

# Information Management Resource Kit

## Module on Digitization and Digital Libraries

### UNIT 1. CONCEPTUAL OVERVIEW

#### LESSON 1. WHY DIGITAL COLLECTIONS AND LIBRARIES?

##### NOTE

Please note that this PDF version does not have the interactive features offered through the IMARK courseware such as exercises with feedback, pop-ups, animations etc.

We recommend that you take the lesson using the interactive courseware environment, and use the PDF version for printing the lesson and to use as a reference after you have completed the course.



© FAO and UNESCO, 2005

## Learning Objectives

At the end of this lesson you will be able to:

- recognize the **benefits of digital libraries** in terms of preservation and access to digital material;
- define the **concept of digital libraries** and identify their functional components, and
- have a basic understanding of the **processes** related to the creation, distribution and provision of access to digital libraries.



## Introduction



Internet and the Web have given impetus to new ways of communication and knowledge sharing.

Digital libraries are considered by many to be a key application of Internet and Web technologies.

They enable the creation, organization, maintenance, management, access to, sharing and preservation of digital document collections.

## Introduction

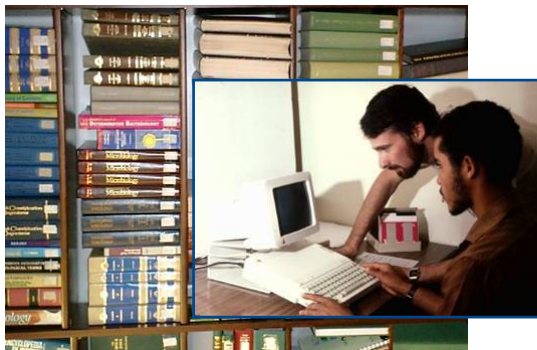


Digital libraries are being created today for diverse communities and in **different fields**: education, science, culture, development, health, governance and so on.

With the recent availability of several free digital library software packages, the creation and sharing of information through digital library collections has become **an attractive and feasible proposition for library and information professionals around the world**.

Let's consider some of the key benefits of digital libraries...

## Benefits of digital libraries



Libraries have traditionally dealt with the organization and provision of **access to physical material** like print-on-paper publications.

Library automation has helped libraries to provide improved access to their collections. However, online access has largely been limited to the computerized library catalogue (OPAC).

Digital libraries differ significantly from such libraries because they allow users to **gain online access to and work with electronic versions of full text documents** and their associated images.

Many digital libraries also provide access to other multimedia content like audio and video.

### Benefits of digital libraries

Before continuing, let's focus once again on the main difference between digital libraries and computerized library catalogues.

What is new about digital libraries?

- They allow users to directly access and use the documents they have found.
- They allow users to quickly find documents by using a computer.
- They allow users to quickly find documents by using a computer and a network.

Select the answer of your choice

### Benefits of digital libraries

Digital libraries bring significant benefits to the users. Please have a look at the following features: are some or all of them interesting for your work environment?



#### Always available

Digital libraries are typically accessed over the Internet and Web. They can thus be accessed from virtually **anywhere**, and at **anytime**. They are not tied to physical location and operating hours of a traditional library. Digital library collections can also be delivered on CD-ROM media to users with inadequate network connectivity.



#### Improved access

Availability of full text electronic documents enable digital libraries to support advanced and novel search and display features including full text search, relevance ranking, and hierarchical document browsing.

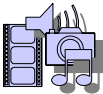
### Benefits of digital libraries

Digital libraries bring significant benefits to the users. Please have a look at the following features: are some or all of them interesting for your work environment?



#### Wider access

A digital library can meet simultaneous access requests for the same electronic document by easily creating multiple instances (or copies) of the requested document. A digital library can thus meet the requirements of a much larger population of users.



#### New forms of access

A digital library can provide access to content in different and more appealing forms, including animation, graphical, audio and video formats; support post processing of information (e.g. conversion of a spreadsheet to graphical form); and adapt to the special needs of physically disadvantaged users.

