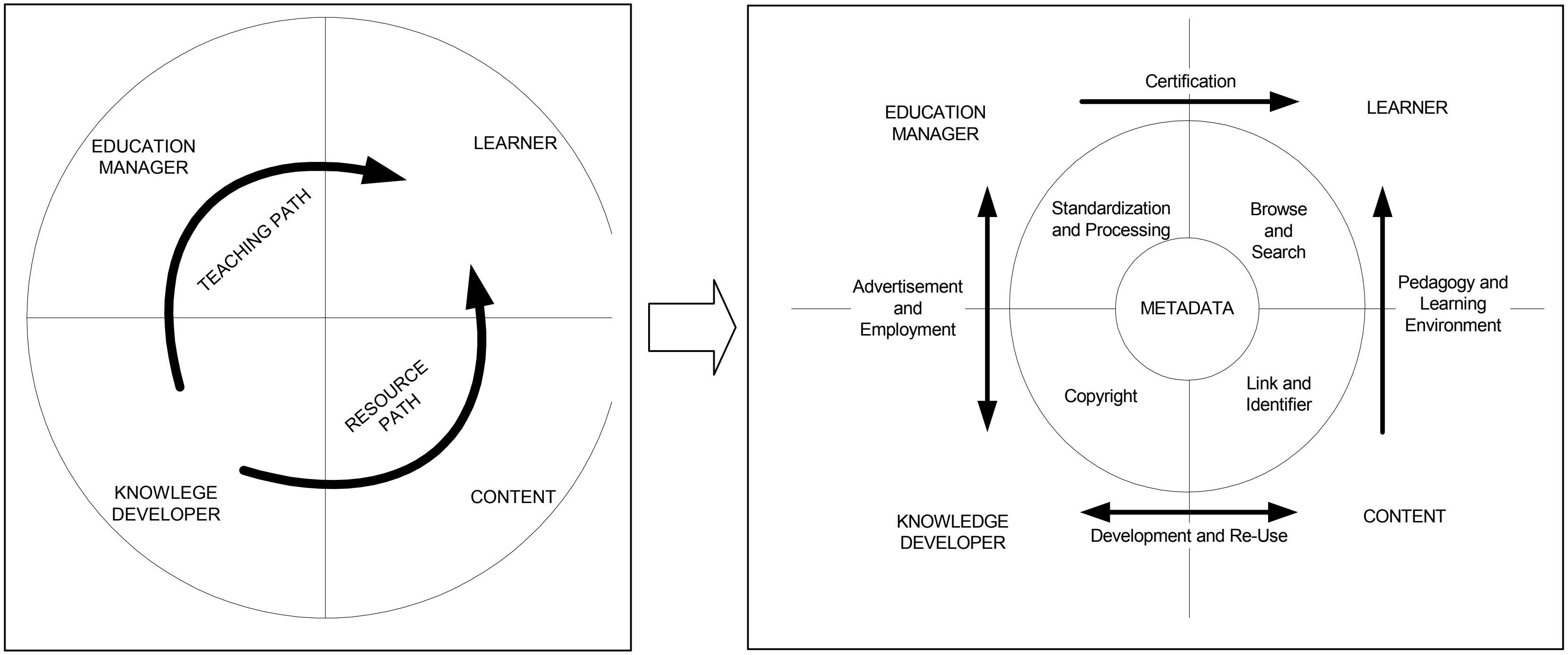
Above is an excerpt (quoted verbatim) from an article written by Steven Gnagni [7]. Though technologies to support such futuristic classrooms are fast becoming a reality, it is important to realize that a fundamental change is occurring in the way pedagogy is endowed from the teacher to the pupil. Hence, before one attempts to solve problem affecting digital libraries, it is essential to evaluate the nature of pedagogy within them. Figure 2.1 shows how pedagogy has evolved along with the digital age. This model shows how learning (and resources associated with it) can be transformed from teacher-centric to learner-centric education, and how content can be transformed from multimedia resources (application or technology specific data) to metadata-based learning modules.

## 2.1 Traditional Model

The traditional model, Figure 2.1(a), has been the centerpiece of imparting education up until the last century. It is a closed model with little room for flexibility and enhancement. Schools and universities, which represent education managers here, structure curricula and standards for judging the level of learning. Here a teacher (the knowledge developer) creates relevant educational resources (the content) to provide step-by-step learning instructions in their respective field of expertise. The students (the learners) use these resources that fit the curriculum, of their desired field of study, to gain education. Thus, there are two distinct paths to pedagogy: the learning path (attending designated standardized in-class lectures etc) and the resource path (via recommended books, class hand-outs, home-works, etc) from the educator to the learner.

## 2.2 Modern Profile

Recently, a new dynamics to gain pedagogy has evolved. The overall architecture representing it is shown in Figure 2.1(b). It is an open system, which represents the objective modularity of the underlying digital nature of relationship between various parties involved in education. At the core is Metadata, which is defined as data about data. It enables smooth flow of control among different entities, actions and functionalities provided by the digital library. Among others, the three main types of Metadata are: content-centric, user-centric, and pedagogy-centric metadata. Content-centric metadata deals with content description and packaging and is used while searching for the content. User-centric metadata deals with the creation and maintenance of user profiles. Pedagogy-centric metadata deals with evaluation of learning and result tracking. The value of elements in the metadata is used to initiate actions among various entities within a digital library. The interaction and inter-use of metadata elements between these entities is complex and often overlapping, while the functionality provided by the metadata elements to each entity is unique [3,5,7,13,14].




**(a) Traditional Model** **(b) Modern Profile**

**FIGURE 2.1: Evolution of Pedagogy Endowment via Digital Libraries**

The education manager quadrant encompasses the educational institutions (like schools, universities, technical colleges etc), as well as new breed of standardizing bodies (like IEEE, ASEE, NSDL, etc.) and commercial for-profit educational service industries (like WebCT, Blackboard etc.). Despite their seemingly different job descriptions, all these groups have the same primary objective, i.e. to facilitate learning from the teacher to the student. To ensure seamless transfer of knowledge, the corresponding groups develop, maintain, use and track standards for metadata (like content description, user profile, general statistics etc) implementation across the digital library. These standards are designed to be scalable and inter-operable between different digital libraries as well.

The knowledge developer quadrant encompasses professors, teachers, subject experts, and alike – those who generate content having educational value. Additionally, in the role of a reviewer, they determine the learning value and quality of the contents. Besides this traditional role, their participation in the digital library provides them with added value of copyright and recognition for the content generated by them. The author information and copyright is preserved via the metadata association of the resource. It is imperative that they understand and use metadata standards (especially those relating to content-centric development and use) developed by education managers. This allows easy identification and use of the content among various groups thus maintaining consistency and smoother flow-control across the digital library.

The content quadrant represents the physical resources that contain educational information. These resources can be of any size and format – as long as it can be electronically handled. The resource ownership, description, grouping and association are encoded in its identifying metadata file. Without identifying metadata, though the pedagogical value of the resource may not be lost, the mere fact that it has pedagogical value would be difficult to estimate. Thus, its use across the digital library is in jeopardy.

The last quadrant, the learner, is the ultimate end-user for all pedagogical service activities. They use the metadata to seek content. It may involve many alternate methods, including: basic keyword search, advanced Boolean search, browsing through taxonomy, etc. Without the metadata, any information about the resource or its pedagogy is hard to find. In digital medium,

metadata also helps to identify learner-centric resources. Some advanced systems, when allowed, can filter/seek learner-centric modules based on the user-profile metadata.
It is clear that the interaction between various quadrants occurs via actions invoked by values in the metadata elements. As long as each quadrant can access the desired metadata, they can exist mutually independent of each other. For example, universities do not necessarily need to maintain content banks. It may be more efficient for a third-party entity (possibly a commercial database company) to maintain such content banks. What is required, is that the third party implement the metadata standard set by the university, when accepting content from authors (possibly pay royalty in return), following which it should share (possibly for a fee) with the affiliated university. The university should then, based on the quality of the metadata, set requirements of certification. The learners, may at anytime search the metadata repository and then download (possibly for a fee) the chosen type of content and gain education at the desired pace of study. The type and fashion of lessons studied by the learner is stored in his/her user-profile. If the user so desires, then he/she may share this profile with the university. Following which, the university, when satisfied of the fulfillment of the course requirement, may provide the desired certification (at a fee), to the learner, for the acquired knowledge.

Let us now revisit the classroom scenario (the quote from Steven Gnagni [7]) given at the beginning of this chapter and see how it fits the modern profile of pedagogical endowment. The professors (in the classroom, creator of “animated calculus lesson” and at CalTech) represent the knowledge developer group. Johnny and his classmates are the learners. The education mangers group includes everyone else – the system administrator controlling the “central university computer”, the database where the results are stored and related entities. The quiz and animated lessons etc form the content quadrant. The central figure – the metadata, is information about the content (the fact that it is an animation) and pupils like Johnny (the fact that he learns visually).

The knowledge developers interact with the educational managers via employment or advertisement. The classroom teacher is directly employed. The author of the “animated calculus lesson” and the CalTech professor can be viewed as indirectly contracted to develop tests (a form of content) for certification. All of them have developed content and are independently able to impart and test pedagogy of the learners. The teacher’s content is developed and distributed in a traditional form of pedagogy, in essence in real-time. The animated content, on the other hand, is stored in a central repository and which is access based on the metadata match between the testing requirements of the user and the property of the content. The important underlining theme being that learner (Johnny) gains education irrespective (or prior knowledge) of the source of the content. The educational manger eventually tests the learner and the results of the test are accordingly used to judge and certify the education gained by the learner.

The central role in all these activities has be the streamlining of the content-metadata packaging and use. The following section defines an application independent definition of a "learning object" that best fits the requirements of the modern profile of a digital library.

Chapter 3

# The Learning Object (LO) Model

The modern pedagogical profile requires a strong, re-configurable and re-usable definition that clubs together content and metadata. It is important to bundle the learning modules in a known and prescribed fashion to allow the convenience of its flow between various quadrants of the modern pedagogical model. This in turn helps improve its development, discovery and delivery process. In this chapter a concept of "*learning object*" (LO) is proposed to describe such learning modules. They are analogous to "books", as described in Section 1.1.1, in a traditional library.

As mentioned earlier, technical content comes in varying sizes and formats and sometimes in mixed formats. Moreover as new applications are developed, where is no way of predicting the resource or collection there of, that forms valuable educational content. The material can be presented as text, graphics, animated graphics (that is, graphics that move), audio, video, or a combination of these. Besides content creation, there is a need to judge its "quality" relevant to learner's needs and other similar contents. This is possible only by having rich information on the content creator and the content itself. It further helps in easy discovery and use of the content. Hence there is a need for a uniform protocol that allows organizing resources with emphasis on "quality" and "learning". All this is achieved with minimal human interaction with the content, so as to speed-up the creation to delivery process. The author has proposed an object model protocol for processing the learning module. It is the ***learning objects*** and is defined as follows.

## 3.1 Definition

*A learning object is defined as a structured electronic resource that encapsulates high quality information in a manner that facilitates pedagogy. It has a stated objective and audience. It has ownership and associated intellectual property rights.*

The definition highlights two aspects namely, “learning” and “object” with the underlying theme being “ownership” and “quality”. It is immutable in the sense that it may not be changed or altered by the user. It must function independently (self-contained and standalone). Also note a LO may have intelligence and additionally, have other LOs embedded in it (to be discussed in later Sections).

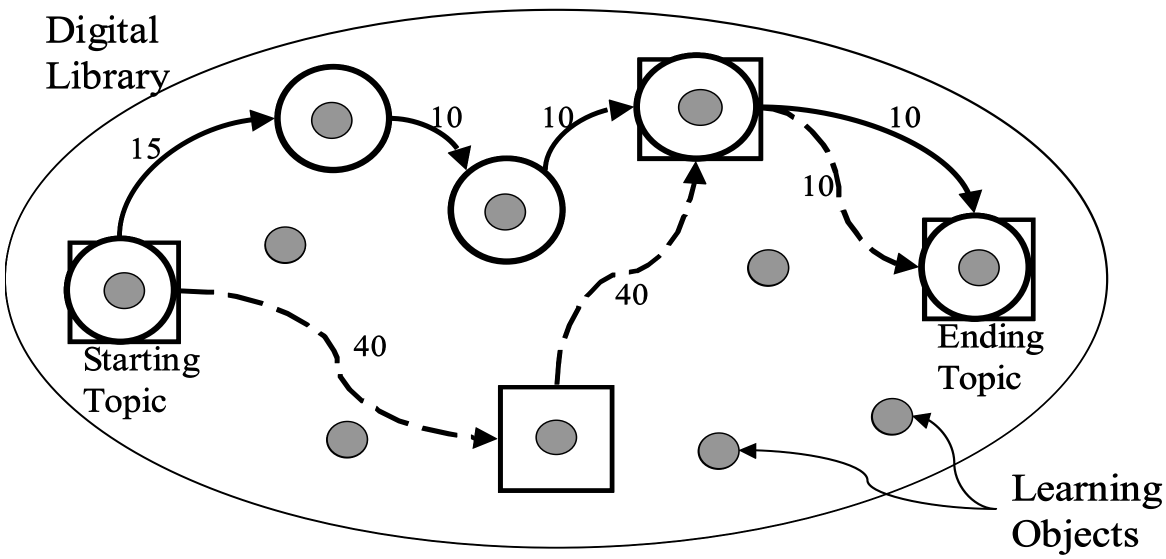
### 3.1.1 LEARNING

The “learning” aspect of the LO definition is very important. Learning refers to education. Education is a process. The main aim of DLNET is to allow continuous education. Digital libraries are more than a simple content repository or a hosting platform, promoting continuous education and life-long learning. It will facilitate the acquisition, assessment and conversion of contents into LOs while fostering the assimilation of those LOs into learning modules and instruction. The purpose of the learning object is to hold information that enables learning. Using this model, e-learning tools have greater flexibility in tailoring courses based on student’s background. As all different quadrants of modern pedagogical profile act independently, the student has no pre-formatted set curricula to follow. This, while being a boon for continuing/electronic education, requires that learning choices be flexible and have successive pedagogical value.

Learning from a LO can be envisioned in two ways. The first approach is for the LO to be an independent entity i.e. as an individual topic. In this case, it is important that it be coherent, self-contained and rich in its pedagogical value. This kind of learning suits best to learners who do not have time to follow extensive course layouts or who are seeking a re-fresher to their topic of interest. Another way to gain learning is via following an ad-hoc path from one learning object to another object. In this case, the learning objects are not only self-contained for that subject but also assists in step-by-step pedagogy from a starting topic to an end topic of desired pedagogical goal.

The original concept was first proposed by Vincenze Carchiolo and others, and is discussed in detail in [10]. Here a database of learning modules is envisioned called “*one-argument lesson*”.

This structure allows for high degree of configurability and reusability. A modified version of the concept is adapted to explaining ad-hoc learning via proposed LOs. Figure 3.1 represents this concept (the figure is adopted from [10]). Each node represents a LO, connected with the other node through oriented and weighted arcs. The weights can represent difficulty of topics, “time to learn”, etc. The logical relation of ‘precedence and succession’ can either be fixed by the original content creator, or by third-party teachers/reviewers or by the learners themselves for a learner-centric experience. Thus, learning can be asynchronous, self-paced or structured.




| Path | Number of LOs | Avg. Time | Difficulty |
|---|---|---|---|
| Solid arrow | 5 | 4 | 45 |
| Dotted arrows | 4 | 4 | 90 |

**FIGURE 3.1: Routing Table to Reach the Target Topic using Ad-hoc LOs [10]**

The overall important criterion of modules, for education, is thus its pedagogical value. Assessment of the pedagogical value of content is a subjective procedure. Sufficient, if not extensive, information should be collected about the creator and the content, so as to enable a reviewer to judge its quality and learning effectiveness.

### 3.1.2 OBJECT

The "object" aspect of the LO definition relates to digital/electronic format of the resources. Simply put it the actual physical resource (the files) that forms the learning object. Since digital construct can be of any size and format, so are the objects within the LO. The digital constructs may encapsulate both information and behavior. Since it's the objects in LOs that have knowledge giving information, the act of accessing, transmitting and distributing them should be as simple for handling across the different spectrum of the modern learning profile. Digital libraries have to assist in streamlining the processing of these objects. Most of these activities occur via the Internet. All the objects composing an LO, must be "playable" via available popular browsers. The tools for developing and validating the resource objects have to be versatile and user-friendly and geared towards web-accessibility. Additionally, it must be capable of capturing high quality and quantitative metadata about the objects. Further the tool must package the object such that it facilitates fast and efficient delivery. Currently many approaches are at works for information delivery.

Traditionally, when an object is accessed, the entire object is sent to the user, disregarding the capabilities of the user's appliance. This unnecessary transmission of data results in network congestion and delays. Alternately, in a precompiled approach, several modalities of an object are generated apriority and stored. This approach has the advantage of sending only useful information. However, it requires that the author of the object stores every possible data format or type the user may want. Thus, it is not practical in a digital library environment. In other cases, only one copy of the object of a single modality is stored and is then converted to the user's requested modality and fidelity on demand. This approach focuses on the translation, but does not deal with mixed modalities, whereas multimedia presentations rely on the interplay of different modalities [8].

The proposed object model uses the combination of the above methods. The users can download one entire LO if desired. Alternately, they can preview the LO via many popular browsers and navigate to the desired section without downloading the entire LO. The essential criterion is that most of the objects, if not all, should be web-viewable, i.e., one can learn by opening and playing the objects in a web-browser.

Hence, the object definition and handling should be streamlined extensively and with view of flexibility and scalability in mind. The "objects" in LOs are analogous to objects used in object-oriented modeling (OOM). The analogy helps visualize how LOs will be packaged, processed and transported across the digital library as well as utilized in course building. OOM concepts such as encapsulation, classification, polymorphism, inheritance and reuse can be borrowed to describe the operations on LOs in the digital library. It may encapsulate other objects. Besides the size and format of the object, the organization of the resources is also important. This allows a modular approach to development of learning resources. Once the objects become part of the LO, they cannot be altered. An LO hence is an immutable entity.

### 3.1.3 OWNERSHIP AND "QUALITY" of LO

Ownership and quality are the requisite attributes of any resource within a library. Ownership relates to the intellectual property of the LO. Once an LO is created, the author is encouraged to submit the LO to DLNET. This has many advantages. It allows gaining recognition, feedback via peer-review and wide distribution of the content. The incentive of the author to create and submit future LO is to ensure that author is given proper royalties for their creation. In LOs ownership is hence essential. It is maintained by binding the metadata with the content. The metadata relating to ownership are listed in row 1 in Table 1.1.

