Chapter 5
Management of Digital Resources

5.1 Introduction

The high growth rate of electronic publishing has led to the growing collection of digital resources of various formats in the LICs. The growth rate is reflected in a study namely “Digital Universe Study Report” by International Data Corporation (IDC)-2010 where it was found that, in 2010 the estimated size of digital universe is 1.2 million Petabytes or 1.2 Zettabytes and it is projected to be 35 Zettabytes in 2020 (International Data Corporation, 2010). Users also used to prefer the digital resources over the printed resources especially in Science & Technology. To cope up with changing environment and to meet the users’ demand, the LICs have no other option but to include digital resources in their collection. In this changed over situation, Library & Information Science (LIS) professionals have faced immense challenge while managing these resources. The LICs and the promoters of digital resources are facing problem in building and maintaining the technology infrastructure necessary to deliver digital resources due to continuous changes in hardware, software, and network technology.

Selection and acquisition of digital resources have become complex due to various reasons including complex pricing policy, ignorance about the new digital resources due to rapid growth of digital resources. The publishers, aggregators and the other vendors offer endless variation in their digital resources subscription packages. Though the emergence of consortia and its implementation have assisted the LICs with collection building of digital resources, this system also has problems like turn down to deal with individual library requirement. The emergence of Open Access (OA) digital resources has made tremendous impact to the users and the LICs. No doubt users have benefited from the use of the OA digital resources but the management of these resources makes things more challenging to the LIS professionals. Organization of digital resources is another area to be taken care of by the LIS professionals for the proper use. Besides the description of the digital object, categorization, indexing, the issues like storage,
managing databases of the resources, searching, browsing and retrieval have to be taken care of. The LIS professionals have to administer various issues of access management like access control, content security, object identification, license issue, user id and password management, proxy/IP authentication management, etc. Management of the digital resources does not come to an end here. For continuous and future use, easy and timely access and protecting the original digital resources digital preservation is important. Digital preservation is a tough job for the LIS professional because of the use of new storage devices, standards and protocols of file formats, compression techniques, etc. The major problem is the continuous change over from one stage to another in both hardware and software. Discussion has been made here covering various issues of digital resources management.

5.2 Acquisition of Digital Resources

There are different types of digital resources based on creation process, granted access to users and the content of the resources. On the basis of creation process, there are born-digital, hybrid and converted digital resources. Based on granted access to users, we have subscription based and Open Access (OA) resources and they again can be classified according to their content. In this context, digital resources like e-book, e-journal, ETD, e-database, e-zine, etc. can be included into the categories subscription based and OA resources. To select and collect these digital resources of different formats is not an easy task. The steps in the acquisition process are discussed below.

5.2.1 Selection of Digital Resources

The acquisition process of digital resources involves the identification, evaluation and decision of selection of the resources. The first step i.e. identification of the resource is done to avoid duplication of the digital resource. The publishers, aggregators and other subscription agencies used to offer free trials of their products to the LICs and its users. The LICs can take this opportunity to examine and evaluate its products freely for a limited period. Sometimes, the vendors used to come to demonstrate their products in the LICs and the professionals can find out details about various features of specific product(s). Other sources to identify digital resources are publishers’ website, Web-
OPAC of different Organizations like IndCat of INFLIBNET (<http://indcat.inflibnet.ac.in/indcat>), Worldcat of OCLC (<http://www.oclc.org/worldcat>), reviews of various digital resources, discussion forum etc.

After identification, evaluation of the digital resources is done to decide whether to procure the resources for the collection of the particular LIC. The authoritativeness, cost, technical requirement, currency of the resources can be scrutinized in the evaluation stage of the resources. In case of printed resources, the criteria like authority, currency, target user, subject coverage, etc. are checked. While talking about evaluation criteria, Joshipura (2008, p. 51) in “Selecting, Acquiring, and Renewing Electronic Resources” wrote “However, with e-resources the selector must consider additional elements such as easy access to the content, coverage, search capability and functionality of the interface; quality of technical support; method of pricing; and provisions of licensing agreements”. According to her, the following evaluation criteria can help selecting suitable digital resources:

**Content:** Here, the subject coverage, authoritativeness, accuracy, completeness, illustrations, of the resource is reviewed. The probable duplication is also verified.

**Currency:** The frequency of update, presence of updated information, existence of outdated connections or invalid connections are verified.

**Cost:** The various pricing models are checked for the pricing options. According to the need and demand, the appropriate pricing model is selected to determine the cost.

**Technical Support:** It is important to ensure compatibility of the digital resources with existing hardware and software. Training for professional staff of the LICs helps in facilitating for the users before selecting a particular digital resource.

Open Access (OA) resources which are available free for users can be included in the digital collection as many of these resources have great research value. The various types of OA resources available are - OA journals, resources of the different IRs of the institutions of higher education and resources, websites of individuals or organizations,
individual blogs, professional discussion forums and many more. But for the sake of increasing our collection, we should not select OA resources which are of little research or academic values for the users of a particular academic institution. These resources should also be evaluated before inclusion in the digital collection. Besides the criteria like content, cost, currency, the other factors to be considered for selection of the OA resources are – reputation of the publisher and creator, accessibility with the existing hardware and software infrastructure, date of publication and regular update, user requirement, objective of the institution, easy navigation and searching facility, etc.