### Benefits of digital libraries

Digital libraries bring significant benefits to the users. Please have a look at the following features: are some or all of them interesting for your work environment?



#### Improved information sharing

Through appropriate metadata and information exchange protocols, digital libraries can easily share information with other similar digital libraries and provide enhanced access to users.



#### Improved preservation

Since electronic documents are not prone to physical wear and tear and their exact copies can easily be made, digital libraries facilitate preservation of special and rare documents and artefacts by providing access to digital versions of these entities.

### Benefits of digital libraries

Pamela, a university librarian at the University of South Africa, has a problem.

Exams are coming up and many students want to consult question papers from the past few years to prepare for them. My staff cannot handle the demand with just one or two sets of printed question papers... Since the university offers a large number of courses and degrees, there are lots of question papers!



Pamela, Librarian



The student community is unhappy and rightly feels that the library should make better arrangements.

### Benefits of digital libraries

Pamela has just returned from participation in an international conference on digital libraries with lots of ideas. She proposes the possibility of setting up a digital library of question papers to management. They readily agree with the idea but require her to submit a **formal proposal based on a feasibility study**.

A five member committee is formed consisting of Pamela, a controller of examinations, a faculty member, a student community member, and a senior staff member from the computer centre.

The committee conducts a thorough feasibility study and submits the proposal to management.

The study considers various factors: objectives and scope of the digital library, user community, access and delivery options, document formats, storage and other hardware and software requirements, copyright issues, digitization requirements, staff requirements, financial requirements and implementation strategy.

We think that the best way to meet the demand is to set up a digital library of question papers and provide online access to these on the campus network! Here is our feasibility study.



### Benefits of digital libraries

Management accepts the proposal and asks the same committee to oversee its implementation, with Pamela acting as the project manager.

The project is implemented satisfactorily in about six months.



The digital library provides access to question papers through several access points, it can be searched by keyword and papers can be viewed online or printed for offline use.

The student community is now quite happy with the question papers digital library and appreciates the library's key role. The library has earned much respect and admiration from its user community!

Digitization is completely outsourced. The computer centre provides the technical support and also hosts the digital library. The library actively participates in the design, implementation and testing stages. The library also plays a key role in promotion, training, delivery and feedback operations.

### Benefits of digital libraries

Let's look at another scenario.

Mr. Parthan is an agriculture extension worker in the state of Kerala in South India. He works for an NGO (Non-Governmental Organization).

He is responsible for gathering short and long term advisories (guidelines) issued periodically by the three state agriculture universities and getting them to the farmers through the village development offices.

It is not simple to manage this responsibility! Farmers frequently complain of not having access to relevant advisories as related documents are either lost or misplaced in local development offices.

Coordinators in the development offices have expressed their difficulty in managing, retrieving and making copies of the advisories on specific demand from farmers.



Mr. Parthan

### Benefits of digital libraries

Mr. Parthan observes the recent **installation of personal computers** (PCs) in development offices by the state informatics centre to gather land records, data and other development information, as part of the state government's e-governance initiative. He also notices that several Internet kiosks have been set up recently in some villages. He wonders how he can take advantage of these.



He decides to talk with his librarian friend at the university.

..my idea is to digitize the advisories using a scanner. But, how should I organize the digitized documents so that they are easy to find and retrieve?



Why not use a digital library? You could develop an organized collection of electronic versions of advisory documents using suitable software...

### Benefits of digital libraries

Mr.Parthan is excited and discusses the possibility with his NGO director. The director encourages him to explore further.

You have my full support. I can work with the state agriculture department in obtaining financial support, but you first need to establish its feasibility and carry out a concept demonstration. You have several issues to address...