The quality implies the accuracy and the pedagogical value of the learning object. This is of prime importance for digital libraries to be credible entities of knowledge. Following four facets represent quality in the context of DLNET.

- **Pedagogical effectiveness**: It is the measure of the education value of the LO. This is relevant for judging the target audience and learning duration for the given LO. An LO providing introductory lessons may be a short text and thus have low pedagogical effectiveness. At the same time it may be possible that, a short applet or a few minute simulations may provide high pedagogical value to its audience.

- **Subject matter accuracy**: It represents the relevance of the data in the resources to the said objective of the LO. This is the most vital match that reviewers have to watch for. LO development tools currently do not have enough intelligence to decipher or judge the relevance of the title with the content given in the resource. They must match to give high quality results when searching and browsing for the desired learning module in digital libraries.
- **Production value appropriateness**: It represents the effectiveness of putting together the LO for development, discovery and delivery. It also alludes to post-publication usefulness, scalability, and backward compatibilities. These parameters are important to sustain the quality of the LO for consistent use as, newer and better versions of the digital libraries are unfolded for backward compatibility.
- **Proper technical behavior**: It represents the presentation and technical "soundness" of the LO. It is the machine handling as well as the ability of the users to use the widely available software and hardware to work with the LO. It reinforces the ease-of-use of the LO.

Judging "quality" is subjective and is done via review process. One of the main aims of such review processes, besides making "quality" determination, is to judge the credibility of the creator. This in turn allows DLNET to become a "trusted source" of the content. Credibility of the creator is ranked based on the years in profession, organizational affiliation, other similar participations etc. It is thus, essential to collect as much information as possible on the creator and the corresponding content. It must comprehensively describe all the characteristics of the LO so as to enable quick discovery. This, at the same time, supporting know "natural discovery" process of the most simplistic searcher.

## 3.2 Characteristics of a LO

Apart from imparting knowledge, a learning object may exhibit some or all of the following characteristics (not necessary all and at the same time). These elements have been adopted from the IMS-IEEE LOM standard (v1.2) [20].

1. **Objective**

   The objective of the LO tells what the resource is trying to do. The authors use this field to state the purpose of their learning module. Every learning endeavor should have a set goal. This can be a single dedicated or multiple steps towards a cohesive concept.

2. **Pre-requisites**

   Pre-requisites implies the required background knowledge the person must have, in order to understand and use the LO. It is a guideline for prospective users to judge its relevance to their learning requirements. From a continuous education or lifelong learning perspective, it gives a measure of the educational level that learners should before attempting to use this LO.

3. **Interactivity Type and Level**

   Interactivity gives the nature of inter-action that occurs between the user and the LO. Most learning objects will have some form of interactivity levels associated with them. A learning activity using a learning object that involves simple web browsing can be considered a passive type of interaction with low to medium interactivity level. On the other hand, a complex LO, such as a simulation tool or an applet that involves feeding inputs and monitoring output/results etc, is an active type with high interactivity level. The following table shows the various ranges for interactivity types and levels.

**TABLE 3.1: Interactivity Type and Level**

| Interactivity Types | Interactivity Levels |
|---|---|
| Passive | Very low |
| Active | Low |
| Expositive | Medium |
| Mixed | High |
| Undefined | Very High |

4. **Learning Resource Type**

   Learning resource type defines the intended use or the context of use by the user. IMS-IEEE LOM [20] defines following types for a collection of resources, to be identified as LOs (one or more applicable):

- Exercises
- Simulations
- Questionnaires
- Diagrams, Figures or Graphs
- Indices
- Slides
- Tables
- Narrative Texts
- Exams
- Experiments
- Problem Statements
- Self Assessments

A course, defined as a LO, may be a collection of slides or a web-based narrative with text, pictures and tables.

**5. Semantic Density**

This category measures the LO's usefulness as compared to its size or duration. A small (few bytes size) interactive java applet may have a higher semantic density, as opposed to a collection of web pages (of few megabytes size), due to its effectiveness and better learning usefulness as well as its high interactivity and demonstrative nature. The vocabulary defines the semantic density:

**TABLE 3.2: Semantic Density**

| Semantic Density |
|---|
| Very low |
| Low |
| Medium |
| High |
| Very High |

**6. End User Description**

End user description encapsulates two important roles of the LO:

(i) The type of user (person)

(ii) The context of use (based on educational level)

The type of user gives the typical user of the LO, most dominant first. These types are: the ***knowledge developers*** (also the *authors* or *teachers*), the ***reviewers*** and the ***learner***. This is based on how LOs are handled with regards to its intended use. All groups of users can search and download LOs from the library. The teacher uses the LO to conduct a class. The author uses the LO to developed course modules or research papers. The reviewers assist in quality determination. While the learners use it for gaining information and knowledge. Note teachers become knowledge developer if they develop a new course module or paper (which may contain LOs downloaded from DLNET) and submit it to the digital library. But this is more of a user role in the context of DLNET as opposed to context of person using the LO.

The context of use gives intended user environment i.e. the educational level at which the intended use of the LO takes place. For DLNET these categories are (modified version of IEEE/IMS LOM categorization):

- Higher Education
- University First Cycle (I and II yr)
- University Second Cycle (III and IV yr)
- University Graduates (Masters)
- University Post-Graduates
- Technical School
- Practicing Engineers
- Continuing Education
- Vocational Training

During the DLNET discovery process, the learner's groups like engineers, university students, etc. use this field to identify the LOs relevant for them. The teachers and authors use this field to identify LOs for their intended audience. On the other hand during the DLNET submission process, the creators use this field to inform searchers of the target audience of their submission.

**7. Difficulty and Learning Time**

Every learning object has an associated difficulty level for corresponding intended user. It also has defined minimum time required to work with the resource.

**TABLE 3.3: Context of Use**

| Difficulty | Learning Time |
|---|---|
| Very easy | Few days (1 to 30) |
| Easy | Few weeks (4 to 10) |
| Medium | Few months (2 to 6) |
| Difficult | |
| Very difficult | |

## 3.3 Functional Requirements of LO

The functional requirements of the LO defines parameters that allow user-friendly interface to the LO. This at the same time allows efficient and uniform machine handling of the resource. LOs must meet the entire criterion at all time to be defined as a LO. The functional requirements are:

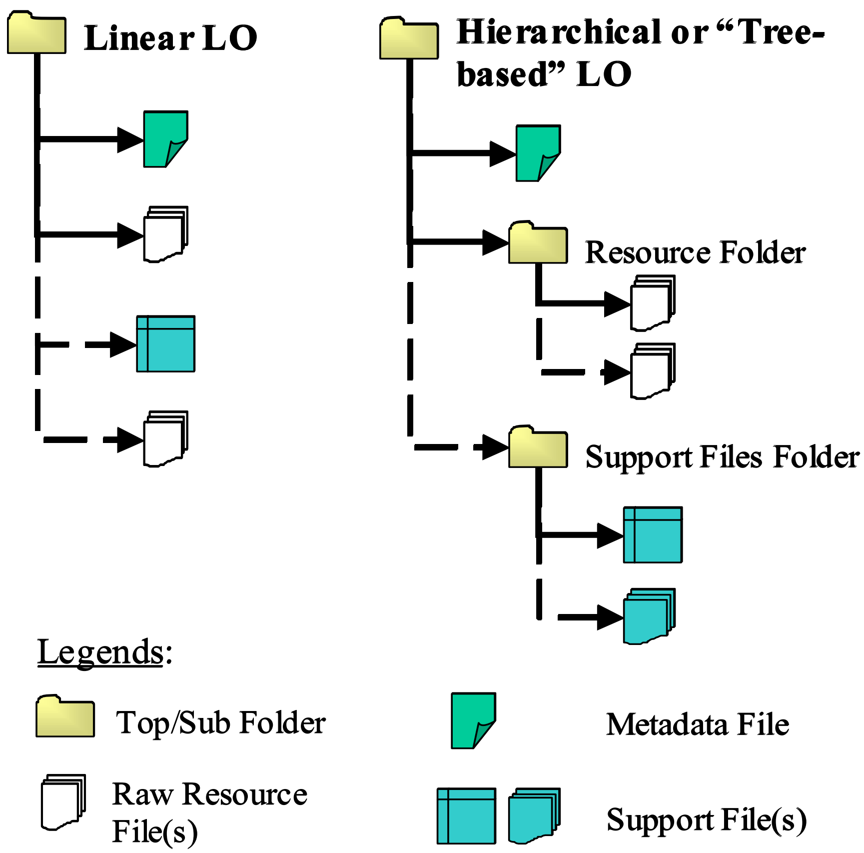
- The LO must have an associated file that contains information (like structure, ownership, usage rights, target audience, etc) about it.
- The LO must playable in know and easily available popular software. (The chosen application in this case a web browser)
- All the objects of the LO must be contained within the same folder.
- It is immutable in the sense that it may not be changed or altered by the user.
- The objects in the LO must be organized in a fashion that will facilitate learning.
- It must function independently (self-contained and standalone).

## 3.4 Learning Object Model

Given the above definition, characteristics and functionalities, there exist some unique ways in which resources and metadata can be physically put together in a learning object model. There are many advantages in exploiting known file structures to represent LO. They provide varying degree of configurability and scalability within the digital library architecture.

### 3.4.1 FILE STRUCTURES

The two most common types of file structures are shown in Figure 3.2. It is the linear and the hierarchical ("tree") based model. In the linear model, all the components of the LO are put in the top level of the file system. The components are distinguished by their unique names and linkage embedded within them. This style of packaging is not scalable and places many constrains on configurability of components. For example, the author has to ensure that all files are unique with respect to the underlying operating system – this limits the creative ability to put together learning modules.  From the server side, having linear structure makes programs difficult to identify and extract medias for selective archiving.  The advantage of this structure is convenience of access for applications "playing" the LO.




**(a) Linear** **(b) Hierarchical**

**FIGURE 3.2: Possible File Structures for LO Model**

The hierarchical model of packaging LO is considerably more scalable compared to linear model.  Here metadata files, resource files and support files (if any) can easily be placed in distinct folders and thus preserve their independence.  Additionally there is no conflict with file name of the components.  However care must be taken to ensure that top-level folders are unique.  This is a small concession, and at time advantageous for learning object development tool.  It also allows to separate different medias, while maintaining robust linkage for applications to "play" the LO.  All the DLNET LO use the hierarchical architecture for packaging.

### 3.4.2 LO ARCHECTURE

Having decided on file structure for the LO, let us now look at blocks that required to implement an LO. It is split into four groups:

- **LO Preamble**

  Unlike in a physical book, where one knows that the appropriate side of the cover is the starting point which is then followed by a table of content and then the text, digital contents as such inherently do not have any structured starting or continuation point. Asking authors to do this job is tedious and automatic chronology extraction tools have so far been not quite successful. We have proposed to implement a simple wrapper as a part of the LO which allows for its flexible use. Figure 3.3 shows the conceptual model. The LO will always play within a web browser. The front of the LO will be a designated frame-based HTML page that will list the essential set of data describing the author and usage in the top frame. The top frame initially contains a selected summary of the metadata and a click able link that starts "playing" the LO. Upon initiating the link, the contents of the LO play in this frame. The bottom frame is used to navigate through the LO. It many contain many options, for example: re-start the LO and/or to view more complete set of LO related information. LO preamble initializes the LO and displays information as to what is to follow. It is a HTML file that has two frames -one on top of another. The LO is initialized through this file. It always has the same filename and is the designated starting point of the LO. This file is located at the top level within the folder that contains the LO.

- **LO Metadata File**

  LO metadata file contains information about the LO. It is based on the desired metadata standard, a system parameter uniform across the digital library. Generally, it is an XML file (to be discussed in the next Chapter), which contains the creator's name, organizational affiliation, LO's ranking, objective, pre-requisites, and many other types of information. The recommended location of the file is at the top level within the package

- **Support Files**

  Support files are optional. Due to modular format of the LO and need for versatility of use, additional support files are provided, which give instructions to users for appropriate use of the LO. They may be required in order to perform functions provided in the bottom frame of the wrapper. All support files must be located in a sub-folder labeled ***SupportFiles*** within the top LO folder.

- **LO resources**

  LO resource files are the collection of digital constructs (the objects) that convey educational information. These files are located in a sub-folder labeled ***Resources*** within the top LO folder. It may contain sub-folders as desired by the knowledge developer. The relative paths are accordingly adjusted, to allow proper functioning.

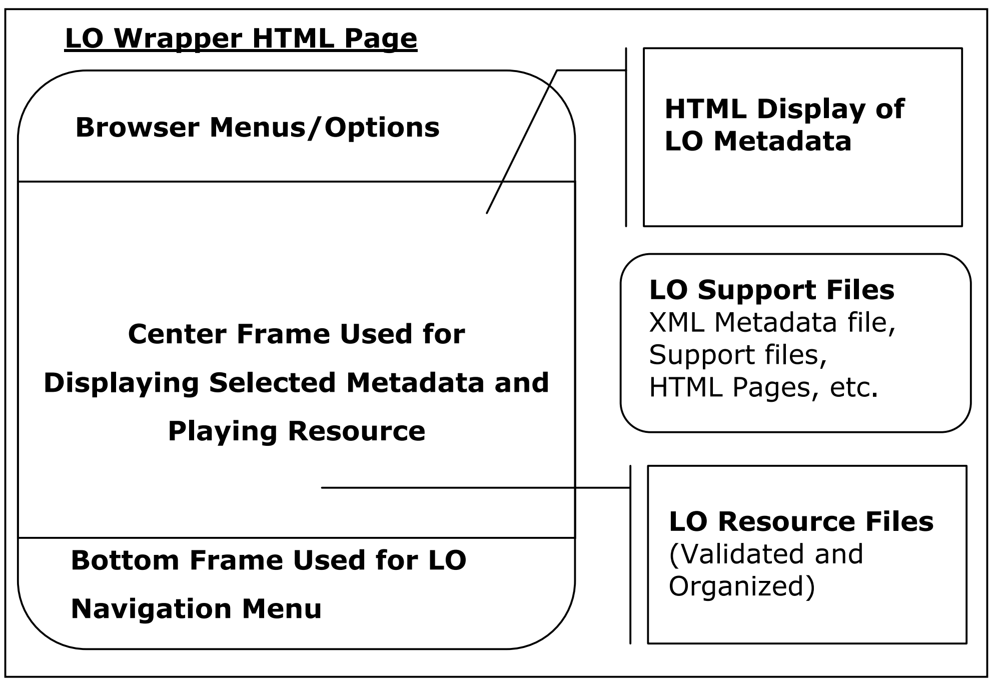



**FIGURE 3.3: The LO Architecture**

Figure 3.4 shows the complete file structure of the proposed LO file package. When an LO is embedded into another LO, the same structured is preserved, but it is placed within the *Resources* folder as it is an "object" for the upper level LO.

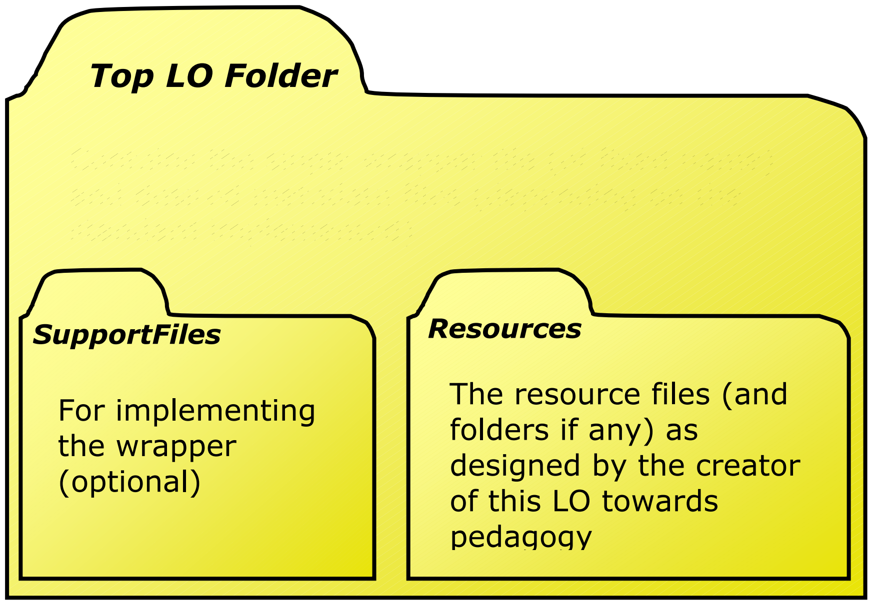



**FIGURE 3.4: File Structure of a DLNET LO**

## 3.5 Types

The digital constructs that form the "object" in LO are available in different forms. The type LO is distinguished based on the type of the digital construct (the content in the *resource* folder). Here a simple distinction is provided namely; web-content based LOs and non-web (browser) supported content LOs. Note that other distinctions can be made to classify LOs. The HTML/non-HTML separation is convenient as all LOs must play in a browser thus it is easier to distinguish the examples explain LOs. The main advantage of this typecasting of the LO is processing them via the content development tools. Such a tool is discussed in the next chapter.

### 3.5.1 WEB CONTENT

Web contents include a broad form of LOs that play in a browser. They can be HTML pages, with pictures, graphics (GIF, JPEG), active codes (JavaScript, JAVA applets, flash, etc.), among others. Note LOs, depending on the browser setting and plug-ins may or may not play a designated web-type LO in some browsers as opposed to others. As noted the LO typecasting is loose and as such holds no rigid standards. A more appropriate classification is the learning resource type as discussed in Section 3.2.

### 3.5.2 NON-WEB CONTENT

Non-web content-based LOs are those that don't play in popular browsers. They are specialized codes based on proprietary software like MATLAB, C, etc. As more and more companies migrate to enable these software to work with Internet, the content and LOs within them may become web-content based LOs.

## 3.6 Extension to LO Model

LO is defined as a structured electronic resource that enables pedagogy. Hence any added functionality geared to providing additional support in assisting with development or discovery of learning module within LO is very useful to improve its usability. Two such variations envisioned by the author of this thesis are easily extensible (and implement able) from the base definition of the LO.

### 3.6.1 NESTED AND LINKED LO

Consider the example in Figure 3.5 as a demonstration of extending and integrating learning objects into a wider course module, which in itself can be a LO. In order to link the learning object to the course page, an HTML link should be initiated to the designated starting page of the learning object. Note that there may be many starting pages and the course developer should be aware, through information available in the metadata, as to which page is the starting page for the target audience. Once the designated section of the LO is finished, it may be possible to redirect the LO to designated page to exit to on the course developers course module. Such implementation would effectively nest the LO within the course developers resources. The new resources can then be submitted back into the digital library as new LO (Note that the copyright will automatically propagate as metadata is embedded within the nested LO). This type of formation is called nested or embedded learning object structure. The learning object development tool can be easily extended to accommodate such LO management environment.