5.2.2 Procurement of Digital Resources

The LICs can collect the commercial digital resources from publisher or aggregators or from different consortia. During present times, consortia based collection development is found to be existed. An LIC can be a member of the group/consortium and for the individual member the consortium procures the e-resources and the individual LIC gets the right to access those resources. The payment may be made from the consortia or the funding agency for the particular institution. However, for the OA resources, the LICs do not require to pay. The OA resources can be brought to the users by simply providing the links in the library portal. LICs generally subscribe to e-journal, e-database and e-book. Different subscription models are available for these subscription-based resources.

a) Subscription model for e-journal

Different subscription models are available for e-journals offered by the publishers (Arora, 2001). Some of them are –

i) Electronic only: Only electronic version of a journal is offered in this model. Sometimes, special discounted price is offered for electronic version only.

ii) Electronic subscription with print subscription: Access to electronic form is given along with the subscription of print form.

iii) Bundled electronic subscription: The entire range of e-journal access is offered for a fixed payment and access to an individual journal or a subset of the bundle is not allowed.
iv) Electronic subscription with back volume archive: Sometime, access to entire back volume is offered for onetime payment and sometimes access is given for few years with moving wall option where the range can be adjusted.

v) Electronic subscription with campus licenses: On payment of a fixed amount of fees, unlimited access to the e-journal is offered.

vi) Consortium licensing: Most attractive price and access rights are given under consortia licensing.

vii) Pay-per-look: User can search the entire database of e-journal, select a particular article and can download the full-text for which payment has to be made.

b) Subscription model for e-book
Various models for e-book selection are offered by the publisher and the subscription agencies. Some of the commonly seen models are –

i) Print on demand: The user has to pay for the particular printout of selected e-book.

ii) One time purchase: An e-book title can be purchased for forever access.

iii) Free browsing with payment for printing and downloading: The user can freely browse the e-book database but has to pay for downloading or taking printout of a particular title.

iv) Annual subscription: Annual subscription for access a title may be allowed with ownership or without ownership of the title.

c) Subscription model for e-database
The subscription models for e-databases are not clear enough as that of e-journal and e-book. The various models are formed based on some factors like:

- Total number of users of database;
- Size and the number of campuses of an institution;
- Discount according to number of database subscription etc.

5.2.3 Licensing
The licensing agreement should be scrutinized before buying a digital resource by the professionals carefully. To protect their investments, the publisher or the other suppliers
of digital resources take the help of license to define and control the use and access of their products. The Copyright Act of 1976 – USA has the provisions for the educational fair use of copyrighted materials within a framework that carefully balances the rights of authors, publishers, and other copyright owners and the need of users for the free exchange of ideas. But the rapid change of technology has compelled the provider of information to adopt some negotiating terms and conditions to control the access of their products.

Obviously, a license is negotiated between the rights holder of the resources i.e. the licensor and the representative of the user community i.e. licensee. The first party may be the creator or publisher or the provider and the second party is representative of a LIC or in case of consortia representative of a group of LICs or the organization. A license may define the right and privileges of the two parties differently than those defined in the Copyright Act of 1976 – USA but may impose some conditions therein. As for example, in the Copyright Act the right to lend to the public is allowed but lending right is only given to the users defined in the license. The right to quote is given for commentary in the Copyright Act but permission may be required for a licensed resource to quote. Right to copy and distribute in limited way for fair use is allowed in Copyright Act but license may prohibit these outside the user defined in the license. Licensing agreement should not be allowed to put restrictions on the fair use and the provisions for the LICs authorized in the Copyright Act. The LICs should be very careful while making a license agreement. Licensing Principles for Electronic Resources Special Committee (2004) of American Association of Law Libraries in “Principles for Licensing Electronic Resources” has put forwarded some important issues for the LICs to be considered while making a licensing agreement, some of which are –

- The access rights of the resources are given to the licensee for permanent use and ownership or subscription-based access rights only;
- The period of time for access rights granted;
- Use of the resources for printing, downloading, copying;
- The means of authentication and access to the resources to authorized users;
• Identification of the user without physical location;
• If permanent use is given in the license a usable copy of the resource should be given with the necessary interface. Specification should be there under what conditions the users can use the archival copy;
• The access to the licensed resources will be available or not after the subscription period;
• The licensee should have the right to access the usage statistics of the licensed resources;
• The financial obligations of the licensee and the licensor if either party terminates the license before the end of license period;
• The terms of reimbursement if in any case the licensed materials are unavailable for use within the license period.

The issue of licensing is not yet has got the proper attention which is reflected from the university libraries websites also. It is found that only three university libraries namely Jayakar Library University of Pune, Ananda Rangapillai Library of Pondicherry University and University Library of Jawaharlal Nehru University have made aware the users about illegal downloading and use of the digital resources available.

5.3 Organization of Digital Resources

The digital resources should be properly organized for easy access and retrieval by the users. Traditional classification schemes and catalogue codes were previously used to organize various types of digital resources, but these were not sufficient for proper organization of these resources. “Traditional catalogue codes and bibliographic formats – such as AACR (Anglo American Cataloguing Rules) and MARC (Machine Readable Cataloguing) – are not useful for digital resources because catalogue systems and MARC formats do not have provisions for describing different types of digital information resources, especially web pages” (Chowdhury & Chowdhury, 2007, p. 123 ). Therefore, metadata and different markup languages like SGML, HTML, and XML are used for organization of digital resources and resource discovery.
5.3.1 Metadata

The application of metadata was developed to cope up with the problems of cataloging and indexing of digital resources. Metadata is a buzzword in Library and Information Science. The classical definition of metadata is “data about data”. It describes the attributes and contents of a document or object. The digital resources are to be described, identified and disseminated to the users and the metadata can be used to achieve this purpose. According to Merriam Webster Dictionary, metadata is “data that provides information about other data” <http://www.merriam-webster.com/dictionary/metadata>. NISO defined metadata as “structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource” (NISO, 2004, p.1). DCMI (Dublin Core Metadata Initiative) described metadata as “data associated with either an information system or an information object for purposes of description, administration, legal requirements, technical functionality, use and usage, and preservation” (Woodley, Clement & Winn, 2005). The structural form of metadata is important to describe a digital resource so that the metadata can be used as a catalogue for managing the collection (Gill, 2008).

a) Types of Metadata

Metadata can be categorized into the following three main categories based on the function for which it is used.