ISSUES	Where can I find this information?
• Are there any freely available digital library software packages?	Unit 5 and 6
• What precautions should they be taken when digitizing?	Unit 4
• What document format is most suitable?	Unit 2
• What additional information should be provided to enable the easy identification of relevant documents?	Unit 3
• Can they deliver the collection on a CD-ROM?	Unit 5 and 6
• What is required to make the collection accessible over the Internet?	Unit 5
• What are the cost implications?	Unit 5



### Benefits of digital libraries

One can visualize several other situations where the application of digital library technology can bring significant practical benefits. The following are examples of such scenarios:



ElectroCon is a consulting firm providing consultancy services to a variety of clients in the electronics industry. The company's consulting experts conduct extensive research, field surveys and market studies. Their significant findings and data are included in their project reports.

Over a period of time the company accumulates a considerable number of these reports which form an invaluable 'knowledge asset' of the company. The managing director realizes that if this asset is effectively used, it can help in quick discovery of related findings and data and avoid their 'reinvention'. Setting up a digital library of these reports to provide fine-grained access to their content and associated data, would thus be a very good opportunity.

### Benefits of digital libraries



Universities have pioneered the setting up of websites. Commercial websites, despite their late start, have moved ahead in terms of advanced features and access to information about products and services offered by the company.

Many universities have woken up to the possibility of setting up 'institutional archives' (or 'repositories') consisting of publications of the faculty and research staff, and providing Internet access to these publications. Setting up a digital library of the intellectual output of a university would contribute to improved visibility and impact of its research.

While reviewing these scenarios, did you find any similarities with some of problems you encounter in your work environment? If so, then this self learning module will help you find solutions.

**What are digital libraries?**

We have seen the use and benefits of digital libraries.

But, what exactly is a digital library?

Imagine the hard disc of a personal computer, containing a random collection of documents, created and/ or downloaded and saved, over a period of time. Do you think we can call this a digital library?

- Yes.
- No.

Please click on the answer of your choice

**What are digital libraries?**

A digital library as “a collection of digital documents (or objects)” is a dominant perception today.

**Digital Library  
as “collection”**



“A digital library is an organized and focused collection of digital objects, including text, images, video and audio, along with methods for access and retrieval, and for selection, creation, organization, maintenance and sharing of the collection.”

Though the focus of this definition is on the document collection, it stresses the fact that digital libraries are **much more than a random assembly** of digital objects. They retain several qualities of traditional libraries such as a defined community of users, focused collections, long term availability, the possibility of selecting, organizing, preserving, and sharing resources. This viewpoint matches with the prevailing interest in the development of one-off, relatively static or special digital library collections. This is elaborated on the remainder of this module.

Digital libraries are sometimes perceived as “institutions” though this is not as dominant as the previous definition.

## What are digital libraries?

### Digital Libraries as "institutions"

The following definition given by the Digital Library Federation (DLF) brings out the essence of this perception.

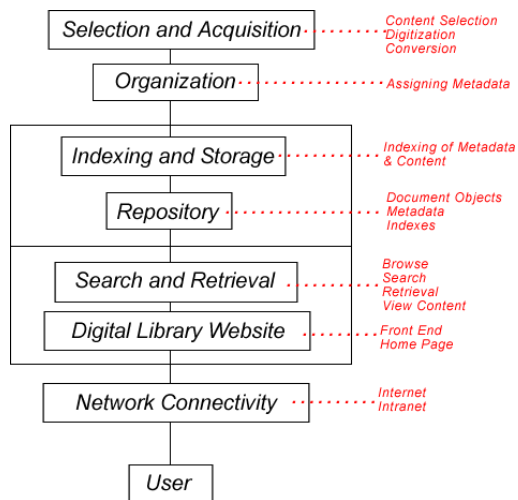
"Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities".

Stress here is on the digital library as a dynamic, growing organism. As digital libraries evolve and become the predominant mode of access to knowledge and learning, institutionalization of digital libraries appears to be an increasing possibility.

## Functional components of a digital library

Most digital libraries share a few common functional components. The diagram illustrates the key components.