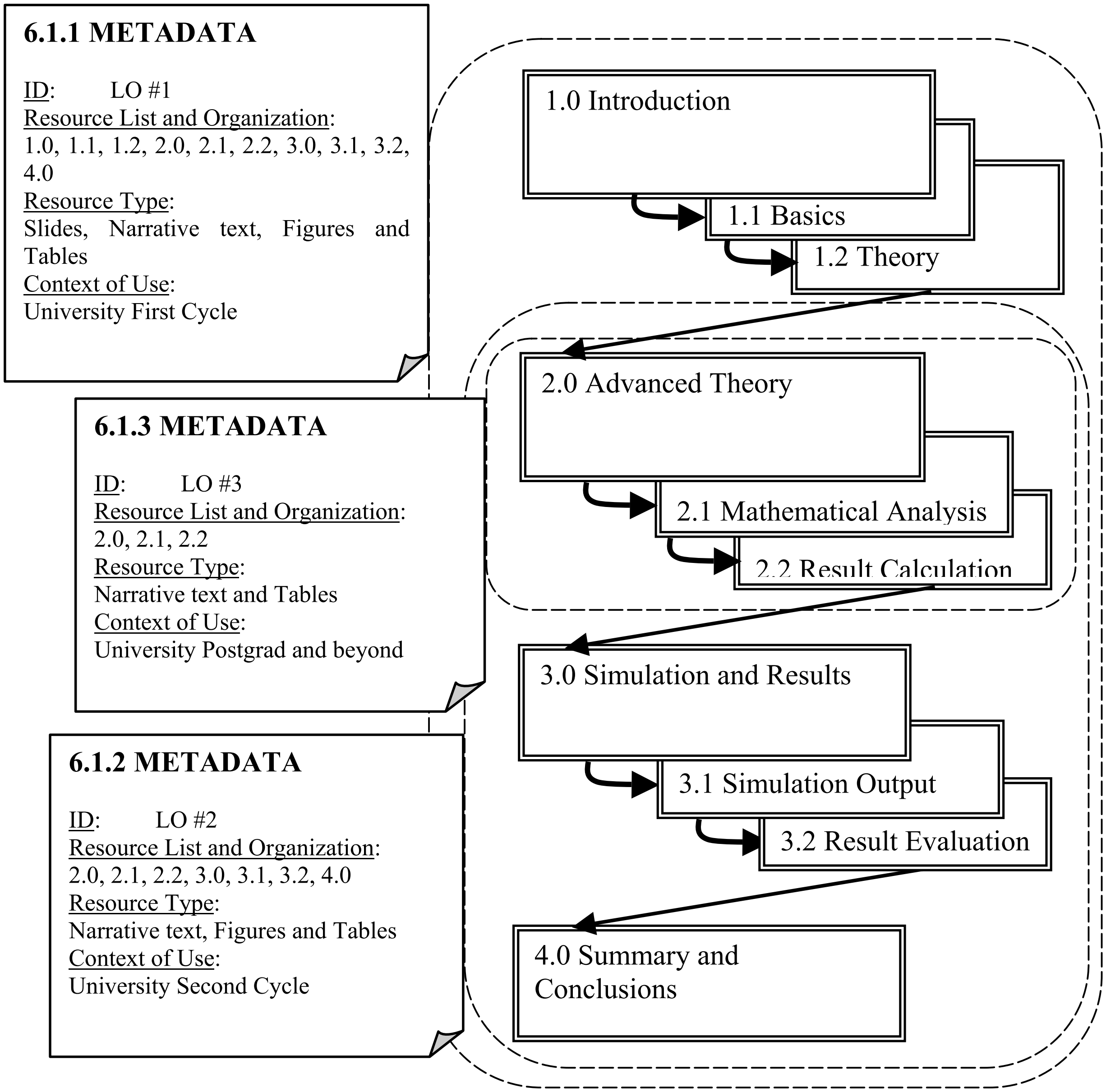



**FIGURE 3.5: Nested Learning object Example**

Figure 3.5 illustrates an example of embedded learning object. Here, the top learning object is the course module with introduction, advanced theory and so on. Note that the section under "Advanced Theory," when recognized as a learning object, can stand on its own as LO #3. A similar situation exists in the case of resources enclosed in sections 2.0 to 4.0, i.e., LO #2. The aggregation level up to which top learning object can be broken up to create new embedded learning objects is left to the creator of the module. The reviewer judges each learning object as a unique entity and should approve of quality at each level.

### 3.6.2 "INTELLIGENT" LO

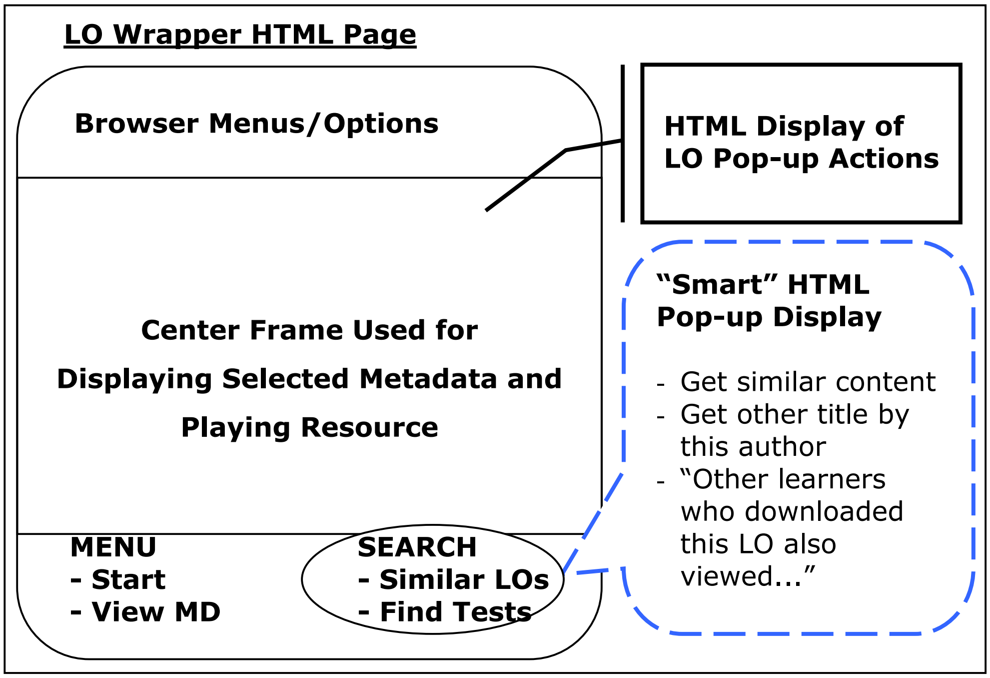



**FIGURE 3.6: "Intelligent" Learning object**

One of the main reasons to use browser as a platform for LO manifestation, is the versatility of different functions that can be implemented or embedded with the LO. This can be done without actually modifying the content and hence is scalable over extended use of the LO. One such idea is to impart intelligence to the LO. The proposed idea is shown in Figure 3.6. Here an additional menu option is proposed in the bottom navigational frame of the LO. By clicking options here, a HTML pop-up window can be opened. Various options can be embedded here to enable an interactive experience. Some click-able parameters are listed within the blue dashed box in Figure 3.6. Upon the choosing the link the appropriate result of the option is displayed on the central display frame. The implementation for this type of "intelligent" LO requires several server side additions and supporting environment.

Chapter 4 discusses the LO Development Tool that can facilitate the creation of the LOs followed by Chapter 5 on the server side architecture for distribution and support such an LO.

Chapter 4

# The LO Development Tool

A learning object is defined as a structured electronic resource that has two components: metadata and resources. In order to facilitates a goal-oriented pedagogy while at the same time preserve copyright, it is important to collect a vast range of the metadata. The quality of the metadata facilitates a natural discovery for the LO. From a technical perspective, the LO is collection of raw resources that are "structured". Hence, resource validation is an equally important part besides metadata collection. Finally, both these component have to be bundled for transmission and distribution within the digital library. The author has proposed and developed a tool to accomplish the above-mentioned tasks. It will collect pedagogy and copyright metadata relating to the resources, validates the resources and packages it. Each of these tasks is described below in details. Following which the architecture and layout of the tool is discussed.

## 4.1 The Metadata Collections

Any proposed tool has to collect, process, store and convey the information about the LO - the ***metadata***. A tool must have well defined fields where user can enter the data. Fields like creator's name, organizational affiliation, occupation, etc. will harvest the metadata relating to the creator. Further fields will be available to capture the characteristics and technical properties of the LO (Section 3.2 and 3.4). These are the educational metadata. Here there are two options. One let the user decide the values of the field based on his/her best experience or prior knowledge. Two, the tool contains pre-defined values for the areas so that the content creator can select the best fit for their content. Some advanced software can automatically parse the resources and extract the required metadata. The author has found that such tools, though somewhat efficient, required considerable knowledge on the part of the operator and are times inconsistent in their approach and hence cannot be completely relied upon to do this task. As high quality of metadata is required (later for quality control and high-yield discovery), the mixed approach to allowing creator to fill-in some metadata areas, while selecting others is judged to be most optimum approach to metadata collection. The collected metadata can be stored and transmitted by many means and methods. The digital library standardization community is currently formulating standards for this action. Here Extensible Mark-up Language (XML) has emerged as the language of choice.

### 4.1.1 WHY XML?

XML is a new and emerging web technology. XML has been designed for ease of implementation, and for interoperability among various data management systems. It enables a convenient and scalable means of communicating data between two nodes, this at the same time is easily supportable by legacy web system. Further XML is an open standard and well supported for implementation. A detailed explanation of relation of XML for metadata handling issues, relevant to digital library, is discussed in [18]. The XML in essence provides a way to bind data into a know format that is easily readable by machines. The binding is structured, based on rules so that errors and duplication can be avoided. A set of required metadata bound in XML is called an XML instance.

Many organizations have proposed different rule for creating XML instances for metadata collection. Dublin core, IEEE-LTSC's LOM, IMS are some to name a few. The difference between them is the depth and variation of metadata required for binding to be classified under their system of standard.

### 4.1.2 COMPARISON OF METADATA STANDARDIZATION

A broad comparison of various current and known metadata standards available for educational metadata binding is given in [14]. The information must be packaged and conveyed using a XML file in keeping with current standards and practices in metadata handling.

The Dublin Core Metadata Initiative is one such initiative. This paragraph is compiled from the information at the Dublin Core website [19]. It is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models. It is developing a specialized metadata vocabulary for describing resources that enable more intelligent information discovery systems. It has two types: Simple (unqualified) and qualified Dublin Core. "Simple Dublin Core" is Dublin Core metadata that has only the main 15 elements of the Dublin Core Metadata Element Set are expressed as simple attribute-value pairs without any "qualifiers" (such as encoding schemes, enumerated lists of values, or other processing clues) to provide more detailed information about a resource. Qualified Dublin Core" employs additional qualifiers to further refine the meaning of a resource. One use for such qualifiers are to indicate if a metadata value is a compound or structured value, rather than just a string. Qualifiers allow applications to increase the specificity or precision of the metadata. Following is a set of metadata elements defined by Dublin Core:

- **Name** - The label assigned to the data element
- **Identifier** - The unique identifier assigned to the data element
- **Version** - The version of the data element
- **Registration Authority** - The entity authorized to register the data element
- **Language** - The language in which the data element is specified
- **Definition** - A statement that clearly represents the concept and essential nature of the data element
- **Obligation** - Indicates if the data element is required to always or sometimes be present (contain a value)
- **Datatype** - Indicates the type of data that can be represented in the value of the data element
- **Maximum Occurrence** - Indicates any limit to the repeatability of the data element
- **Comment** - A remark concerning the application of the data element

IMS Global Learning Consortium, Inc. in a similar effort is developing and promoting open specifications for facilitating online distributed learning activities such as locating and using educational content, tracking learner progress, reporting learner performance, and exchanging student records between administrative systems via metadata. This paragraph is compiled from the information at the IMS website [20]. The most relevant work is relating to the IMS Learning Resources Meta-data Specifications (released August 20, 1999) which creates a uniform way for describing learning resources so that they can be more easily found (discovered), using meta-data aware search tools that reflect the unique needs of users in learning situations. Another one is the IMS Content & Packaging Specification will make it easier to create reusable content objects that will be useful in a variety of learning systems. Related standard in terms of the overall modern profile of digital library (Section 2.2) are the IMS Question & Test Specification addresses the need to be able to share test items and other assessment tools across different systems and the IMS Learner Profiles Specification will look at ways to organize learner information so that learning systems can be more responsive to the specific needs of each user. It has elements similar to Dublin Core but also include many of the characteristics required by the LO.

The Learning Technology Standards Committee (LTSC) is chartered by the IEEE Computer Society Standards Activity Board to develop accredited technical standards, recommended practices and guides for learning technology. This paragraph is compiled from the information at the IEEE-LTSC website [21]. For the purpose of this thesis, the work done by the IEEE P1484.12 Learning Object Metadata Working Group is most relevant. This standard will specify the syntax and semantics of Learning Object Metadata, defined as the attributes required to fully/adequately describe a Learning Object. Learning Objects are defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning. The Learning Object Metadata standards will focus on the minimal set of attributes needed to allow these Learning Objects to be managed, located, and evaluated. The standards will accommodate the ability for locally extending the basic fields and entity types, and the fields can have a status of obligatory (must be present) or optional (maybe absent). Relevant attributes of Learning Objects to be described include type of object, author, owner, terms of distribution, and format. Where applicable, Learning Object Metadata may also include pedagogical attributes such as; teaching or interaction style, grade level, mastery level, and prerequisites. It is possible for any given Learning Object to have more than one set of Learning Object Metadata. The standard will support security, privacy, commerce, and evaluation, but only to the extent that metadata fields will be provided for specifying descriptive tokens related to these areas; the standard will NOT concern itself with how these features are implemented.

The above three metadata encapsulation standards are known to be gaining wider acceptance among the digital library community. After reviewing the above available standards, the author selected the IMS as the standard to implement with the LO development tool. IMS has many advantages over other schemes as it already comes with metadata elements for educational information. It further has complimentary metadata standards for content packaging. More importantly, it has accompanying standard on testing and user profile maintenance for learners in a digital library. It thus has all the required set of Metadata to define all the interactions in the modern profile of the digital library. Dublin Core is the chosen standard by NSDL. The author believes that designing crosswalks between is relatively simpler and lot more advantages exist by using IMS as the primary metadata standard within any digital library.

IMS is also the chosen standard for the Sharable Content Object Reference Model (SCORM) used by US Department of Defense for continuous education [22]. SCORM defines a Web-based learning "Content Aggregation Model" and "Run-Time Environment" for learning objects. The SCORM is a collection of specifications adapted from multiple sources to provide a comprehensive suite of e-learning capabilities that enable interoperability, accessibility and reusability of Web-based learning content.

Regardless of the choice of standards, the designer of any LO developmental tool should be aware of all standards and provide support for the most popularly used ones. Additionally work is being done on developing crosswalks between these metadata standards to easily translate one into another. Detail understanding of the IMS-IEEE LOM standards, content packaging or XML file handling is not required to participate in DLNET. As mentioned earlier these processes are handled by the DLNET Tool and are therefore transparent to the user.

## 4.2 Resource Validation

After metadata collection, the next important task is to verify the "structural integrity" of the resource that form the LO. This is the resource validation process. Validation assures that no broken links will exist within the file structure or organization by parsing through the resource folder. In other words, an LO must have all the constituent files aboard. For a web-based content, the validation process goes through all the resource locator tags and ensures that the resources are located in the said path. It will then collect the validated resources for centralized packaging. For non-web type LOs, validation process ensures the presence of the file with the resource basket. A summary of all resource files should be presented to the creator and should be editable up until the last moments, before packaging to ensure technical validity of the content. Once validation is complete, the information about the organization (or structure) of resource is appended to the XML metadata file in conformance with IMS-IEEE content packaging standards. The resource validation is automated with user intervention/verification allowed at anytime.

## 4.3 Packaging

Once metadata collection and resource validation is accomplished, the final step in creation of LO is packaging. Packaging bundles together the two components in a prescribed fashions to allow convenience of quality control, re-use, discovery and delivery. As mentioned in the previous chapter, part of this packaging is to create the preamble and support files. The whole list of file is then structured in the selected file format (Section 3.4.1) and is then ready to be reviewed for archiving with the digital library.

As IMS standards are the adopted method for processing all actions, IMS content packaging standard was adopted to package the metadata and resource in a learning module. This creates a zip file with the metadata information at the top folder. The metadata information is stored in a file called ***imsmanifest.xml*** that has three components to it. Figure 4.1 illustrates the components and is reflective of the bundling within it. The *imsmanifest.xml* file contains three main sections. The metadata information, the organization of the resources and the resource names themselves. The organization of the resource helps in identifying the pedagogical progression from one file to the next. Manifest file may contain sub-manifests within them. These may contain information about the LOs that are embedded with this top LO (Section 3.6.1). This type of architecture hence ensures copyright preservation and propagation along with the actual use of the LO.

Once the packaging is complete the creator of the LO is expected to submit the LO to the digital library, where quality control and further archiving and distribution steps are undertaken. The author of this thesis document has proposed and implemented a tool architecture to assist LO creators in accomplishing the above three functionalities towards transformation of resources into LOs.