**Descriptive metadata:** It includes bibliographic description consisting of keywords or subject descriptors.

**Structural metadata:** The elements within the digital resource that facilitate navigation. The table of contents, index etc. are structural metadata.

**Administrative or technical metadata:** It includes date of creation, version, file format, compression technology etc. It is vital for long term collection management. There are two other subsets of administrative metadata – preservation metadata and rights management metadata which are sometimes listed as separate categories of metadata. Rights management metadata contains intellectual property rights and preservation metadata contains information needed to manage the preservation of digital resources.
Metadata can be appended with a digital resource or can be stored separately. Storing metadata in a database and linked to the digital resources can simplify management of metadata. The main functions of metadata are –

**Resource discovery:** Identify digital resources, bring similar resources together, and give location information.

**Organize the digital resource:** Metadata can help to organize the web resources with their names and location.

**Interoperability:** By using defined metadata schemes and shared transfer protocols metadata can be shared between multiple systems.

**Preservation:** Preservation metadata puts the elements required to make accessible the digital resources in future.

**b) Selected Metadata schemes**

Various metadata schemes are being used by different users depending upon their requirements. Some of the popular and common metadata schemes are discussed below.

**Dublin Core (DC):** The name of Dublin Core (DC) metadata is evolved from the place Dublin, Ohio in USA where a workshop was held in 1995 to develop a metadata scheme and its element set to describe web resources. DC metadata scheme consists of 15 elements – Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights. DC metadata can be classified into two types – Simple or unqualified and complex or qualified according to the refinement made in the elements. All elements of DC metadata are repeatable and optional and can be put in any order <http://dublincore.org>.

**MODS (Metadata Object Description Schema):** Derived from selected elements of MARC 21, MODS is expressed in XML schema language. MODS elements are simpler to apply than MARC 21 and are richer than Dublin Core <http://www.loc.gov/standards/mods>.

**EAD (Encoded Archival Description):** The EAD Document Type Definition (DTD) is a standard for encoding archival finding aids using Extensible Markup Language (XML). The standard is maintained in the Network Development and MARC Standards Office of
the Library of Congress (LC) in partnership with the Society of American Archivists
<http://www.loc.gov/ead>.

**LOM (Learning Object Metadata):** LOM was developed by IEEE Learning Technology Standards Committee (LTSC) to enable to facilitate search, evaluation, acquisition and use of learning objects, for instance by learners or instructors or automated software processes. This multi-part standard also facilitates the sharing and exchange of learning objects by enabling the development of catalogs and inventories while taking into account the diversity of cultural and lingual contexts in which the learning objects and their metadata are reused <http://www.ieeeltsc.org:8080/Plone>.

**TEI (Text Encoding Initiative):** TEI is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of guidelines which specify encoding methods for machine-readable texts, chiefly in the Humanities, Social Sciences and Linguistics. Since 1994, the TEI guidelines have been widely used by libraries, museums, publishers and individual scholars to present texts for online research, teaching, and preservation. In addition to the guidelines themselves, the consortium provides a variety of supporting resources including resources for learning TEI, information on projects using the TEI, TEI-related publications, and software developed for or adapted to the TEI <http://www.tei-c.org/index.xml>.

d) Metadata creation

Informational professionals, creator of the digital resources are involved in creation of metadata. The creator of a digital resource or the technical staff involved in digitization of a document assigns the basic structural and administrative metadata. Metadata can be created either manually or automatically. The professionals involved in the management of digital library or digitization input the metadata in a defined format manually, while in automatic process metadata is extracted from the digital resources. There are various types of tools for creation for metadata. They are as follows –
Templates: Templates are predefined sheets which provide an outline of schema elements where metadata values can be entered and it will generate a formatted set of the elements and their respective values.

Mark-up tools: Metadata elements and values are generated in XML or SGML Document Type Definitions (DTD) by the mark-up tools.

Extraction tools: Metadata is extracted from the analysis of textual digital resources automatically by extraction tools.

Conversion tools: These tools translate one metadata format to another format.

There are a number of softwares available for automatic generation of metadata. Some of them are DC-dot <http://www.ukoln.ac.uk/metadata/dcdot>, Dublin core generator <http://www.dublincoregenerator.com> etc. DC-dot was developed by Andy Powell, UKOLN (UK Office for Library and Information Networking), University of Bath. It retrieves a Web page and automatically generates Dublin Core metadata, either as HTML <meta> tags or as RDF/XML suitable for embedding in the <head>...</head> section of the page. DC-dot copies resource “identifier” metadata from the Web browser’s “address prompt,” and harvests “title,” “keywords,” “description,” and “type” metadata from resource META tags. If source code metadata is absent DC-dot will automatically generate “keywords” by analyzing anchors (hyperlinked concepts) and presentation encoding such as bolding and font size but will not produce “description” metadata. DC-dot also automatically generates “type,” “format” and “date” metadata and can read source code programming that automatically tracks date.

e) Interoperability and Metadata harvesting

After the rapid growth of e-publishing, the number of digital repositories of various educational, research institutions is also increasing. The accessibility of these vast and diverse resources is a very difficult task. The lack of interoperability is one of the most significant problems that digital repositories are facing today. In general, interoperability is the ability of systems, organizations and individuals to work together towards common or diverse goals. “Interoperability is a broad term, touching many diverse aspects of archive initiatives, including their metadata formats, their underlying architecture, their openness to the creation of third-party digital library services, their integration with the
established mechanism of scholarly communication, their usability in a cross-disciplinary context, their ability to contribute to a collective metrics system for usage and citation, etc.” Sompel & Lagoze (2000). In the technical arena, it is supported by open standards for communication between systems and for description of resources and collections, among others.