The following slides give more information on each component



### Functional components of a digital library

*Selection and acquisition* – Typical processes covered in this component include: the selection of documents to be added, the digitization and/ or conversion of these documents to appropriate digital form.

*Organization* – A key process involved in this component is assignment of metadata (e.g. bibliographic information) to each document being added to the collection.

*Indexing and storage* – This component carries out indexing and storage of documents and metadata, for efficient search and retrieval.

*Repository* – This is the core component of the digital library, consisting of the document objects, metadata and indexes created for purposes of search and retrieval.

### Functional components of a digital library

*Search and retrieval* – This is the digital library front-end used by the end-users to browse, search, retrieve and view the contents of the digital library. This is typically presented to the users as an HTML page.

*Digital library website*: This is the server computer that hosts the digital library collection, and presents the collection to the user in the form of a website home page. The user selects a suitable link on this page to go to the search and retrieval front-end mentioned above. The digital library delivers the content based on search and retrieval operations. The digital library home page itself may be integrated with the library website through an appropriate hypertext link.

*Network connectivity* – For online access, the digital library website computer should have a dedicated connection to the intranet and/or Internet. Depending on the target user community, access may be restricted to the intranet (organizational LAN) or extended to external users through the Internet.

### Functional components of a digital library

Of course, at the other end there is the user, accessing the digital library using a web browser on her/ his networked desktop computer.

Digital libraries may be implemented in several ways and actual components depend on the access and delivery requirements, chosen implementation strategy and technology used. For example, in the absence of network connectivity (at the digital library and/or user side), it may be possible to deliver the digital library collection on CD-ROM media.

Digital libraries may also have **rights management** and **e-commerce components** to handle security and payment aspects.



These subjects are covered in the following sections of this module:

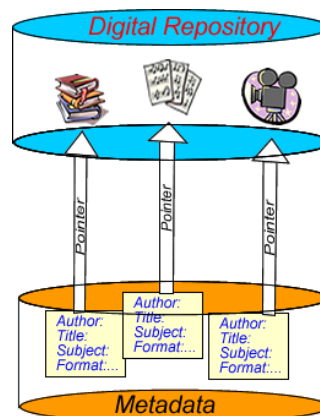
- Unit 5: Creation and sharing digital libraries
- Lesson 5.1: Creating a digital library collection
- Lesson 5.2: Facilities and requirements

### Functional components of a digital library

In addition to the digital documents stored in the digital library, metadata is probably the next most important characteristic of a digital library which separates it from other collections of online information.

As shown in the diagram, **each digital document** (and multimedia digital object) in the digital library **is assigned a set of metadata values that describes various attributes of the document** for purposes such as search and retrieval; identification and selection; administration; access management; preservation; information sharing; and so on.

Metadata is assigned manually and/ or extracted automatically when the document is added to the digital library. Various metadata schemes and standards are available today for use in new digital library projects.



This subject is covered in the following section of this module:  
Unit 3: Metadata standards and subject indexing

## Creation, distribution and provision of access to digital libraries

How does one go about creating a digital library collection?



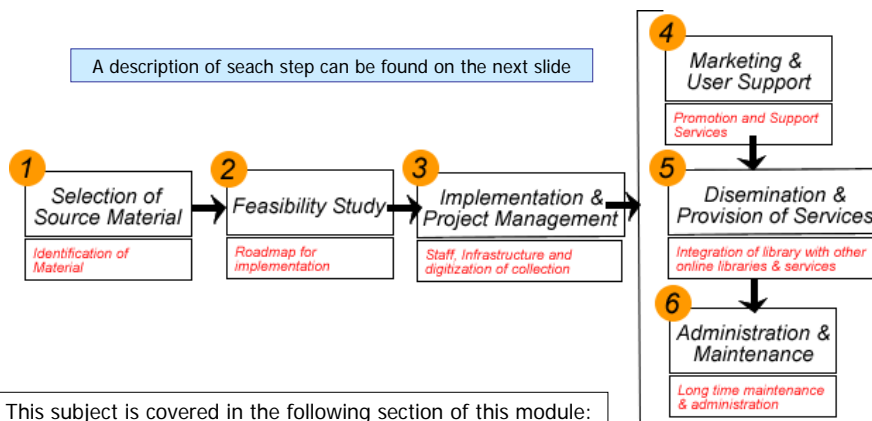
Simple and very small digital library collections can be created with minimal resources and effort.