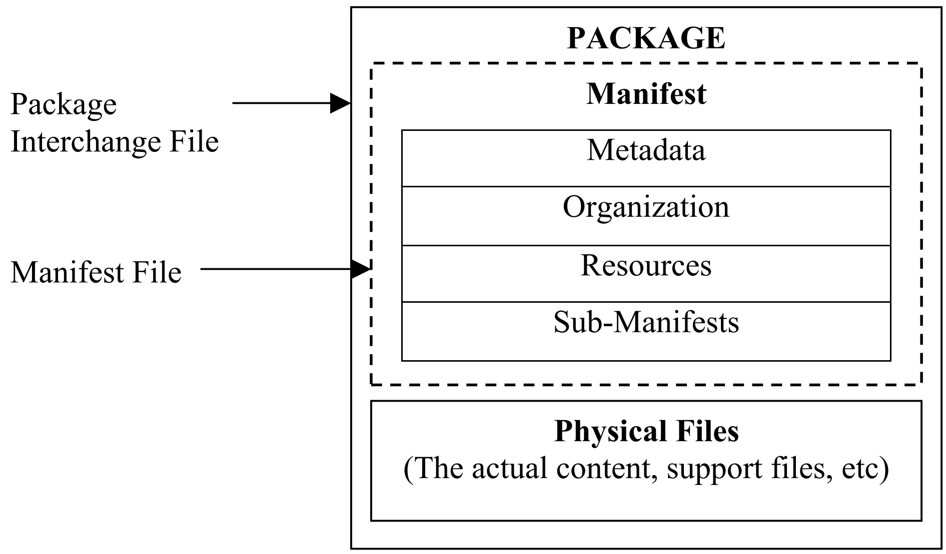



**FIGURE 4.1: IMS Content Packaging Scope**

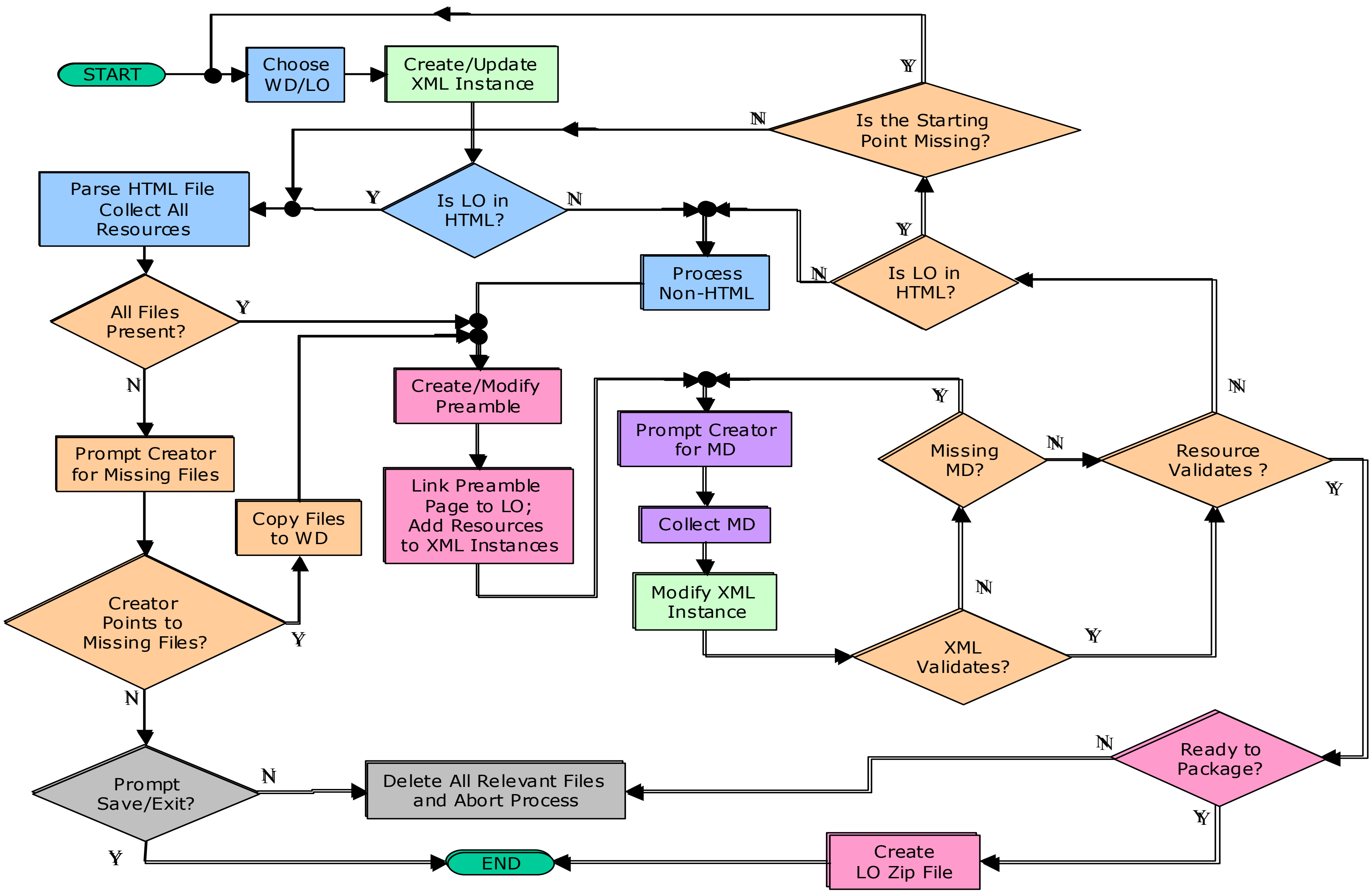
## 4.4 Architecture

The layout of the tool is designed with mind that many authors are contributing material voluntarily and may have varying adaptation-curve or aptitude towards learning new applications. The tool must take them through a step-by-step process of resource to LO transformation. It was also noted that collection of metadata like author name, contact information, title, description, etc. may have privacy concerns and hence care must be taken to explain the need to seek such information from the creator. As mentioned earlier the information for some of the fields must be provided by the creator, while others may be selected from preset options. Overall it is envisioned that collection of metadata should not take more than 5 to 10 minutes.

Validation and packaging time is dependent on the total size of the resources. Validation for web-based content can be automated, as usage of HTML tags is well known and structured to easily be parsed via software algorithms. For non-web based content, the creator is expected to manually add the resource to the validation pool.

It is proposed that the tool be implemented as a client-based application. Being a client-side application, overhead of required server-side or online connection is avoided and additionally provides better management of the LO creation process. This at the same time allows uniform machine processing, easier access and delivery to other users. The creator can stop at any moment and re-visit the project for completion at other times. As mentioned earlier the major functions of the tool will be to facilitate content submission process by guiding authors through the procedure. The Tool is best described by the list of "action verbs" listed below in rough sequence of their execution. The process is shown in detail in the flowchart (Figure 4.2).

1. **Harvest Metadata**: Information regarding the content and the contributing author(s) will be collected using a form that the contributing author(s) is expected to fill out.
2. **Create XML Metadata Record**: A first-cut metadata record will be compiled from the returns of the Form in Item 1. The IMS/IEEE Learning Object Metadata standard will be used in the Tool. The record will be in XML format and in compliance with the designated schema.
3. **Get Content**: The contributing author(s) will be prompted to upload their content to the Tool. Content is expected to be in either html (default) or non-html format.
4. **Validate**: Both the content and the associated XML metadata record will be validated at this stage. The latter involves validating the metadata record against the designated schema. Validating content involves verifying the file structure and the embedded links by parsing through the document. At the same time, the Tool will extract additional metadata regarding the file structure.
5. **Update XML Metadata Record**: The XML Metadata Record of the content will be updated based on the additional manifest info obtained at the Validation stage.
6. **Encapsulate and Zip**: The content and its metadata will be packaged and zipped into a learning-object (LO) file ready for uploading to DLNET's content repository. The zip file will also contain the DLNET LO support files.
7. **Upload Content**: This functionality will be implemented at a later stage. For now, authors will be asked to upload their content (created in stage 6) by logging into their DLNET accounts or website.




Color Key corresponds to the following actions:

**Harvesting Metadata**
**Validating Content/Metadata**
**Creating/Updating XML Metadata**
**Packaging Content**
**Gathering Content**

**FIGURE 4.2: Algorithm for tool implementation**

### 4.4.1 JAVA CODING

The JAVA coding of the tool architecture follows the given flowchart. The code structure is provided in Appendix A. Class files are setup to allow easy flow control from one block to next. Metadata collections, retention and presentation are monitored via set of required Boolean parameter. An essential set of metadata items are mandated to be entered in the tool to allow eventual packaging. The creator is prompted for those metadata if any are missing. The list of the items given in Table 1.1 is mandated as required metadata. Since this information is stored in XML file format, a special class of JAVA code was created to embed the data into XML file tags. This Class file is for IMS encoding. Special Class files would be required if the encoding is to be done in other metadata standards or alternately crosswalks would be required.

Validation is the next step in progression towards creation of the LO. Validation algorithm is done in two parts. The selection of the algorithm is transparent to the creator. The creator of the LO may add or remove any number of resources at any time from the validation process. It occurs at the discretion of the creator. But note in web-based content, the creator only has to provide the initial HTML file. The tool's validation algorithm, then parses through this file and collects all the HREF links within it. The relative HREF linked files are then recursively parsed until all the resource files are collected. The absolute links (the outside web-links) are ignored. The creator cannot change the set of collected HTML validation resource as they are essential for technical integrity of the resource or in other words are required file to play the LO. The creator can provide a different starting file and thus ensure a different set of valid web-content resource is collected for packaging.

The final step is packaging. Up until this point the creator has full control to change either the metadata or the resources. The IMS packaging standard is that of zip file. Thus the tool packages the learning module containing the metadata and the validated resource into a single zip file. It further creates the HTML encapsulation and required support files as a part of the package. The final package is now ready for quality control and distribution.

Note the package can directly be used as a LO and distributed by the creator.  As such if the tool is designed by a third party (which conforms to the above mentioned LO design and standards) then the use of the free standing LO is left to the discretion of the creator or the copyright owner of the resource.  The author of this document recommends the LOs created via the tool be submitted to a supporting digital library for further processing as it enhances the value of the resource, provides wider copyright recognition with the knowledge developer community.

## 4.5 Layout and Usage

Before beginning to use the LO Tool, it is recommended that the creator of the LO assemble all files related to the resource you are contributing in one folder on the desktop.  Name the folder appropriately, i.e., do not use special operating system characters such as open space and : \ / * ?. As the tool is based on JAVA, the user is expected have Java Runtime Environment setup in their PC.

Figure 4.3 shows the starting page of the tool as soon as it is successfully invoked.  Step-by-step instructions follow on how to handle specific actions.

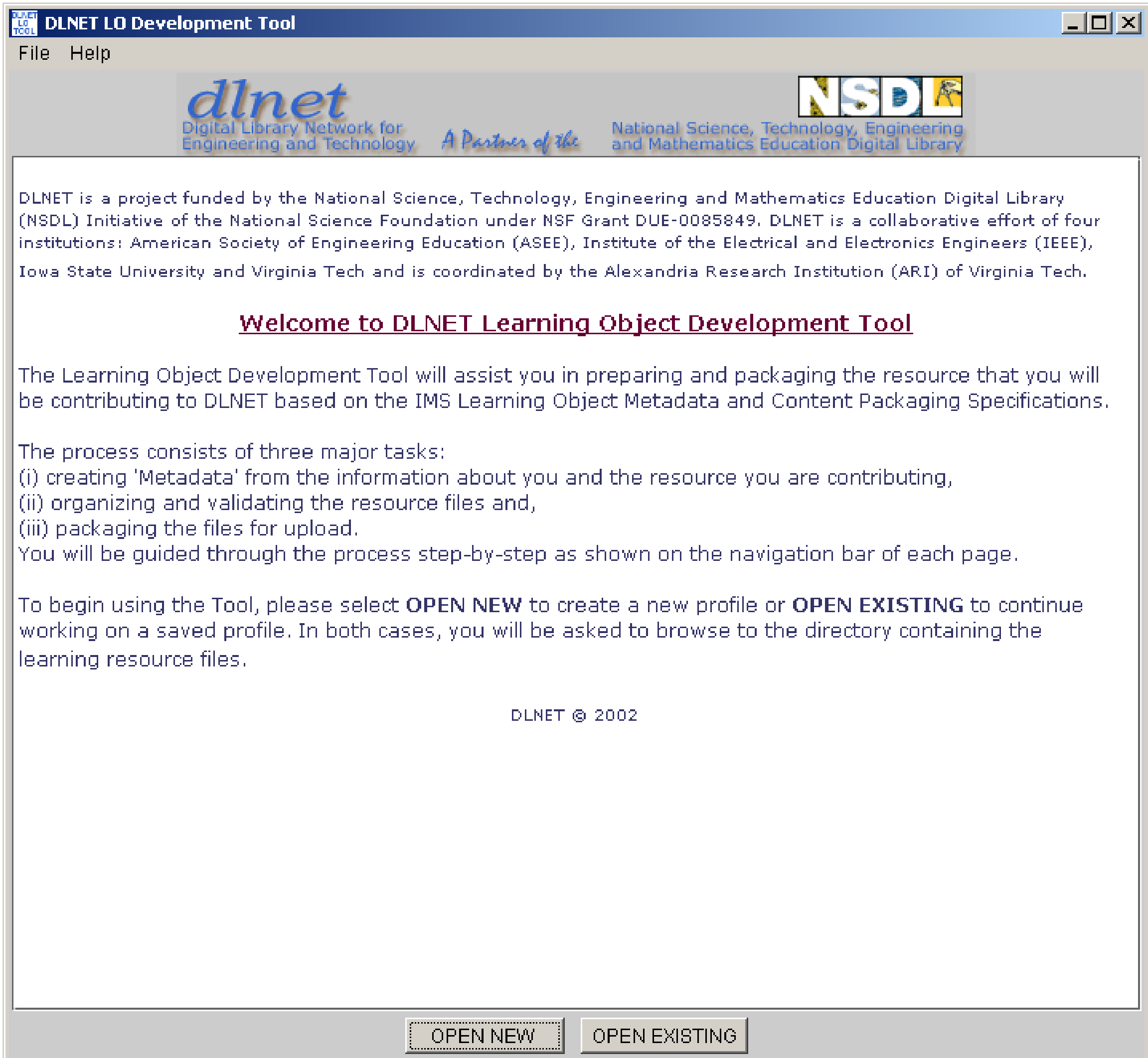



**FIGURE 4.3: Welcome Page of the LO Development Tool**

Double-click on the LO Tool icon to start the tool. The screen as seen on the top appears.

- Select **OPEN NEW** to begin a new session.
- Select **OPEN EXISTING** to continue on a previously saved session.

In both types of sessions, you must first point to the workspace or folder where you saved the content for submission. For new sessions, you will need to create a project file (an XML file with a ***.dln*** extension). Provide a filename for the project <*project_filename.dln*> or use the default filename shown that corresponds to the name of the folder containing the resource. For previously saved sessions, navigate to the workspace or folder and point to the project filename <*project_filename.dln*>. This will bring up the pages of the saved session.

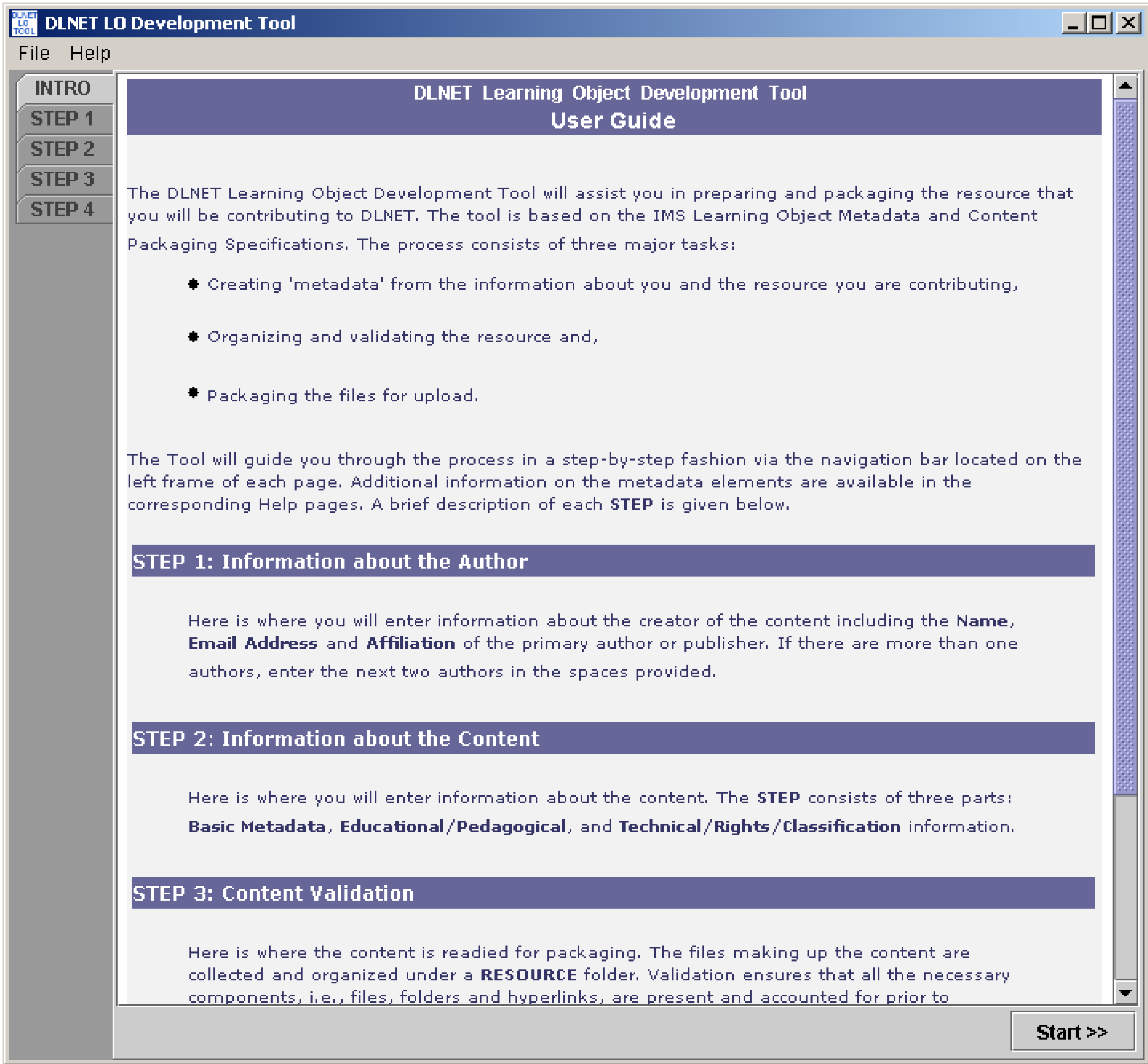



**FIGURE 4.4: LO Development Environment within the tool**

Once the project file is successfully initialized, the *Introduction Page* (Figure 4.4) appears. In case the project file is not properly initialized, a message appears with the reason and the user stays in *Welcome Page*. The *Introduction Page* provides a summary of the steps in the Tool. If first time user, it is useful to go over this information to know what is expected and when it is expected. Click the **Start >>** button to go to **Step 1**. This begins the actual creation of the LO.