The evolution of Open Archives Initiative- Protocol for Metadata Harvesting (OAI-PMH) is one of the solutions to overcome the problem of lack of interoperability. The OAI-PMH was designed to facilitate the technical interoperability among distributed digital repositories and archives. It provides an “application independent interoperability framework based on metadata harvesting that can be used by a variety of communities who are engaged in publishing content on the Web” OAI (2002). The objective of OAI-PMH is to develop a low-barrier, lightweight framework to facilitate the information discovery of content in distributed archives. OAI-PMH has two main components: Service Provider and Data Provider. The Service Provider sends request to the Data Providers for sending the required information regarding metadata and the records. It uses HTTP as internet protocol to send and receive the metadata information from Data Provider to Service Provider. Again, XML is used for encoding and exchanging information and qualified and simple DC as metadata scheme. But it also supports EAD, METS, etc. Arc, Citebase, METALIS, NCSTRL, OAIster, etc. are some software which support OAI-PMH.

One of the most widely used metadata harvester software is the Open Harvester Systems (OHS) which is a free metadata indexing system developed by the Public Knowledge Project (PKP) through its federally funded efforts to expand and improve access to research <http://pkp.sfu.ca/?q=harvester>. PKP-OHS allows creating a searchable index of the metadata from Open Archives Initiative (OAI)-compliant archives such as sites using Open Journal Systems (OJS) or Open Conference Systems (OCS).
Many LICs are using PKP-OHS worldwide. In India, PKP-OHS is used by SEED of IIT Delhi <http://eprint.iitd.ac.in/seed> which has currently 6176 papers from 4 archives indexed in it. The University of Glasgow Open Archives Harvester <http://daedalus.lib.gla.ac.uk/pkpharvester/harvester/index.php> is also using PKP-OHS which has 3059 papers from 4 archives indexed in it.

5.3.2 Institutional Repository

The digital resources produced by an institution can be separately organized to give a user interface to its users. This is popularly known as institutional repository. According to Lynch (2003), institutional repository of a university is “a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members”. Institutional repositories manage and create supporting services to store, preserve, and disseminate an organization’s digital information or knowledge assets created by faculty, research staff, and students. The concept of capturing and making available the research outputs of an institution and other relevant documents locally to the users of a particular institution
through intranet and globally through internet is institutional repository. An institutional repository might include electronic versions of documents such as research papers, project reports, patents, theses and dissertations. It may also include many of the digital assets generated by an institution such as working papers, lectures, conference proceedings, learning objects, administrative documents, course notes, etc. The learning objects among others include study materials, assignments, question papers, audio-video materials and multimedia presentations like interactive e-learning modules.

Establishing an institutional repository enables a university to publicize its research, teaching programmes, and other resources to the outside world. IR of a university can be a base for evaluating the standards of the research outputs from which the quality of a university’s academic research output can be easily judged. The benefit of OA can be achieved by establishing institutional repository. Different universities have started creating IR by using either OS software or proprietary software. The following is the homepage of the IR of Jadavpur University, Kolkata <http://dspace.jdvu.ac.in>.

Fig. 5.2: Home Page of IR, Central Library, Jadavpur University
(May 27, 2012)
Software for creating IR

Different proprietary and OSS are available to create an institutional repository. A discussion is made below about OSS and proprietary software along with examples.

a) Open Source Software

The open source software “promotes software reliability and quality by supporting independent peer review and rapid evaluation of source code. To be certified as open source, the license of a program must guarantee the right to read, redistribute, modify, and use it freely” <www.opensource.org>. Open source software is different from freeware and shareware. Freeware software is released free of cost in binary format only and its licenses usually prohibit modifications and commercial redistribution. On the other hand, shareware is software that is released free of cost in binary format but only for a limited trial period after which users are encouraged to purchase the software. The important advantage of open source software is the availability of source code under a copyright license that permits users to study, change, and improve the software, and to redistribute it in modified or unmodified form. If the budget of a university library is limited and has its own computer experts then it should go for open source software. A discussion on the popular open source software used for creating IR is made here.

i) DSpace <http://ww.dspace.org>

DSpace is designed by MIT in collaboration with the Hewlett-Packard Company. DSpace architecture supports the participation of the schools, departments, research centres and other units typical of a large research institution. As the requirements of these communities might vary DSpace allows the workflow and other policy-related aspects of the system to be customized to serve the content, authorization and intellectual property issues of each. Supporting this type of distributed content administration coupled with integrated tools to support digital preservation planning makes DSpace well suited to the realities of managing a repository in a large institutional setting.
ii) CDSWare <http://cdsware.cern.ch>
CDSWare is maintained and made publicly available by CERN and supports electronic preprint servers, online library catalogs and other web-based document repository systems. CDSWare was built to handle very large repositories holding disparate types of materials including multimedia content catalogs, confidential and public sets of documents, etc.

iii) Eprints <http://software.eprints.org>
The University of Southampton develops the Eprints software for managing large institute oriented digital archive for scholarly objects. Eprints worldwide installed base affords an extensive support network for new implementations. The size of the installed base suggests that an institution can get it up and running relatively quickly and with a minimum of technical expertise.