However, careful planning and greater resources are required for creating large digital library collections involving extensive digitization and/or conversion efforts, 'live' collections that require frequent updating, collections that need to handle rare and special material, and collections that require extensive search and browse features.

Even small digital library projects can benefit from careful planning and implementation.

## Creation, distribution and provision of access to digital libraries

Here is a brief overview of steps involved in the creation and sharing of digital library collections.



This subject is covered in the following section of this module:  
Unit 5: Creation and sharing digital libraries  
Lesson 5.3: Creating a digital library collection: a Case Study

## Creation, distribution and provision of access to digital libraries

### Selection of source material

A digital library project usually begins with the identification and selection of the source material that is going to constitute the content of the digital library.

### Feasibility study

After obtaining management approval for the project, it is essential to conduct a feasibility study to assess the viability of the project, to address various factors associated with the digital library collection and to formulate a road map for implementation. Key factors include: objectives and goals; target user community; stake-holders; source material; copyright issues; file format options; access and delivery options; hardware and software requirements; personnel; cost; and implementation strategy and timeline.

### Implementation and project management

This includes setting up the necessary IT infrastructure (hardware and software), deployment of staff, managing the digitalization, and establishment of the digital library collection.

### Marketing and user support

This involves systematic promotion of the digital library for effective use and setting up of support services.

### Dissemination and provision of services

This involves integration of the digital library collection with other online library and information services.

### Administration and maintenance

The digital library should be administered and maintained to ensure its long term availability and use.

## Summary

Digital libraries enable the **creation, organization, maintenance, management, access to, sharing and preservation** of digital document collections.

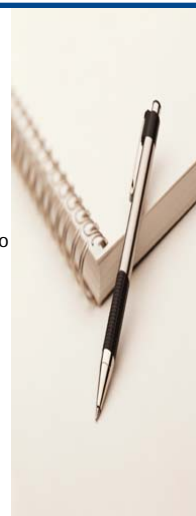
Digital libraries differ from traditional libraries by enabling users to **access and work directly with electronic documents**. This offers several significant benefits: access anytime anywhere; improved and wider access; new forms of access; improved resource sharing; and preservation.

Digital library applications can be conceived in a **very large number of areas** today including education, science, culture, development, health, governance and so on.

Digital libraries share several **common functional components**: document selection and acquisition; organization and loading; indexing and storage; repository; search and retrieval; digital library website; and network connectivity.

'**Metadata**' assigned to digital documents is a key component as this enables effective resource discovery.

Creating a digital library requires **careful planning**, followed by the implementation and management of the collection.



**If you want to know more...**

**Online Resources:**

Berkely Digital Library Sunsite (<http://sunsite.berkeley.edu/>)

Sun Microsystems. Digital library technology trends. August 2002.  
([http://www.sun.com/products-n-solutions/edu/whitepapers/pdf/digital\\_library\\_trends.pdf](http://www.sun.com/products-n-solutions/edu/whitepapers/pdf/digital_library_trends.pdf))

Peter Noerr. Digital Library Toolkit. Sun Microsystems, January 2003  
(<http://www.sun.com/products-n-solutions/edu/whitepapers/digitaltoolkit.html>)

**Additional Reading:**

Ian H. Witten and David Bainbridge 2003. How to build a digital library.  
Morgan Kaufman Publishers

Willaim Y Arms 2000. Digital libraries. The MIT Press

Christine L Borgman 2000. From Gutenberg to the Global Information  
Infrastructure. The MIT Press