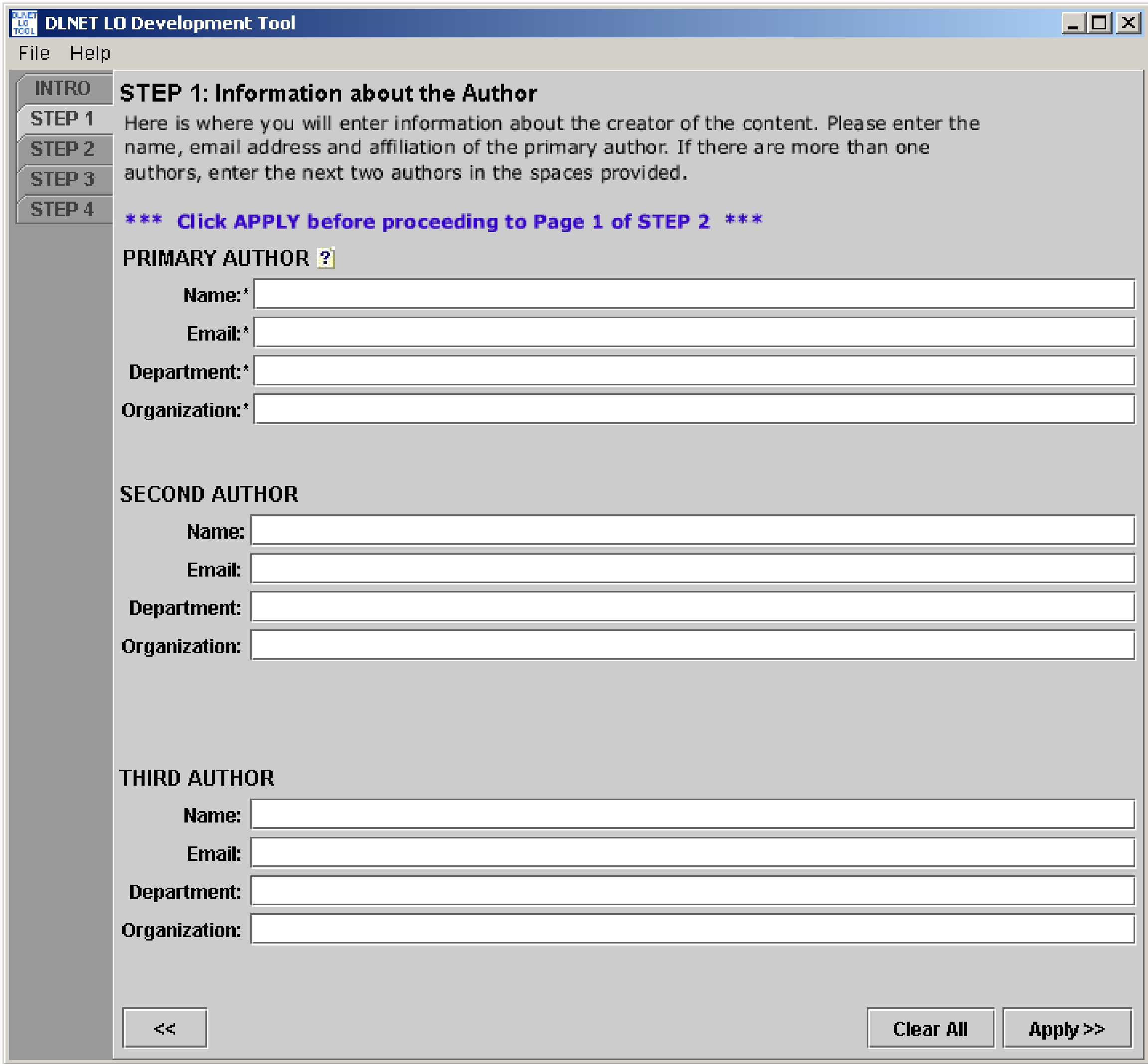



**FIGURE 4.5: User Metadata Collection Environment within the tool**

Figure 4.5 shows the GUI for Step 1. Here is where the creator will enter information about the owner (i.e. himself or herself) of the content including the Name**,** Email Address and Affiliation of the primary author or publisher. If there is more than one author, enter the next two authors in the spaces provided. In case the tool is being used by a third party to create the LO, please provide the information of the person or institution holding the copyright of the resource. Text areas are provided next to the metadata tag names to fill in the appropriate information.

Now click on the **APPLY >>** button to proceeded to the next page. Choosing **APPLY** writes the information into the XML project file and hence in the event it is to be invoked at a later time.

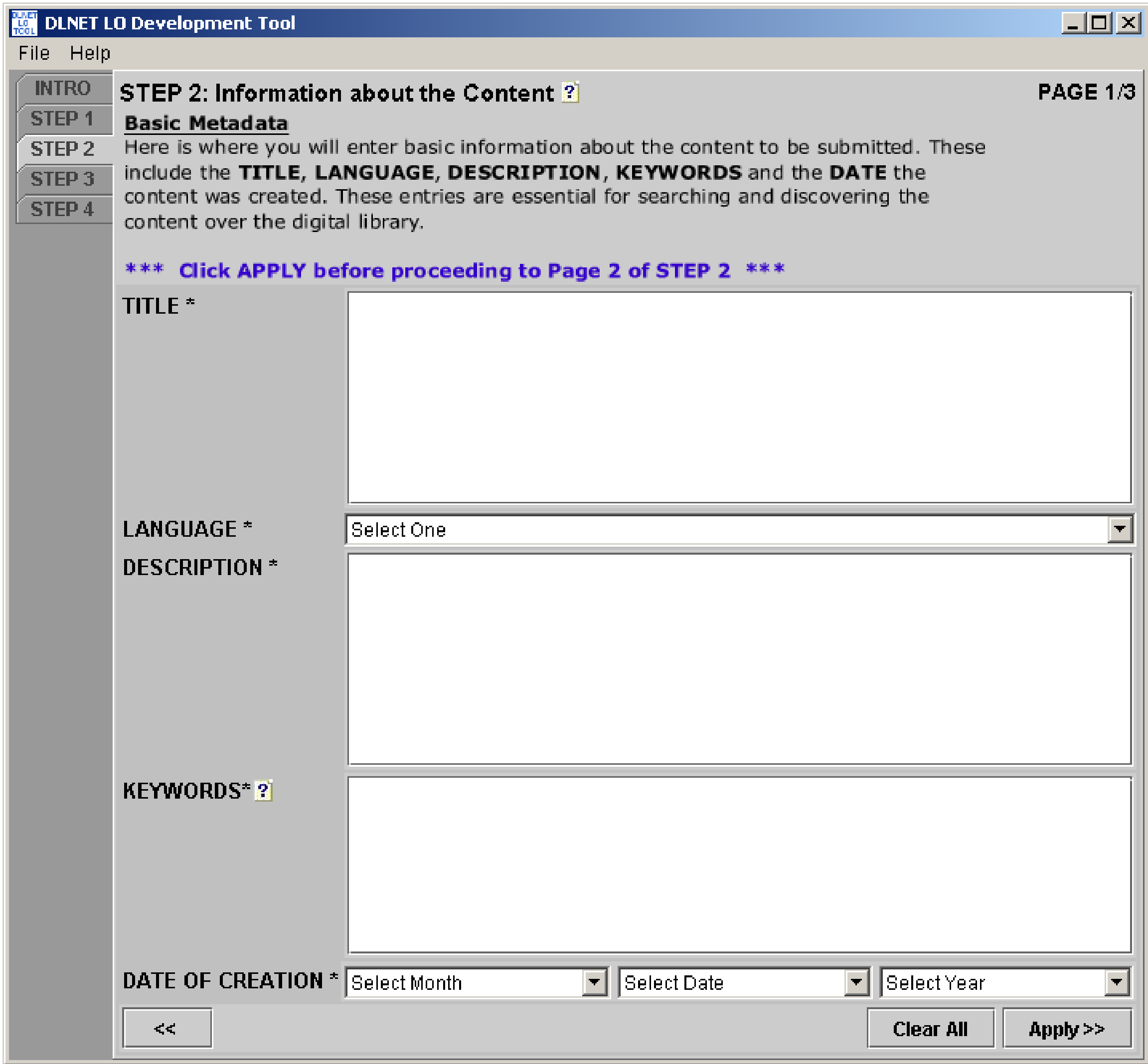



**FIGURE 4.6: Content Metadata Collection Environment within the tool**

STEP 2 is used to collect information about the content. It has three parts. Here is where the creator is expected to enter basic information about the content to be submitted. The first part (Figure 4.6) includes the **TITLE**, **LANGUAGE**, **DESCRIPTION**, **KEYWORDS** and the **DATE** the content was created. These entries are essential for searching and discovering the content over the digital library. The second part (Figure 4.7) deals with the Educational or Pedagogic information. Here, additional information about the content describing its learning or pedagogic features can be added. These include **RESOURCE TYPE**, **INTERACTIVITY TYPE** and **LEVEL**, **INTENDED USER**, **CONTEXT OF USE** and **TIME DURATION** required to learn the resource. Each item is provided with a list of possible values to choose from. Use the CTRL key to select multiple values that may apply to your content.

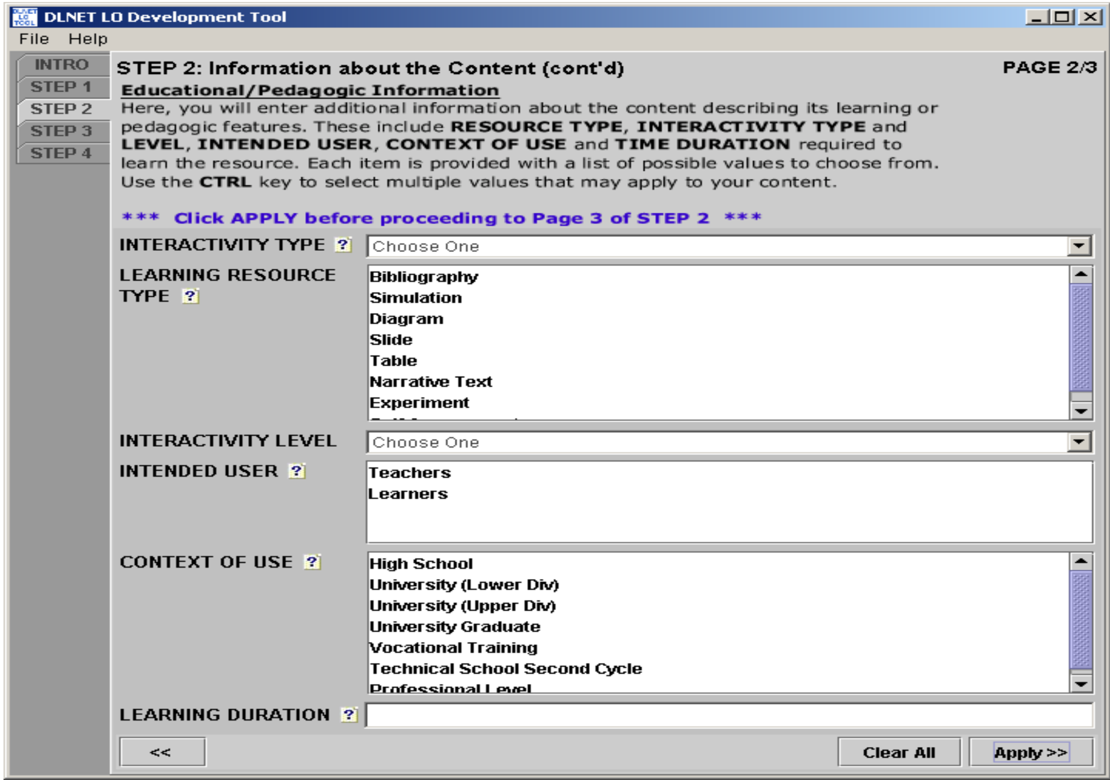



FIGURE 4.7: Content Metadata Collection Environment within the tool

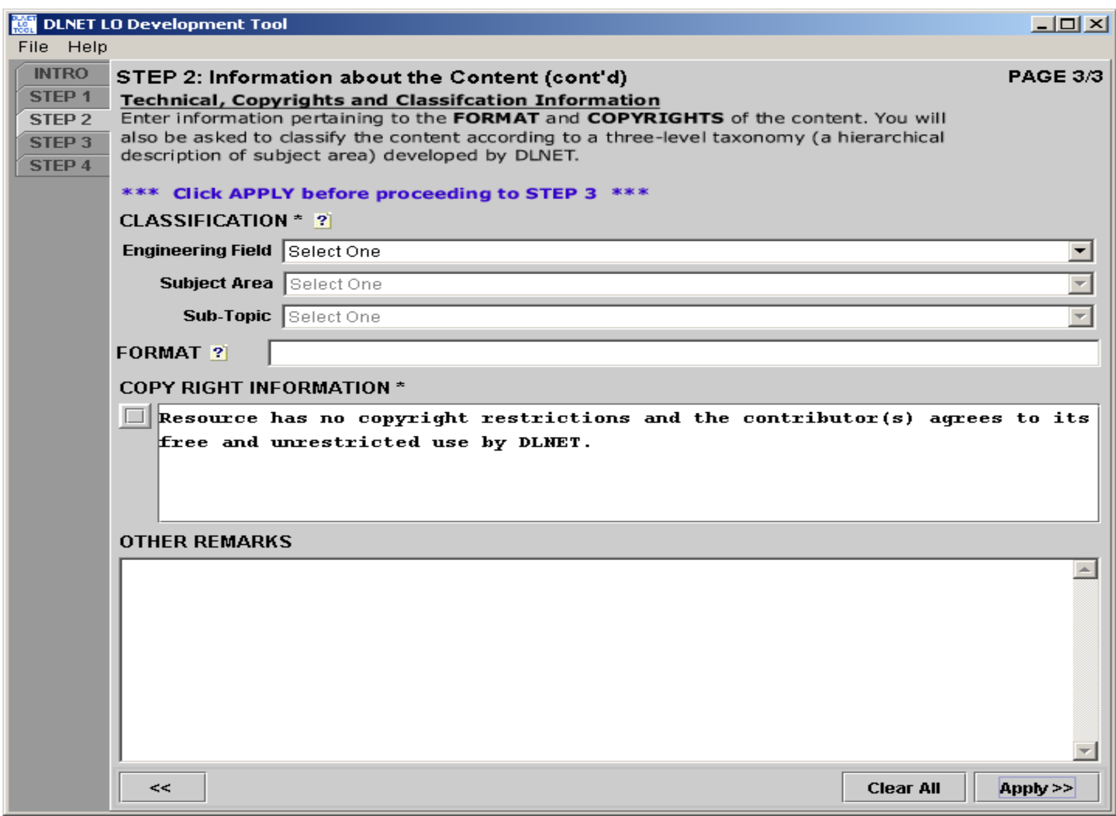



FIGURE 4.8: Content Metadata Collection Environment within the tool

The third part of STEP 2 (Figure 4.8) deals with the Technical, Copyrights and Classification Information. Here, the creator is expected to enter information pertaining to the **FORMAT** and **COPYRIGHTS** of the content. The content is classified according to a three-level engineering taxonomy. Remember to click **APPLY>>** to proceed to the **NEXT** part.

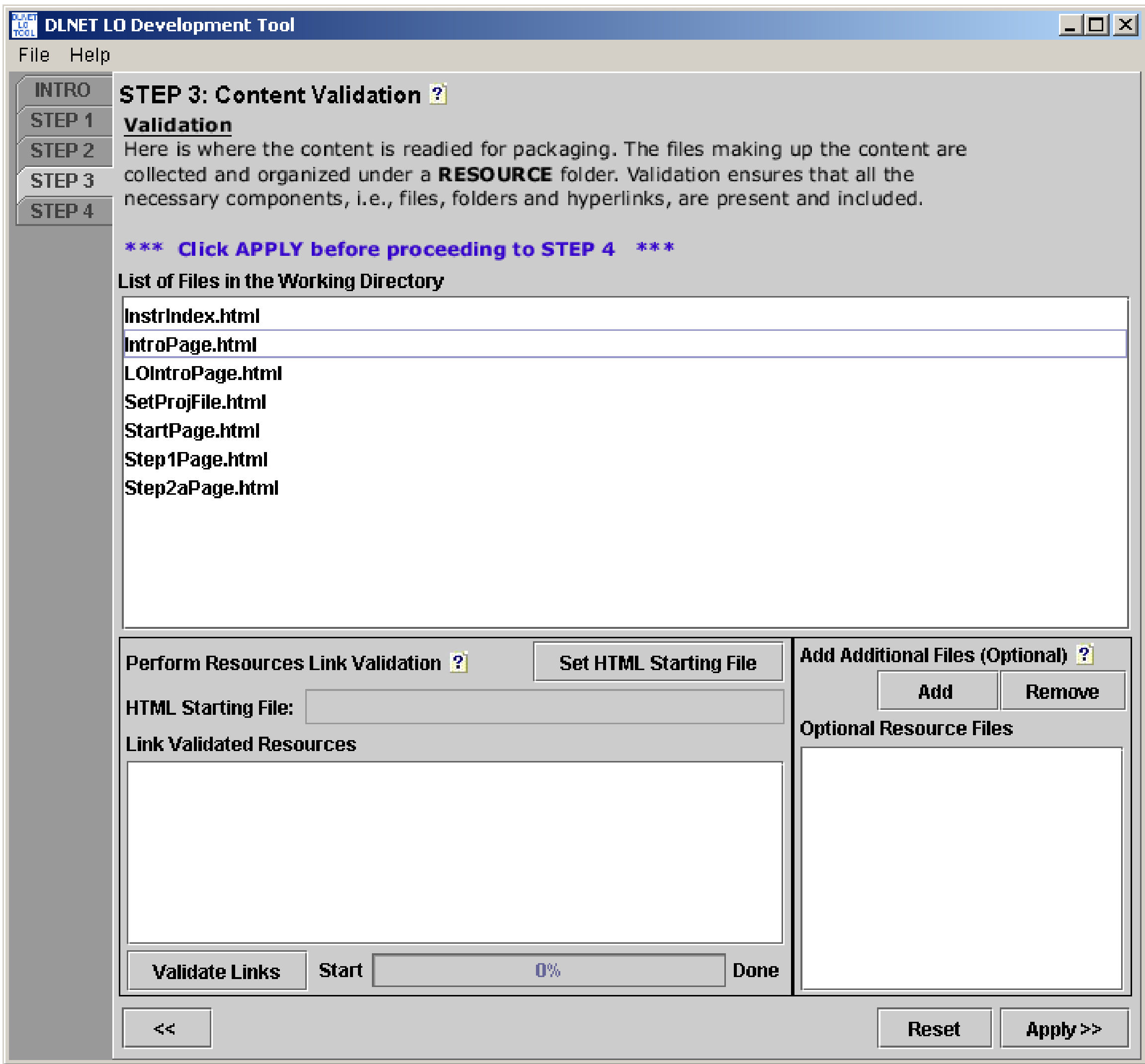



**FIGURE 4.9: Resource Validation Environment within the tool**

STEP 3 is used for resource validation. The files making up the content are collected and organized under a RESOURCE folder. Validation ensures that all the necessary components, i.e., files, folders and hyperlinks, are present. Click **APPLY**>> before proceeding to the next STEP.