iv) Fedora <http://www.fedora.info>
The Fedora digital library software is based on the Flexible Extensible Digital Object and Repository Architecture (Fedora). The system is designed to be a foundation upon which full-featured institutional repositories and other interoperable web-based digital libraries can be built. Jointly developed by the University of Virginia and Cornell University, the system implements the Fedora architecture adding utilities that facilitate repository management. The system’s interface comprises three web-based services – a management API that defines an interface for administering the repository, operations necessary for clients to create and maintain digital objects in the repository and a streamlined version of the access system implemented as an HTTP-enabled web service.

v) MyCoRe <http://www.mycore.de/engl/index.html>
MyCore grew out of the MILESS project of the University Of Essen. The MyCOre system is now being developed by a consortium of universities to provide a core bundle of software tools to support digital libraries and archiving solutions (or Content Repositories, thus “CoRe”). The bundle is designed to be configurable and adaptable to local requirements, without the need for local programming efforts. The core contains all
the functionality that would be required in a repository implementation, including distributed search over geographically dispersed repositories, OAI functionality, audio/video streaming support, file management, online metadata editors etc.

b) Commercial Software

Some of the popular commercial softwares for creating digital library are –

i) CONTENTdm <http://www.contentdm.org>
CONTENTdm is a Digital Collection Management Software which is able to handle the storage, management and delivery of digital resources of any format — local history archives, newspapers, books, maps, slide libraries or audio/video to users across the Web. Though CONTENTdm can be installed locally OCLC’s Hosting Services makes it even easier by offering operational support and reliability at an affordable price. PDF documents, theses, dissertations and reports can all be easily stored and managed in CONTENTdm and it helps to develop institutional repository.

ii) DLXS <http://www.dlxs.org>
The University of Michigan Digital Library eXtension Service (DLXS) provides the foundation and the framework for educational and non-profit institutions to fully develop their digital library collections. The newest DLXS enhancement - XPAT - is a powerful, SGML/XML-aware search engine and an ultra-versatile tool for the development of digital libraries. XPAT provides excellent support for word and phrase searching, indexing of SGML elements and attributes, a baseline of support for XML (without Unicode), fast retrieval, and open systems integration.

iii) Veridian <http://www.dlconsulting.com/veridian>
Veridian easily manages newspaper collections with multiple titles and millions of pages without slowing its blazingly fast search engine. It was built specifically to display newspaper and other text-based collection but it is equally capable of managing and displaying collections of photos, video and other file types. Veridian does not change data to a proprietary format. The content remains the same regardless of the format.
iv) VITAL <http://www.vtls.com/products/vital>

VITAL provides every feature—ingesting, storing, indexing, cataloging, searching and retrieving—required to handle large text and rich content collections. Workflows have been adapted to meet the specialized requirements of libraries and archives. VITAL offers complex object support, authority management, flexible relationship building, the VALET submission workflow tool and consortium features allow multiple institutions or separate departments to share the repository while retaining control over their materials. VITAL includes its own web access portal but can also be accessed through Chivas Social Discovery. Customers wanting to implement a more unique and distinctive interface for their collections can use VTLS custom Drupal modules and Drupal design and implementation services. VITAL supports the metadata format OAI-PMH and also includes conversion to/from the various formats to allow data exchange.

5.4 Access Management of Digital Resources

Due to the diverse types of digital resources, it becomes difficult for the users to access the information of a digital library. The collection may have e-book, e-journal, conference proceedings, databases, audio, and video collections. Every type has different formats and different searching techniques. So, to acquaint the users to the digital resources available in a digital library, there should be a proper user interface from where they can start their search. Users should be informed about different types of available digital resources in a library for its best use. We can do this by providing the list of the resources in the form of a web page where links to different resources like e-journals, online databases and internet resources are provided. These jobs are being performed mostly by the library professionals while doing these they have to face number of hurdles like enormous and continuous growth of digital resources, their changeable characteristics and the associated tools and method related to these resources. Following tools and methods of building a useful user interface can be helpful in giving the user a proper platform to use the digital resources available in an LIC.
5.4.1 Library Portal

A library may develop digital resources collection from different angles – digitized documents, born-digital, also from web resources. The resources are of various types: e-book, e-journal, e-database, ETD, etc. A proper way to search these resources is a must failing which users may be in a complex situation to identify, locate those resources. Otherwise, users have to browse each and every resource available in different platforms which require multiple log-in to access these resources. This problem can be largely addressed by creating a library platform. The library portal gives a single search interface of the digital resources from OPAC, e-journals, online databases to their in-house collections available in their IR.

Different experts have identified different benefits of library portal which can be extended to users for easy location of the resources for the users. Cox (2003) has identified the following five benefits of library portal.

- Easier access for users: Provides search facilities ay a single platform.
- Simplified authentication: Provides IP authentication to users.
- Unified presentation of quality resources: Supports searching from bibliographic search to full text options.
- Personalisation: Groups of users can be offered cluster of resources, or they can store their favourite databases and searches for quick access and reuse.
- The portal may be a mechanism by which to offer services to an institutional portal or Virtual Learning Environment.

Library portal is also known as library web portal which accommodates the web based digital resources of the LICs. Several Content Management System (CMS) softwares are available to manage library web portal. There are different OSS (Open Source Software) and Proprietary CMS software available to manage library portal. A CMS provide web browser and WYSIWYG (What-You-See-Is-What-You-Get) type editor to the creators to upload their creation into the system. After the creation of the content, it is stored in a central repository along with metadata values. The content stored in the repository is
delivered to the users by publishing engines which generate web pages of the content and in a variety of formats. CMS software has many features such as search engine, file upload and download, tools for synchronous & asynchronous communication, wiki functionality, guest book, newsletters, Really Simple Syndication (RSS), multiple platform support, etc. Following are some OS CMS software available for developing a library portal.