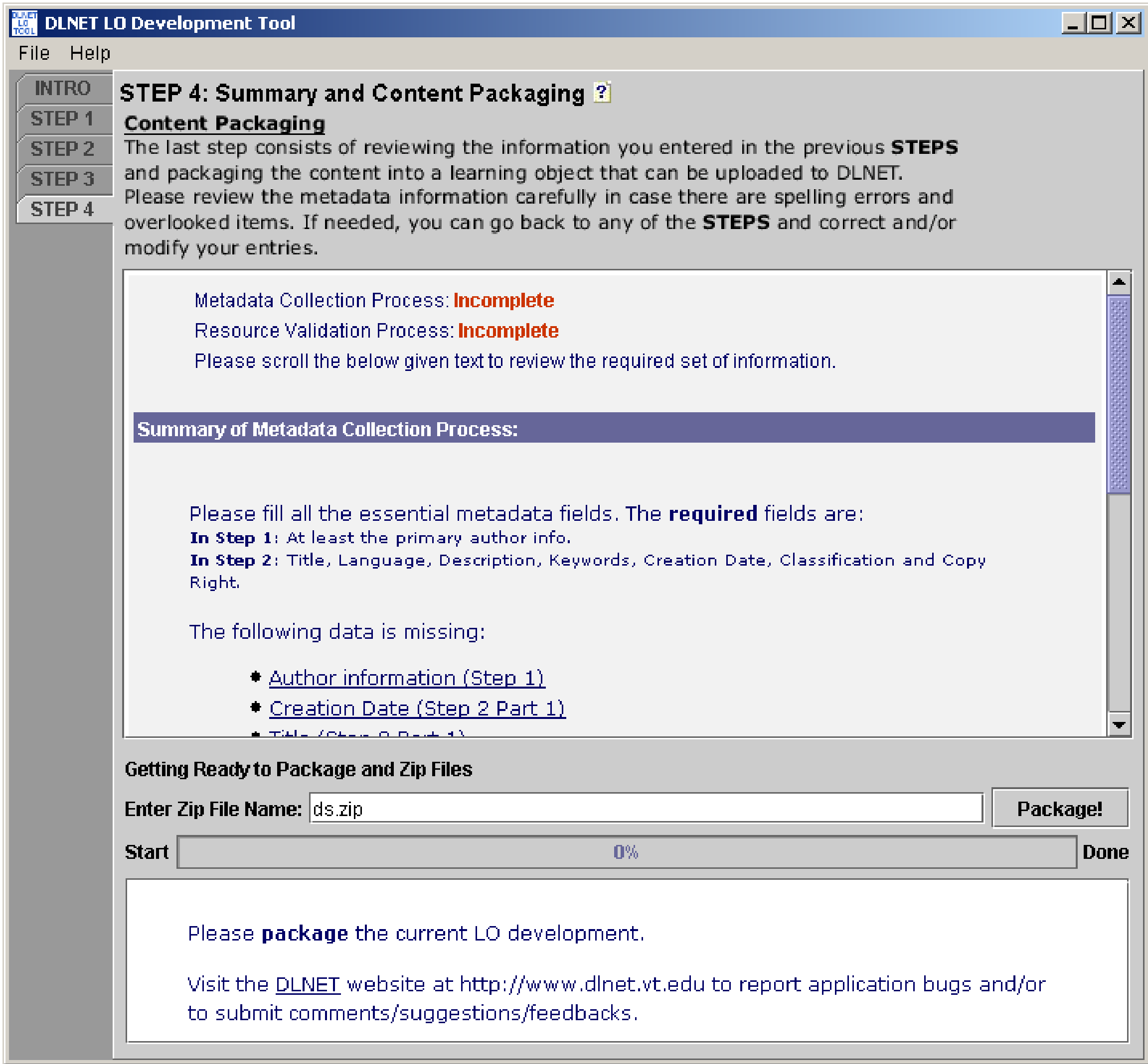



**FIGURE 4.10: LO Packaging Environment within the tool**

The last step consists of reviewing the information you entered in the previous STEPS and packaging the content into a learning object that can be uploaded to DLNET. Please review the metadata information carefully in case there are spelling errors and overlooked items. If needed, the creator at this stage can go back to any of STEPS and correct and/or modify the entries. The **Package!** button in the Tool will create a zipped folder containing the Learning Object. This is the item you upload into the digital library website for review and distribution.

Chapter 5

# The DLNET Architecture

As mentioned in Section 1.4, DLNET aims to promote lifelong learning for engineering community. It was thus a very unique platform where various concepts discussed within the previous chapters (3, 4 and 5) can be implemented and tested with the digital library context. Like any other digital library it can be though of as a large collection of distributed, heterogeneous information managed autonomously by Alexandria Research Institute, Alexandria, VA – part of Virginia Tech, Blacksburg, VA. It further encompasses all four quadrants as discussed in Chapter 2. The learning object (LO) model allows to easy access, control and maintenance of the system. The architecture is part of technology handling and delivering learning programs (Section 1.2). But before, the actual nuts-and-bolts are discussed, it is important to understand the kinds of patron who would be interested in using such a library.

## 5.1 Patron Categorization

Broadly, the categories of people interfacing the library are classified as: the knowledge developers, the reviewers, partners and affiliates, and the learners. Following sections discuss their role and functionality vis-à-vis the library required by each group

### 5.1.1 KNOWLEDGE DEVELOPERS

The knowledge developer; also known as content developers, educators, teachers, etc; are the people who generate resources. The resources containing pedagogy form the content quadrant of the digital library (Figure 2.1). Thus their primary task is resource development. They may create the knowledge via indigenous information privy to them or may compose/compile them from multiple sources. One such source may be the host digital library itself. It is important that knowledge developer have certain expertise in the subject matter of the content. This lends credibility to the resource to be archived within the library.

Having created the (pedagogical) value added resource; the knowledge developer can make use of the tool to package the content for submission to DLNET. As discussed in Chapter 4, the tool has been designed with view that knowledge developer is required to follow minimum set of instructions to convert the resource into LO. The critical step here is collection of the metadata. As this process is subjective and is manually entered, the knowledge developer should be fully aware of the quantity and quality of the data to be filled within the tool. Considerable degree of abstraction has been imparted into the tool, so that the knowledge developers need not be aware of the rather complex XML specifications and syntaxes. The resource validation and packaging are transparent to the knowledge developer. The user of the tool should examine the summary of all actions, before continuing with packaging. At this juncture the knowledge developer may choose to host the content it his/her website or may submit it to the digital library for quality control and distribution. In a proprietary system, it may not be possible to host, view or use the package via ones own distributions network.

### 5.1.2 REVIEWERS

Reviewers are a category of users that have volunteered for quality control of the LOs submitted to the digital library. Like knowledge developers, reviewers should possess certain level of expertise in the subject matter under assigned to them for review. Some of the quality parameter they evaluate is the pedagogical effectiveness, quality, ease of usage, suitability and conformity to the area of submission. Digital library can either solicit subject experts to review the content submitted to it or may present the option to register as reviewers when patrons sign-up with the library. This category can further be split into pre- and post-publication reviewers.

The pre-publication reviewers are the first to judge and rank the content i.e. before the LO is archived with the library. The LO is accessible to them alone to view and mark the content. They are selected by an apt algorithm, which matches the content classification (specified during learning object creation) with the area of specialization of a reviewer (specified during registration). The post-publication reviewer are composed of learners and users who have use the LO for pedagogy. The feedback provided by them can be used to judge the effectiveness and popularity of the LO among the target audience and the learner community at large.

### 5.1.3 PARTNERS AND AFFILIATES

The partners and affiliates of digital libraries such as DLNET, provide an important support function in terms of maintaining, editing and updating the collection. Some of the common functionality provided by them includes content "*collection*" (upload LOs created by third party knowledge developers), cataloging (this is task done by librarians etc), version control (done by system administrator and editors) among others.

In the dynamic world of data mining and data warehousing, many advantages are realized by collaborating digital library efforts with other entities. Pooling resources in such fashions allows serving a wider audience. Important consideration for such partnerships to work is the distribution of functionality and resources. Access control either direct or proxy means is also essential. Currently metadata sharing and copyright is one contentious issue being discussed extensively among the digital library community with regards to sustain viable partnership between various entities within the modern profile of the library.

DLNET has partnered with several such entities, in academic, government and professional arena. The next chapter discuss in detail the nature of this partnership and how such grouping enhances the services provide to the patron.

### 5.1.4 LEARNERS

The learners are the end-users in a digital library comprise of all users seeking a learning experience from the LOs. They are the largest patron group within any digital library. It is essential to provide easy-to-use interfaces for users to register and maintain their profiles; while at the same time interact efficiently with the learning objects. Access control allows providing individually customized services to the user and hence improving the user's experience with the library. On a larger scale this also allows to view and capture essential data on usage and trends in the type of LOs being sort by the patrons. This helps to maintain and monitor the system scalability etc.

## 5.2 Content Categorization

Having analyzed the patrons interfacing with the library, let us now look at the LOs that can be handled by the digital library. As seen in Chapter 3 and 4, all content must be wrapped into an LO. The LO can contain actual physical resource or be a pointer to a resource such as URL link.

The categorization thus made between the LOs being submitted to DLNET. The library has to accordingly handle these LOs.

### 5.2.1 ZIP LO

The ZIP-LO are nothing but resources packaged together with metadata using the LO development tool. They can within themselves be of web-type or non-web supported type (Section 3.5). When the ZIP-LO is submitted to the library, it is essential to check for integrity of the package. Since this type of resource is stored in-house at the library, the system has greater flexibility in handling and processing across various blocks within the library.

### 5.2.2 URL LO

The URL LO is metadata with link to an external web resource. Here the resource is hosted at an external location and is "recommended" by the knowledge developer or content collectors to be indexed with the digital library. This type of LO is useful to enrich the hits returned for a given query, the quality and importance attached to such resources should be watched for credibility and quality control.

Knowing the different category of people who are served by the library, the structure and type of learning object expected as content and the overall flow control (the modern profile) within the various components of the library, we can now envision the full system architecture of the digital library.

## 5.3 System Architecture

The system architecture described in this section is for the custom built DLNET library. It is generalized enough to represent any commonly known models of the digital libraries that are currently in use through the digital medium. Figure 5.1 shows the system implementation.

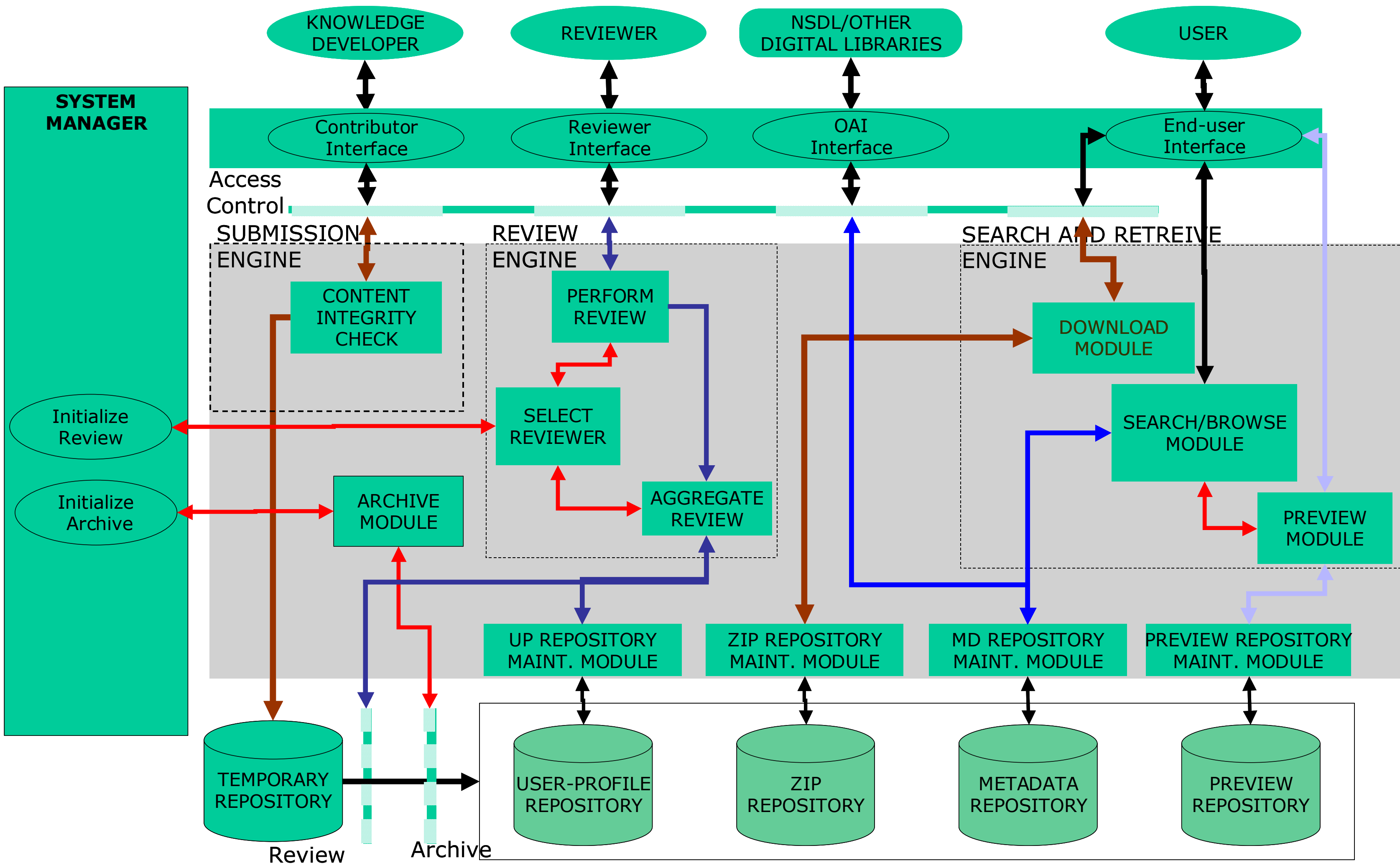



**FIGURE 5.1: DLNET System Architecture**

The system architecture is envisioned to contain four core components: web-server, application-server, databases and resource repositories. All the patrons interface with the library via the web-server. The web-server host the web pages and other relevant feature rich front to seek and retain patrons. Essential parts of the information here include, guidelines for various categories of patrons, instructions on building compatible LOs, registration, privacy and help information.

The application server is the core engine of the digital library. It handles information and maintains flow-control across different blocks within the library. Once the patron login, the application server ensures they are assigned privileges based on their profile. Knowledge developers can submit new LO or view the current status of their earlier submitted LO. Reviews can check if any LOs have been assigned to them for review. Learners can search, retrieve and download LOs of choice. Affiliates and partners can be assign additional privileges of patron services, cataloging and maintaining the collection. System administrators have full control over every aspect of the digital library operation.

Once a package is uploaded at the website and passes content integrity check, the application server moves it into a temporary storage. Here the LO is un-packed, and the metadata and resource information is retrieved. For URL-LO the content integrity step is omitted as the content is capture at the web-server program and thus robustness of the LO is guaranteed. The LO is processed via this location until it is to be published. Only the knowledge developer and designated pre-publication reviewers can view the content during this phase. The review process is invoked by the application-server at regular intervals.

### 5.3.1 REVIEW PROCESS

The review process starts by comparing the characteristics of the submitted LO with the background of reviewers. Some of the parameters checked are the subject classification, recent review load, reviewer’s years-of-experience in the subject matter etc. Based on this match an optimum number of reviewers per LO are assigned (in DLNET the number is selected to be three reviewers per content). In the event not sufficient reviewers are found, the process iterates the next time it is invoked. Note if no reviewers are found for the LO over a long duration, then system administrator is notified to solicit external reviewers for the specified LO.

The selected review at this point may start reviewing the LO. The reviewer is presented with a form containing a questionnaire. The questions relate to quality and pedagogical value of the LO. The answers to the questions would be used to assign a rank to the LO. The feedback from the reviewers is aggregated and analyzed by the review engine and the final rank assigned to the learning object.

If the rank is below a certain pre-determined standard set by the library, then the LO is withdrawn. If the LO is ranked over this standard, it is designated to be archived. The above review process is typical of a ZIP-LO, as LO is hosted on the library's temporary repository. For URL-LOs, the review process is similar but is a striped down in terms of question designated for quality control.

### 5.3.2 ARCHIVE PROCESS

Once the LO is approved for archiving, the system invokes the final archiving process. This can be set a scheduled task that is invoked at regular interval for all the LOs designated for archiving. Archiving involves several sub-processes. The first action is to recreate (or compile) comprehensive and complete metadata information. Now the metadata will contain additional information about review and ranking. This is basically the "cue card" (Section 1.1.2) for the content. The metadata stored in specialized and separate repository, since is the most extensively used component of the LO. It should be a fast access repository implementation. The information from the metadata can be used in two ways: one, directly from the XML file format or second, by extracting selected metadata element and entering them within a database like Microsoft SQL or Oracle 9i etc.

Using the XML metadata file can provide a more extensive option of fields to search the LOs from. But parsing XML metadata files, especially for large volumes of LOs is not scalable and may reduce the efficiency of the system. Using the database entries is an efficient method to handle search and seek requests. However database columns may or may not exists for the fields the users are interested in searching. Hence a combination of the two techniques is required. An initial set of results for search query can be retrieved from the database and the learner is additionally provide with the option to view the metadata of these results to refine their result. This way both broad and in-depth search of LOs is possible to given set of query.

The second step in archiving is the creation of the downloadable LO package. Here, new LO preamble is created to reflect the copyright information and optionally "intelligence" is added to it (Section 3.6.2). Once all the new LO components are created, the LO is package into a ZIP-LO learning module. This allows for easy convenience of downloading. The ZIP-LOs are stored in a special ZIP repository. They can be designed (or selected hardware) to handle high volume interactions for large size LOs.

To improve learner experience with the library and its content, it is useful and helpful to preview the LO before downloading it. Hence in DLNET a special repository is created called the preview repository. Here the un-zipped version of the LO is stored. The learners once decided on a set of LO to preview, the application-server sets the links to the starting file within the selected LO. This link is invoked and LO plays on the learner's browser.

### 5.3.3 CONTENT MANAGEMENT

Content management is the task of application server and performance operation on the content and the related metadata. One of the main tasks is to assist learners is search and recovery of the desired LO. This process should support the natural discovery process of the learners to zoom-in on the LO via provided search elements. In DLNET, three types of search are implemented.