### Table 5.1 CMS Software (Selected)

<table>
<thead>
<tr>
<th>Open Source Software</th>
<th>Proprietary Software</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Software</strong></td>
<td><strong>Website</strong></td>
</tr>
<tr>
<td>imCMS</td>
<td><a href="http://www.imcms.net">http://www.imcms.net</a></td>
</tr>
</tbody>
</table>

(Source: http://www.opensourcecms.com)

Libraries have started to develop library portal using different CMS software available. The Central Library of IIT Bombay has developed a library portal using OS CMS software Joomla. Joomla is an award-winning content management system (CMS), which enables building Web sites and powerful online applications. Many aspects, including its ease-of-use and extensibility, have made Joomla the most popular CMS available. Best of all, Joomla is an open source solution that is freely available to everyone <http://www.joomla.org>.
The website of IIT Bombay is a platform where all the resources and services are brought together for the users in a single attractive interface.

5.4.2 User Interface

User interface bridges the user and the digital resources available in the digital library. A good user interface can make difference of the uses of a digital library than that of a bad user interface. User interface design includes the screen appeared to the user, how they view it and how they manipulate it. It also includes searching and browsing option available, getting the searched results, help menu. To design the user interface, various graphical representations are used to give a better and attractive look. The following points can be considered as guiding principles while designing user interface

- Strive for consistency: Terminology, layout, instructions, fonts, and colors should be used consistently throughout the interface;
- Provide shortcuts to skilled users;

98
• Provide informative feedback: The system should provide users with appropriate feedback about the sources and what is being searched for;
• Design for closure: Users should know when they have completed searching the entire collection or have viewed every item in a browse list;
• Permit reversal of actions: Users should be able to undo or modify actions;
• Support user control: The user should be able to monitor the progress of a search and should be able to specify the parameters to control a search;
• Reduce short-term memory load: The system should keep track of some important actions performed by the users and should allow them to jump to a formerly performed action easily;
• Make error handling facilities simple: Users should be able to rectify errors easily, and all error messages should be clear and specific;
• Provide plenty of space: A large space should be provided for entering text in the search boxes.

The LIS professionals can consult the web page developer of their university to give an idea regarding the access points, facilities, and links to be provided to the users before launching the library webpage. This is because the LIS professionals working in a particular LIC can better understand the users need and their expectations regarding the search interface.

5.5 Preservation of Digital Resources

Digital resources have several advantages over the printed form but their storage media are prone to decay and the hardware and software on which they run become obsolete rapidly. Digital preservation deals with the issues that help to manage the risk of inaccessibility of the digital resources due to technological obsolescence. There are two facets of digital preservation – the first one is digital preservation of the storage media and the second one is the hardware and software required to run the digital resources. We have to consider the digital preservation of both digitized and born-digital resources. Taking concern only about the storage media but ignoring the future accessibility will
lead only to a store house of digital resources with of no use (Lahkar, 2012). Digital preservation has become difficult due to the following factors –

**Diverse Forms:** Digital resources have several forms. Some digital resources can be converted to printed form and vice versa. But there are some resources like web pages for which the case is different. So, the nature of digital objects creates problems in digital preservation.

**Hardware/ Software obsolescence:** Digital resources need specific hardware and software to access them. But the rapid change in hardware and software technology is a threat to digital preservation.

**Fragility of storage media:** The storage media like optical and magnetic media deteriorates quickly due to heat, dust, humidity, etc. This fragile nature creates problem in digital preservation.

**5.5.1 Strategies of Digital Preservation**

The preservation strategies for digital resources may be long-term or medium-term or short-term. To preserve digital resources, a single strategy will not work for all types of digital resources and conditions. Arora (2001) referred to the following 13 strategies for digital preservation given by Claire Tristram.

i) **Bit-stream copying:** Taking the back-up of digital resources is bit-stream copying. It is not a long-term strategy of digital preservation.

ii) **Refreshing:** To face the problem of decay of storage media it is applied. It is simply copying digital resources from one storage media to another same type media.

iii) **Technology preservation:** In technological preservation, the technological environment required to access the digital resources which include operating system, application software, media drives, etc. are preserved.

iv) **Digital Archaeology:** In this strategy, specialized techniques are used to recover the content of the digital resources from damaged storage media or from obsolete hardware and software environment.
v) **Analogue backups:** Analogue backup in durable media like taking good quality printout or creation of microfilm of the digital resources can help to preserve it without losing the quality of the digital resources.

vi) **Migration:** In this strategy, the digital resources are transferred periodically from one hardware/software configuration system to another system. This process is applied to preserve the integrity of the digital resources and to keep the resources accessible in the changing hardware/software system.

vii) **Replication:** Multiple preservation strategies are represented by replication. For example, LOCKSS (Lots of Copies Keeps Stuff Safe) supports a system which allows libraries to collect, preserve and provide their users with access to material published on the Web. The system applies the replication process. Another process called peer-to-peer data trading comprising archiving sites is used to handle data for reliable replication.

viii) **Reliance on standards:** To rely on a well recognized format and reject proprietary and less used standards will help to upgrade the standards to work it on changing environment. This strategy is called reliance on standards.

ix) **Normalization:** In this strategy, digital resources of same type are preserved in a particular selected format, which is believed to be the best for the longevity and preservation of the resources.

x) **Canonicalization:** It is a method which is proposed to verify that the converted form of a digital resource from its original form has or has not lost its fundamental properties.

xi) **Emulation:** Here, a special kind of software called emulator is used to translate instructions from the original software to run the digital resource in a new platform of hardware/software.

xii) **Encapsulation:** A digital resource combined with its metadata like reference, provenance, fixity, context information reduces the possibility of running the resource in different environment. This is what is done in encapsulation.

xiii) **Universal virtual computer:** It is imagined that a type of programme will imitate the computer environment for any digital resources called universal virtual computer.
5.5.2 Preservation Metadata

Preservation metadata is considered as the subset of administrative metadata designed for support continuing access of the digital resources. All the metadata are essential for preservation and the preservation metadata is an amalgamation of all types of metadata. For the long term accessibility of the digital resources, preservation metadata describes the attributes necessary for it. Unlike the descriptive metadata which are used for identification and discovery of digital resources, Preservation metadata is intended to store technical details like the format, structure and use of the digital content, the history of all actions performed on the resource including changes and decisions, the authenticity information such as technical features or custody history, and the responsibilities and rights information applicable to preservation. The scope and depth of the preservation metadata required for a given digital preservation activity will vary according to numerous factors such as the “intensity” of preservation, the length of archival retention, or even the knowledge base of the intended user community.

According to OCLC/RLG (2002, p.1) preservation metadata “is the information necessary to maintain the viability, renderability, and understandability of digital resources over the long-term. Viability requires that the archived digital object’s bit stream is intact and readable from the digital media upon which it is stored. Renderability refers to the translation of the bit stream into a form that can be viewed by human users, or processed by computers. Understandability involves providing enough information such that the rendered content can be interpreted and understood by its intended users”.

**OAIS reference model:** OAIS (Open Archival Information System) reference model (2003) of International Standard Organization (ISO) was a major initiative of digital preservation metadata. OAIS reference model identifies the functions in the long-term storage of and access to digital resources. These functions include acquisition and processing, archival storage, preservation planning, access, data management and administration of the archive. The OAIS model includes an information model to support the management and preservation of digital resources. This model has been used as the basis of preservation metadata initiatives. OAIS reference model was developed by the
Consultative Committee for Space Data Systems (CCSDS). It is a framework for understanding and applying concepts needed for long term digital preservation. It is also a starting point for a model addressing non-digital information. The model establishes terminology and concepts relevant to digital archiving, identifies the key components and processes prevalent to most digital archiving activity and proposes an information model for digital resources and their associated metadata. The OAIS reference model can be applied at a broad level to achieve handling digital image files, born-digital objects. The OAIS reference model enjoys the status of a de facto standard in digital preservation. It provides a high level overview of the types of information needed to support digital preservation that can broadly be grouped under two major umbrella terms called Preservation Description Information (PDI) and Representation and Descriptive Information (RDI).

i) Preservation Description Information: It consists of four major types of metadata elements namely, reference information, provenance information, context information and fixity information.

Reference information enumerates and describes identifiers assigned to the content information such that it can be referred to unambiguously both internally and externally to the archive (e.g. ISBN, URN).

Provenance information documents the history of the content information (e.g. its origins, chain of custody, preservation actions and effects) and helps to support claims of authenticity and integrity.

Context information documents the relationship of the content information to its environment (e.g. why it was created, relationships to other content information).

Fixity information documents authentication mechanisms used to ensure that the content information has not been altered in an undocumented manner (e.g. checksum, digital signature).

ii) Representation and Descriptive information: It facilitates proper rendering, understanding and interpretation of a digital object’s content. At the most fundamental level, representation information imparts meaning to an object’s bit-stream. For example,
it may indicate that a sequence of bits represents text encoded as ASCII characters and furthermore, that the text is in French. The depth of the representation information required depends on the designated community for whom the content is intended. Descriptive information metadata contains more ephemeral metadata, the information used to aid searching, ordering and retrieval of the objects.

There are several preservation metadata formats. Two popular preservation metadata formats are discussed below –

**a) PREMIS:** PREMIS (PREservation Metadata: Implementation Strategies) is a sponsored group of the OCLC and the RLG composed of more than thirty international experts in preservation metadata. It aims at to –

(i) Define a core set of implementable broadly applicable to preservation metadata elements, supported by a data dictionary; and

(ii) Identify and evaluate alternative strategies for encoding, storing, managing and exchanging preservation metadata in digital archiving systems. The PREMIS data dictionary is a comprehensive, practical resource for implementing preservation metadata in digital archiving systems. It defines implementable, core preservation metadata, along with guidelines and recommendations for management and use.

**b) METS:** Metadata Encoding and Transmission Standard (METS) is an XML schema developed by the Digital Library Federation (DLF) and maintained by the Library of Congress (LoC). The objective of METS is to provide an XML based document format for encoding metadata that supports the management and exchange of digital objects among repositories. The METS schema is not prescriptive about what metadata elements have to be included. However, in order to sustain interoperability, METS endorses specific XML schemas for a number of metadata element sets. A METS document consists of seven sections, to carry the various types of metadata related to the access and management of digital objects.
5.5.3 File Format for Digital Preservation

A good number of digital resources are available in different file formats and the number of file formats is continuously growing. For long term preservation of digital resources, proper selection and upgrading to new file format is very important. Brown (2008) defined file formats as “file format encode information into forms that can only be processed and rendered comprehensible by very specific combinations of hardware and software”. Complex natures of digital resources and increasing number of file formats have made it difficult to choose a proper file format for digital preservation. We face lots of problems in digital preservation like digital resource created long time ago need specific hardware and software to access it, choosing a file format may help in one aspect like compression but may result loss of originality of the resource and while converting from one file format to another there may be loss of information. Williamson (2005) in “Strategies for managing digital content formats” mentioned “one of the key components in ensuring resource longevity is the choice of file and media formats used to create, store, and deliver digital content, and the strategies that are employed to manage these in the long run”.