- Simple Keyword Search

  A keyword search is executed by collecting the search query and parsing it for given words. Stop words are dropped and stemming is done wherever possible [23]. Once this is complete a sequence of queries is executed (either in a sequential manner or simultaneously) using sets of stored procedures and results are obtained from the database containing metadata stored as reverse indices. The results of the search are sorted in the order of degree of match and precedence, and presented to the patron. The user then has the option of viewing the metadata, previewing or even downloading the learning object.

- Boolean Search

  A similar functionality is provided for Boolean search (also called advanced search), but with better precision. Currently, DLNET support advanced search based on author, title, description, keyword and format. Here the patron can selectively include, mix and match the desired set of metadata elements for search. This helps in expanding the precision of the search. Additionally this functionality can be expanded to enable search by interactivity type, intended user, context, semantic density, difficulty and other metadata fields.

- Taxonomical Search

  DLNET provides its users the opportunity to browse through its collection using a pre-defined taxonomy. A three-level taxonomy has been designed to sub-classify engineering fields into two subgroups. This provides a unique opportunity to a user to track his learning resource based on taxonomical classification. Future plans include making available a combination of advanced/keyword search along with browse functionality.

The other content management activities include support for downloading the LO. The download functionality is provided in DLNET so as to enable users personalize their learning experience and even to perhaps re-use parts of the learning module in creating new improved learning objects. The requirement however for downloading a learning object is that the downloader must be a registered patron of the library.

Besides direct interfacing with the in-house LOs, content management module of the application server has to manage sharing of the libraries resources with affiliate and partner libraries. The NSDL community comprises of federation of digital libraries including the DLNET portal. DLNET plans to share its collection-level and item-level metadata resources using the Open Archives Initiative (OAI) Protocol [24]. If need be, a handshake will be developed to grant special access to these federations. Partner digital libraries from NSDL domain will have proxy access to DLNET's metadata repository thus harvesting DLNET's metadata.

Some of the other content management activities include version management, content usage statistics collections, converting the content into different formats to make it playable in different medias, monitoring the use or abuse of the system. etc. Due to limited nature of the DLNET project, these functionalities have currently not being implemented.

### 5.3.4 USER MANAGEMENT

Through out the above discussion, access control has been the central issue in patrons interface with the library. When patrons register with DLNET, an XML profile is created that is stored in the user profile repository. This is equivalent to the "library card" in traditional libraries with added security of password protection. Registered patrons can track their activities via information contained in this file. Knowledge developers can view the current status of their submissions. Reviewer can view the LOs available for review and also the ranks of the past reviews. The learners can keep track of their recent downloads etc. All these functionalities are supported by the application-server via the personalized ***My DLNET*** page presented via the web-server.

## 5.4 Hardware and Software

DLNET's web services are powered by Apache's Tomcat Service running on Microsoft's IIS web server. The front end and user interfaces are provided by a combination of cutting-edge technological elements. Java Beans, JSP and Java Servlets are the primary components that provide front and back-end support. In addition a powerful database service (SQL Server 2000) aids in information retrieval. Use of XML requires extensive usage of java based XML parsers and language transformers.

Chapter 6

# Digital Libraries – "The Big Picture"

As seen from previous chapters robust methodologies, architectural models and hardware/software exist that allow fast time-to-conceive and hosting of digital library. Due to their growing popularity, DL projects are being initiated by many entities across a wide spectrum of discipline. This chapter summarizes these activities and attempts to foresee the future of this endeavor and related pedagogical model.

## 6.1 Current Academic and Commercial Activities

The drive to build digital libraries, both in public and private sector, is motivated by the advantage of the technology to offer continuous education, training and pedagogical services at all levels of users.

### 6.1.1 NSDL AND THE DLNET

This paragraph is quoted from [17]. It details the vision and implementation of NSDL.
The National Science, Mathematics, Engineering, and Technology (SMET) Education Digital Library (NSDL) has been conceived, and is being constructed, to support excellence in SMET education for all Americans. The NSDL will be a comprehensive information system built as a distributed network. The NSDL will develop and make accessible collections of high-quality resources for instruction at all levels and in all educational settings. It will also establish and maintain communication networks to facilitate interactions and collaborations among all SMET educators and learners, and will foster development of new communities of learners in SMET education. Multiple services will be available to help users effectively access and use NSDL resources. The potential impact of the NSDL on the nature of SMET education is great. Building the NSDL to meet the diverse needs of its users presents many challenges, including:

- Developing a shared vision for the form and function of the NSDL.
- Meeting the needs of diverse learners.
- Meeting the needs of the many disciplines encompassed by the NSDL.

- Acquiring input from the community of users to ensure that the NSDL is both used and useable.
- Developing a governance structure and Core Integration System that balances community needs with technical applications.
- Integrating technologies that already exist, and looking ahead to the development and integration of new technologies.
- Providing mechanisms for sharing knowledge and resources, and building cooperation among NSDL collaborators.
- Evaluating NSDL and its impact on SMET education.
- Coordinating these many interests and functions to provide an integrated whole.

Significant progress has been made in developing technology fundamental to a successful digital library such as interoperability, authentication, persistence, and archiving functions. Building on this base, the National Science Foundation initiated the NSDL effort with a program solicitation (NSF 00-44) to establish an educational digital library for science, mathematics, engineering and technology,

> "To catalyze and support continual improvements in the quality of science, mathematics, engineering, and technology (SMET) education, the National Science has established the National Science, Mathematics, Engineering and Technology Education Digital Library (NSDL). The resulting digital library, a network of learning environments and resources for SMET education, will ultimately meet the needs of students and teachers at all levels—K-12, undergraduate, graduate, and lifelong learning—in both individual and collaborative settings."

The DLNET (Digital Library Network for Engineering and Technology) is a project funded by the National Science, Technology, Engineering and Mathematics Education Digital Library (NSDL) initiative of the National Science Foundation under NSF Grant DUE-0085849. It is a repository of LOs as well as a platform for information discovery, interaction, content-building and distribution that will support pedagogy and learning in Engineering and Technology. DLNET is a collaborative effort of four institutions: American Society for Engineering

Education (ASEE), the Institute of Electrical and Electronics Engineers, Inc. (IEEE), Iowa State University and Virginia Tech. The library itself is hosted at the Alexandria Research Institute (ARI) of Virginia Tech. Under NSDL, the mission of DLNET is to offer a roadmap, including implementation strategies and building blocks that will assist in the lifelong learning of engineering faculty and students, practicing engineers and technical professionals. In this regard, DLNET has been developed as a platform for hosting engineering and technology related content targeted at the practicing engineer and technologist. The platform provides tools for collecting reviewing, validating and cataloguing new materials. It also provides a portal through which contents can be both viewed and accessed. The portal also provides a gateway to education and research materials published by universities and professional associations in the various engineering disciplines including the IEEE and ASEE. More importantly, DLNET serves as a partner in the federation of digital libraries under NSDL. The federated structure is illustrated in Figure 6.1.

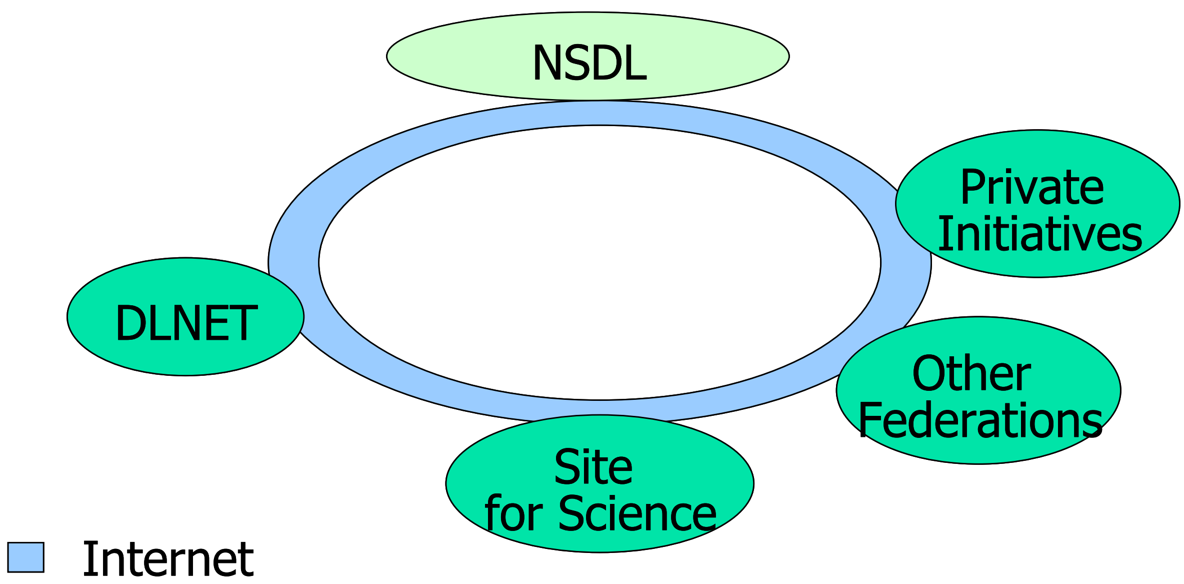



**FIGURE 6.1: Digital Library Network – "The Big Picture"**

As envisioned by NSDL, there exist many digital libraries that work autonomously, but follow known standards of operation. All digital libraries may "talk" to each other via the Internet. Patrons of one library may search metadata of another library and if arranged can also access content from other library.  This type of integration allows cross disciples and high quality education via one portal.

### 6.1.2 PRIVATE INITIATIVES

The majority of private initiatives to build digital libraries come from for-profit companies. The world's largest computer maker IBM estimates a world e-learning market of $43 billion in year 2004 [25]. These companies develop either full-installable models of digital libraries or provide support services based on the quadrant their customer fall in (Figure 2.1). There are two approaches to developing a full-scale digital library model. As many standards are still at works, some companies such as Blackboard and Web-CT have commercialized proprietary systems that handle all the required functionalities and protocols for the library. Others (like IBM) have developed a generalized model based on most popular metadata standard for digital library and currently are selling commercial products based on these models.

The imperative among various organizations is to develop a business model that supports all the quadrants of a modern digital library. In this situation, a need for standardization becomes apparent. Slowly but steadily, consensus is emerging on need for common architecture, services, protocols, data models and open interfaces.

## 6.2 Future Activities

Digital library is a multidisciplinary effort. Thus the future work and subsistence of digital library will be influenced by both technical and non-technical paradigm. In the non-technical arena the type of knowledge, the way people choose to interface to gain that knowledge, the dynamics or need for information will play critical role in shaping the patron interface and content management modules in the library. In the technical arena, new application, hardware and software developments are shaping the inspiration, efficiency and sustainability of digital library. In the immediate future, though standardization work is the most critical.

Some of the areas where additional work may be required are, in handling and sharing of metadata among various quadrants of the library. Currently metadata is view as a proprietary commodity. This is well suited in an integrated business model where all the functionalities are supported under "one roof" by a single organization. But for the growth and proliferation of this novel technology, a certain degree of freedom of use and sharing of the metadata is essential. As noted in Section 2.2 of the modern profile of digital library, given permission the learner should be able to easily transfer their "credits" of using certain LOs to one or many educational managers for certification or degree. Similarly, knowing some basic characteristics of the learner helps content banks to target or "advertise" their product to the right people seeking education. Similarly, the knowledge developer should have full understanding of various metadata elements set forth by the education managers so as to develop good quality content.

Currently many digital libraries are collecting many kinds of metadata, regardless of the ownership of the content. Within DLNET itself, two kinds of LOs are accepted: the ZIP-LO is hosted by DLNET and another is URL-LO that contains pointer to third party LO. The author of this thesis perceives that the second category of LOs would eventually become redundant. This is so since digital libraries are more than merely search engines. Search engines like Google, are vastly efficient in seeking and targeting third party content for their users. The aim of digital libraries is to provide education. Hence there should be a traceable "learning" component associated with the LO. Thus the author believes that a structured approach, however loose, to digital library such as NSDL (Section 6.1.1) is bested suited in terms of overall learning management system. If a certain digital library cannot fulfill the patrons requested course or information need, it can pass the request on to the next level up in the hierarchy of the library where then the patron can be pointed to the appropriate library that can serve the request. Thus in this process the patron actually "learns" and can get "credit" for gaining such education.

With regard to ownership of content in digital libraries, copyright transfer is the key. Copyright can be set in three forms: free with no copyright restrictions, free with copyright restrictions, and payment-based resource. Free with no copyright restrictions means the resource can be used by one and all as they please. Free with copyright restrictions means the resource can be used by everyone, but in its original format and with due recognition given to its creator. Payment based resource are proprietary in nature and one has to have privileged access for using it. For non-profit public domain digital libraries, the first two forms of copyright are most suitable. It is envisioned that future digital library tools will allow patrons to choose, maintain and track their copyright preferences via the appropriate metadata element. Majority of resource (knowledge) developers would tend to choose the 'free with copyright restrictions' as it allow gaining recognition without legal hassle of fighting infringements.

The digital libraries of the future would thus be more than content repositories, but would be full-fledged learning support system. Teachers can independently create content and seek copyright. Learners can seek education in what they want and when they want and get tested for it. Thus the glue between content and metadata – the LO should be robust.

In the future it is expected that the physical nature of the LO and the supporting tools would dynamically upgrade based on user input and available technologies. But the core concept of the LO and the need for such model is expected to remain constant. Regardless of the underlying digital library architecture, the content and metadata has to move freely between various entities interfacing with the library. Technological changes may allow to develop a faster and efficient wrapper or alternately may eliminate the need for support files with the LO file structure. Given time, a more comprehensive tool sets can evolve from the LO development tool as discussed in Chapter 4. This, say "Learning Assistance Tools Set" may contain not only the LO development tool but may also have course builder/binder, LO to third party resource packagers, metadata crosswalk software and copyright/royalty management application. For learners, they may eliminate the dependency on browser and have a machine independent platform for viewing LOs. Additional plug-ins may allow to test the learner based on the LOs currently available in their folder.

It is expected that the pedagogical endowment and learning management model will be further refined to accommodate the changing nature of request from patrons and the nature of the content. The core aim of the model will always be to facilitate learning – anywhere, anytime and for life.

With advent of new technologies like virtual reality (VR), we may actually come back full circle as learners may wear an "avatar" and choose to learn by attending (read logging in to) the digitized classroom set in the virtual space. Once in the “class” they may be presented with different subject options; speed, difficulty or cost of LOs associated with each etc. Through their VR suite, they may actually be able to flip through the LO “pages” and ask real-time question to the “teacher”. The teacher themselves may be real person with an “avatar” or humanized equivalent of an “artificially intelligent” machine. The classroom may be setup by a third party content host who has high bandwidth and application intensive servers that support such rich VR type classrooms. Once the “class” is over, the learners may then login to a secure examination VR environment of a university and give an exam to earn their degree.

Chapter 7

# Summary and Conclusions

E-learning is expected to generate $11 billion in revenue in by 2005 in the corporate training alone [25]. As more and more individuals and corporations realize that education is for life, e-learning is emerging as THE solution for pedagogy. The advantage in terms of convenience and cost saving are just too many to ignore.

## 7.1 Summary

E-learning can be achieved via many ways and through many institutions. Traditionally, education has been imparted in classroom from teachers who give lectures and use slides or handouts as teaching aids. This, though efficient, has limitations in terms of flexibility and convenience of education. Pupils cannot learn via alternate means or at other times. With the advent of Internet technology, all this has changed. Internet allows the flexibility of providing greater variety and additional means of access for education. The "classrooms" in this medium are always "on". Pedagogical endowment has thus evolved to fit the changing needs of their practitioners and the supporting medium. In the modern paradigm, all components that enable education can work independently and yet support learning and testing of its patrons. This paradigm is the digital library. Despite the name of library, it is more than a repository of content; it is learning management and support system.

Many models explaining digital library exist. They generally are "people oriented" i.e. saying from teacher-centric to learner-centric etc. In this thesis, a new and unique model of digital library is proposed. It is metadata centric. Metadata is information about entities. A metadata centric architecture allows for maximum flexibility and scalability while offering custom solution for all the people and entities interfacing with the system. There are basically three types of metadata that any digital library handles: content metadata, user metadata, and pedagogical metadata. They are like "cue cards and library cards" in traditional library model. Content-centric metadata deals with content description and packaging and are used while searching for the content. Content can be a physical resource or links to these resources. User-centric metadata deals with the creation and maintenance of user profiles. The users are categorized as: knowledge developers, reviewer and learners. Pedagogy-centric metadata deals with evaluation of learning and result tracking. The value of elements in the metadata is used to initiate actions among various entities within a digital library.

The proposed model that uses these metadata elements is explained in Chapter 2 and illustrated using Figure 2.1. The groups forming the library are: the knowledge developers, the education managers, the learners and the content. The model represents the objective modularity of the underlying digital nature of relationship between various parties involved in education. It is clear that the interaction between various quadrants occurs via actions invoked by values in the metadata elements. As long as each quadrant can access the desired metadata, they can exist mutually independent of each other. For example, universities do not necessarily need to maintain content banks. It may be more efficient for other third-party entity (possibly a commercial database company) to maintain such content banks. What is required, is that the third party implement the metadata standard set by the university, when accepting content from authors (possibly pay royalty in return), following which it should share (possibly for a fee) with the affiliate university. The university should then, based on the quality of the metadata, set requirements of certification.

The central role in all these activities is played by content-metadata packet. For this we have proposed a "learning object (LO)" model. The model shows how learning (and resources associated with it) can be transformed from teacher-centric to learner-centric education, and how content can be transformed from multimedia resources (application or technology specific data) to metadata-based learning modules. The LO model can be easily molded to use any known and popular metadata and/or content packaging standards. The features and complexity are functions of the owner's (digital library's owners/operators) choice of these standards, thus making it versatile and inherently scalable.

The LO model encompass two aspects namely, "learning" and "object" with the underlying theme being "ownership" and "quality". They are like the "book" in traditional library. Learning refers to education. The purpose of the learning object is to hold information that enables learning. The object aspect of the LO relates to actual physical resource (the digital/electronic files). Since digital construct can be of any size and format, so are the objects within the LO. The LO is an immutable entity and it may not be changed or altered by the user. It must function independently (self-contained and standalone). Ownership and quality are the requisite attributes of any resource within a library. Ownership relates to the intellectual property of the LO. The quality implies the accuracy and the pedagogical value of the learning object. The characteristics of the LO are subjective of the quality of the LO. Certain common characteristics of the LO are: objective, pre-requisites, interactivity type and level, type of end-user, context of use, difficulty, learning time, etc. This definition of learning object is quite unique as it allows many flexible yet scalable abstractions for learning. Additionally, it can be extended and made intelligent (Section 3.6) as desired for creating feature rich implementations.