While choosing file format for long term digital preservation, Guercio & Cappiello(2004) suggested some criteria to be evaluated. They are as follows –

Openness: To fulfill this criterion, a file format has to be OAIS compatible; support all internet protocol, support metadata, support authenticity information, data integrity has to be guaranteed, and data backup must be simple.

Portability: A file format should be independent of hardware and software platform. Besides the above two the qualities that should be checked in file format selection, are low space cost, robustness, simplicity, highly tested, loss-free, supports metadata, etc. Based on the selection criteria, they suggested file formats for long term preservation of three mostly used types of digital resources – textual, image, sound and video type of digital resources.
Table 5.2 File Formats for Digital Preservation

<table>
<thead>
<tr>
<th>Sl No.</th>
<th>Type of file</th>
<th>Format suggested</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Text</td>
<td>Unicode (ASCII), XML and PDF/A</td>
</tr>
<tr>
<td>2</td>
<td>Image</td>
<td>raster: standard TIFF for master copies (no-compression, high resolution), JPEG for safety copies or distribution vector: CGM, EPS, DXF, SVG</td>
</tr>
<tr>
<td>3</td>
<td>Sound</td>
<td>compressionless WAV (PCM-coding)</td>
</tr>
<tr>
<td>4</td>
<td>Video</td>
<td>MPEG</td>
</tr>
</tbody>
</table>

Most of the digital resources of a university library are available in textual format. XML (eXtensible Mark-up Language) and PDF/A are the two file formats which can be chosen for long term preservation of these resources.

Developed in the year 1996 by XML Working Group created within the World Wide Web Consortium (W3C), XML offers two great advantages – transparency and interoperability. On the other hand, Portable Document Format (PDF) is a file format developed by Adobe Systems for representing records independently from the original application, software, hardware and operating system. A PDF file can describe records containing any combination of text, graphics, and images in a device independent and resolution-independent format. Often XML and PDF are regarded as rivals for long-term preservation. PDF and XML are complementary and it is actually more appropriate to use both XML and PDF for preservation of a document than to choose between XML and PDF. The PDF/A, the archival version of PDF format was expressly introduced for the preservation of digital resources independent of the tools and systems used for creating, storing and rendering them.

For the long term preservation of digital resources, PDF/A should be considered instead of PDF. PDF reproduces the visual appearance of documents whether they were originally created in PDF, converted from another electronic format or digitized from paper or microfilm. This format is mainly used to collect and disseminate information
over the internet and store electronic records PDF cannot be used as an archival format. Long-term solutions are needed to keep digital PDF records accessible for long time. Anderson (2005) provided a to-do-list for managing file formats used in long term preservation. These are –

- Inventory all files and make a list of formats;
- Migrate older materials to newer versions;
- Try to limit the number of file format; and
- Watch the technology market for news regarding file formats.

To choose and manage file format for long-term preservation is a complex issue. Different file formats need different software platform to access them. Therefore, attention should also be paid to the necessary hardware and software which will be required to run the file saved using a particular file format. Preference should be given to open file format over proprietary file format for long term digital preservation.

5.5.4 Storage Management for Digital Preservation

The preservation of digital resources depends upon the storage media used for preservation. Storage media are vulnerable to technological obsolescence and natural decay, etc. Three main types of storage media used are hard disk, optical media and magnetic tape. We should always go for the latest available in the market when we select the storage media. The optical media such as CD-ROM, DVD-ROM, etc are low cost affordable storage media for digital preservation. Incorrect storage, handling and use can reduce the lifespan of these types of storage media.

The following factors should be taken care of for optical storage media like CD-ROM and DVD-ROM –

**Temperature:** Optimum temperature should be maintained in the storage environment. High or low or fluctuating temperature is not suitable for storage. The storage temperature should be 0-23° C. The copies which are used for access should be stored at temperature similar to the access environment.
**Relative Humidity (RH):** High or low or fluctuating RH can cause damage to optical storage media. The optimum RH should be maintained between 30% to 50% for storage of optical disc.

**Air Quality:** The air pollution agents like particulate and gaseous pollutants should be controlled in the storage environment of the optical storage media.

**Light:** Direct sunlight should be avoided and should be stored in dark environment.

**Handling:** Poor handling like touching the bottom surface by hand, bend disc, leaving the disc in the optical drive etc. can cause damage to the optical storage media. Wearing close-fitting vinyl gloves, avoiding archival copies for user access, preparing written handling guidelines to users for these storage media will help to preserve them.

**Labelling:** Avoid applying labels or writing on the disc. When labeling is must use a xylene-free, soft felt-tipped pen with water-based ink and only write on the clear inner hub on the label side of the disc.

These are some precautions to be looked into for preservation of optical storage media. But the optical storage media are the secondary storage media and we should concentrate on Hierarchy Storage Mechanism (HSM) and redundancy for the primary storage device like hard disc. The digital resources can be stored in an HSM where most frequently used data are kept in the first disc and the less used data are kept in an automated tape library. An HSM system can automatically migrate data from disc to tape and vice versa. Multiple storage location can be used for digital preservation for the digital resources of a distributed network.

A system where different types of hardware and software are completely dependent on each other for storage of digital resources has a drawback. The system may breakdown if any one of the subsystem fails. Redundant Array of Inexpensive Discs (RAID) technology models can help in this regard which can give greater security and performance. In this system, the data are stored across a number of discs in such a way that even if some disc fails the system functions while the failed component is replaced.
5.6 Summing Up

Due to rapid developments in ICT and its application in creation and management of digital resources, the management of digital resources the LIS professionals should also update their knowledge and practical skills of management of digital resources to cope up in the changing environment. The issue of digital preservation should be given due importance to avoid the inaccessibility of the digital resources in near future. The present status of digital resources management in the university libraries of India under this study are discussed in the next chapter based on survey made.