The physical file structure of the LO model is based on hierarchical ("tree") file structure. This is one of the most configurable and scalable file representations. The top package is split into four groups (see Figure 3.4 for illustration): the HTML preamble, the XML metadata file, support files, the actual resources. The concept of the preamble or the wrapper is unique to the LO model. Unlike in a physical book, where one knows that the appropriate side of the cover is the starting point which is then followed by a table of content and then the text, digital content as such inherently don't have any structured starting or continuation point. Asking authors to do this job is tedious and automatic chronology extraction tools have so far been not quite successful. Thus the need for the HTML preamble – a new and unique concept, proposed and implemented as part of the thesis work (Section 3.4.2). The LO's preamble is the starting point from where the LO will always "play". When an LO is embedded into another LO, the same structured is preserved, but it is placed within the *Resources* folder as it is "object" for the upper level LO.

In order to test the LO model, as part of the thesis work, an LO developed was built. It collects pedagogy and copyright metadata relating to the resources, validates the resources and packages it into an LO. As high quality of metadata is required (later for quality control and high-yield discovery), the mixed approach to allowing creator to fill-in some metadata areas, while selecting others is judged to be most optimum approach to metadata collection. The metadata is stored in IMS based XML metadata standard. The XML in essence provides a way to bind data into a know format that is easily readable by machines. After metadata collection, the next important task is to verify the "structural integrity" of the resource that form the LO. This is the resource validation process. Validation assures that no broken links will exist within the file structure or organization by parsing through the resource folder. Once validation is complete, the information about the organization (or structure) of resource is appended to the XML metadata file in conformance with IMS-IEEE content packaging standards. The package now is the LO.

The layout of the tool is designed with mind that many authors are contributing material voluntarily and may have varying adaptation-curve or aptitude towards learning new applications. The tool takes them through a step-by-step process of resource to LO transformation. The tool is implemented as client-based application and is code in JAVA, making it independent of the underlying operating system. The LO is now ready to be uploaded into the digital library for review and wider distribution.

The objective of the Digital Library Network for Engineering and Technology (DLNET) is to develop a specialized collection of engineering and technology-related content targeted at the practicing engineer and technologist [12]. Author has taken part in implementing a limited model of the modern pedagogical endowment for the DLNET. The corresponding LO model was also implemented for flow-control and user/system interactions. The model basically allows content creators to "package" resources for the educator to the learners. The core metadata in this case is the content-centric metadata. Once the LO is created, the knowledge developer login to the DLNET website and uploads the LO. The LO is then separated into educational content and the metadata. Members of the peer group review the content and metadata, and upon passing, the metadata is added to the metadata repository and content is added to the content repository. Public can search or browse through the metadata at anytime. But, they are required to login to download the content. This way a user-profile database is developed. As of now no plans to certify are envisioned, but can be easily implemented to complete the digital library picture.

Apache's Tomcat Service running on Microsoft's IIS web server provides DLNET's web service. Java Beans, JSP and Java Servlets are the primary components that provide front and back-end support. In addition a powerful database service (SQL Server 2000) aids in information retrieval. Use of XML requires extensive usage of java based XML parsers and language transformers.

## 7.2 Conclusions

A digital library can be successful only if the underlying system architecture is robust. While the hardware and software aspects play a major part in implementation, it is the interaction and functionalities that are provided by the entities associated with digital library, that will allow efficient pedagogy. The modern profile of the digital library is such that these entities, while acting independently, have a unifying goal to sustain the paths to knowledge from the educator to the learner. It is modular and hence more flexible in meeting the changing landscape of education endowment. Furthermore, digital libraries must themselves be scalable to integrate various libraries representing different disciples together. All this is possible only because of the central and critical role played by metadata. The educational content is processed based on the values of elements in the metadata. Thus, metadata and content should inherently be grouped together by association. The learning object model defines such an association.

The LO model should in itself be easily accessible across various quadrants of the digital library architecture. The proposed concept of "learning object" helps authors and users alike in streamlining the learning program development process. Its physical representation is hence standardized. The system designer, based on the available underlying hardware and software capabilities, can decide the metadata and packaging formats. It, thus, integrates the digital library abstraction into the actual digital media.

### 7.2.1 RECOMMENDATIONS

Digital libraries such as DLNET, besides providing core functionalities, can retain and enhance the user experience by supporting add-on feature like smart browsing, efficient user profile management, intelligent collections compilation, etc. Smart browsing uses correlation data between what kinds of LOs are downloaded and who downloads the LOs. By tracking these events, the library may “suggest” LOs for download to patrons who are interested in similar topics. Thus providing targeted services to subject specific communities – enhancing user experience with the library. Currently DLNET tracks who downloads what LOs. This information can then be mapped with the already present data about the nature of LOs archived in DLNET. By noting the grouping of LOs, “suggestions” may be made to patron of other related LOs they may be interested via the search website. In terms of coding, the LO identifiers within the user profiles would be compared, in real-time, together with the LO identifiers of the subject the user is interested in for display. The extension suggested about is reasonable to implement and should be considered for future version DLNET.

An efficient user profile management helps to build a strong learning community around the library. A digital library is more that hardware/software realization of education. It requires seamless interaction between various people – the teacher, the learner, the content developer, the learning manager, etc. Efficiently correlating and bringing the right group of people and content together is essential for proper pedagogy. This can be achieved via strong and robust database and metadata management tools. With regards to LO, an intelligent utility (as discussed in Section 3.6.2) is reasonably easy to implement, both on the client and the server side. This is simple JAVA and JavaScript based coding, where HTML links can be dynamically set to Servlets that perform the requisite action on the server to meet the user’s needs. For example, to search for other LOs by the same author, a HTML link to the ‘advanced search option’ Servlet in the DLNET website can be invoked to perform the task. The results can then be rendered in the LOs’ top frame for viewing. More advanced options such as; direct secure client-server interaction via HTTP or telnet should be explored for better integration of user actions with the system. For this, a more thorough network support protocol needs to be implemented. Here the client application will initiate an Internet connection to the DLNET server. The DLNET server will then authenticate the client and allow privileged information to pass to client. The patron can then render this information at the client site for custom view. Though user profile management can be done both at the client or the server –side of the architecture, it would be more secure if data were stored on the server.

The unique style of the LO proposed here, is a strong candidate for standardization of LO with digital libraries across various disciplines and architectures. The type of file structure and a fixed LO preamble format allows easy use among all kinds of patrons. Also from a systems perspective, the model can be served interoperable among different libraries. It can easily be created, broken up, quality-checked, transmitted and reassembled at the destination. These ideas are unique to this model and should be further investigated, via the recommendations (and their implementations) made above, to extract additional mileage in terms of hardware/software usage, user-friendliness, inter-disciplinary interoperability, etc.

A beta-version of core LO concepts and supporting learning management system has been successfully tested with volunteers and implemented as a part of the DLNET project. Above all, digital libraries, like the DLNET, are evolving to be more than just a repository of learning materials. Digital library patrons can judge and selectively choose what they want to learn about and then have a system that supports that chain of learning. Hopefully, in the longer run these ideas will promote other digital libraries to standardize not only the development of learning programs, but also the overall system architecture so as to allow easy across-the-board acceptance of digital libraries as a credible and qualitative source of education - anytime and anywhere.

# Appendix

## A: DLNET LO Development Tool - JAVA Coding

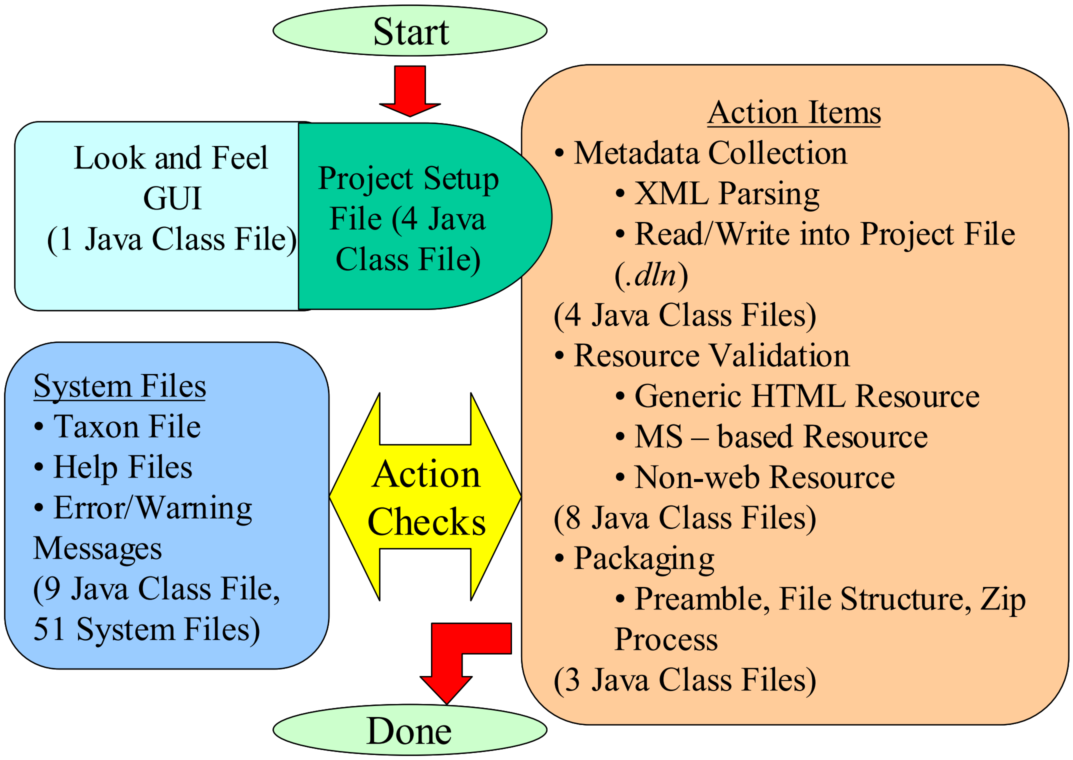



**FIGURE A: LO Development Tool JAVA Code Structure**

Figure A illustrates the coding structure and succession of actions as handled by the tool software. The user is expected to double-click to initiate the tool. This reads the GUI file and renders the GUI. Once the tools (Figure 4.3) is seen, user can start actual LO creation process. For this a project file needs to be set (or a pre-existing one read) into the system. This project file is checked for integrity and compatibility with the system. For example if it in XML format, if another file of such name exist at the prescribed location, etc. Once initialization is over, core tool task can begin. These are metadata collection, resource validation and packaging.

Metadata collection involves collecting the values provided in the designated fields and writing them to the project file. Metadata is collected in batches: user data, general data, educational data and classification data. The values are written as the tool progress from one set to next.

Author can edit them at anytime as long as the values are "applied" after each set. Another aspect of metadata collection is the check that ensures all essential metadata is collected. A special universally defined Java file sets (and keeps track of) variables that are required and prompts the author if any are missing. Since the metadata is saved in XML file format, a special Java file was coded to ensure correct file interface for reading and saving the data.

Resource validation is an automated process. The top resource validation Java file takes the index HTML file and parses the HTML code and then recursively goes through linked files for collection. If the top file is in propriety format like Microsoft, a different Java file is called, that parsers the collection based on the resource setup within the content. For non-web resources, the validation ensures the file exists in the said location. The master list of validated files is maintained for eventual packaging.

Once the above two process are successfully complete, packaging can begin. This process is initiated by clicking the appropriate button in the GUI. The user can choose to assign a unique name for the package or let the default name be set by the tool. The packaging begins by collecting the data in the metadata project file and creating the LO preamble from it. The LO preamble is a HTML file. Following which the navigation bar for the bottom frame is created. Then the tool creates a virtual file structure. The appropriate files are set within this file system. The preamble file and the resource files are then set in a zip file based on the virtual file system in the tool's memory. In this step the packaging essentially engineers the links in the fashion that files are placed within the zip package based on their virtual file location in the file architecture. Hence when the file is unzipped the LO spills out in the file structure set of it.

At this point the LO creation is done. The author may then choose to close the project and start a new one or exit the application. Note the author can stop at any time during the above processes and restart the project at a later time – as long as the project file is saved up to that step.

# List of Acronyms

| | |
|---|---|
| ARI | Alexandria Research Institute |
| ASEE | The American Society for Engineering Education |
| DC | Dublin Core |
| DL | Digital Library |
| DLNET | Digital Library Network for Engineering and Technology |
| IEEE | The Institute of Electrical and Electronics Engineers, Inc., |
| LMS | Learning Management System |
| LTSC | Learning Technology Standards Committee (chartered by the IEEE Computer Society Standards Activity Board) |
| NSDL | National Science, Mathematics, Engineering, and Technology (STEM) Education Digital Library |
| STEM | Science, Technology, Engineering, and Mathematics |
| VT | Virginia Polytechnic Institute and State University (Virginia Tech) |

# References


[1] Learnframe, “Facts, Figures and Forces Behind e-Learning”, *http://www.learnframe.com/*, August 2000.

[2] Ubel R., “Engineers turn to e-learning”, *IEEE Spectrum*, October 2000, pp. 59-63.

[3] Mallak L. A., “Challenges in Implementing e-Learning”, *Portland International Conference on Management of Engineering and Technology*, 2001, pp. 298-299.

[4] Ozsoyoglu G., Balkir H., et al., “Electronic Books in Digital Libraries”, *IEEE Proceedings on Advances in Digital Libraries*, 2000, pp. 5-14.

[5] Saul C., “An Overview of Online Learning”, *http://www.lakewoodconferences.com/*

[6] Google, *http://www.google.com/*

[7] Gnagni S., The University Business Magazine, *http://www.universitybusiness.com/*, February 2001.

[8] Adam N. R., Atluri V., et al., “A Dynamic Manifestation Approach for Providing Universal Access to Digital Library Objects”, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 13, No. 4, July/August 2001, pp. 705-716.

[9] Wiley D., “Peer-to-peer and Learning Object: The New Potential for Collaborative Constructivist Learning Online”, *IEEE International Conference Advanced Learning Technologies*, 2001, pp. 494-495.

[10] Carchiolo V., Longhew A., et al., “Learning through Ad-hoc Formative Paths”, *IEEE International Conference Advanced Learning Technologies*, 2001, pp. 96-99.

[11] The National SMETE Digital Library, *http://www.smete.org/*

[12] Anido L., Llamas M., et al., “A Distributed Object Computing Approach to E-Learning”, *3rd International Symposium on Distributed Objects and Applications*, 2001, pp. 260-269.

[13] Mukaiyama H., “Next Generation Digital Library – Architecture and Implementation”, *International Conference on Digital Libraries*, Kyoto, 2001, pp. 105-112.

[14] Anido L., Llamas M., “A contribution to the e-learning standardization”, *2nd IEEE Conference on Standardization and Innovation in Information Technology*, 2001, pp.295-308.

[15] DLNET Team, "Learning Object Definition", *http://www.dlnet.vt.edu/*, *White Paper*, October 2001.

[16] DLNET Team, "DLNET Schemas for Metadata and Content Packaging", *http://www.dlnet.vt.edu/ White Paper*, November 2001.

[17] "Pathways to Progress: Vision and Plans for Developing the NSDL", March 20, 2001.

[18] Wason T. D., "Dr. Tom's XML Guide", *http://www.imsproject.org/drtomxml.html*

[19] Dublin Core, *http://dublincore.org/*

[20] IMS Project, *http://www.imsproject.org/*

[21] IEEE-LTSC, *http://ltsc.ieee.org/*

[22] SCROM, *http://www.adlnet.org/*

[23] Korfhage R.R., "Information Storage and Retrieval", *John Wiley and Sons Publications*, 1997.

[24] Open Archives Initiative (OAI), *http://www.openarchives.org/*

[25] eLearning News, *http://brainware.tv/elearning_news.htm*

# Vita

## EDUCATION

2000 – 2002 **Master of Science,** Electrical Engineering
*VIRGINIA TECH, Blacksburg, VA*

- Summary: Thesis provides a complete functional and business model for digital library implementation. Also illustrates a unique learning object model for e-learning. Other projects include; study of CDMA 19-cell network performance analysis with improved front-end base-station filtering using high temperature superconductor filters, develop understanding of different logic blocks within a wireless system via VHDL and SystemC modeling
- Awarded second place in the 18$^{th}$ Annual Research Symposium 2002 (Engineering and Physical Sciences category) sponsored by Graduate Student Assembly
- Publications:
    - Shankar Mahadevan, "CDMA Capacity Improvement With HTS Receiver Front-End", The Propagator, p11-14, Spring/Summer 2001, Blacksburg VA
    - Shankar Mahadevan, Dr. Saifur Rahman, Dr. Peter Weisner, "An Online Process for Posting and Meta-tagging Learning Objects and Integrating Contents into Digital Libraries", Frontier in Education 2002, November 2002, Boston MA

1995 – 2000 **Bachelor of Engineering,** Electrical Engineering
*CARLETON UNIVERSITY, Ottawa, Canada*

- Summary: Projects in logic programming, wireless systems analysis and integrated circuit design
- Twice nominated to the Deans' Honors List
- Winner of A. D. Dunton Scholarship for excellent academic performance

## WORK EXPERIENCE

Aug 2000 – Jun 2002 **Graduate Research Assistant**, Full Time
*ALEXANDRIA RESEARCH INSTITUTE, Alexandria, VA*

- Implemented a 19-cell CDMA system, for system analysis performance evaluation
- Built the Digital Library Network for Engineering and Technology (DLNET) for web publishing

May 2000 – Aug 2000 **Summer Student Position**, Full Time
*CONEXANT SYSTEMS, Ottawa, Canada*

- Performed circuit translation from low-speed to a low-power, high-speed CMOS technology using Verilog HDL and PERL scripts
- Did IC design, synthesis and simulation using Synopsys for Bluetooth applications

May 1998 – Aug 1999 **Internship Student Position**, Full Time
*NORTEL NETWORKS, Kanata, Canada*

- Implemented Quality of Service (QoS) algorithms for an IP gateway using assembly language; allowing the user-control of TCP/IP traffic on the network
- Designed and implemented a server-to-server FTP client and a network management tool in JAVA for configuring port based vLANs and spanning trees for local area networks

Jan 2000 – April 2000 **Teaching Assistant**, Part Time
*CARLETON UNIVERSITY, Ottawa, Canada*

- Undertook tutorial classes for "Switching Circuits" – an undergraduate digital design course
- Lab duties involved student supervision and correcting lab reports

## INTERESTS

- Hiking, camping, white-water rafting, rock climbing, soccer, chess
- Web publishing and design, home-improvement, reading, traveling