

Lectures on Digital Libraries

Historical Background & Definitions

Historical Background of DL

- Griffin (1998) briefly described that **scholarly communication among intellectuals** resulted fundamental basis for developing digital libraries in late **1980s**.
- The **National Science Foundation (NSF)** sponsored Santa Fe planning workshop on digital knowledge which was organized in March 1997 to strengthen this new concept as a system that brings together information collection, services and manpower for developing, preserving and dissemination of data, information and knowledge.
- In 1971, the **Project Gutenberg was started**. After this revolution, many other projects were started by different academic and private organizations. In Europe, China and India some other projects were also initiated.

What is a Digital Library?

- A digital library is an organized collection of information
 - A focused collection of **digital objects**
 - Methods for **finding, access** and **retrieval**
 - Methods for **selection, organization,** and **maintenance** of the collection
 - Methods for **preservation**

What is a Digital Library?

- A digital library is a collection of digital documents or objects.
- Digital library as an organized and focused collection of digital objects, including text, images, video and audio, with the methods of access and retrieval and for the selection, creation, organization, maintenance and sharing of collection (Smith, 2001).

What is a Digital Library?

- 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 (DLF).

What is a Digital Library?

Digital Library is a library in which a significant proportion of the resources are available in machine-readable format (as opposed to print or microform), accessible by means of computers.

The digital content may be locally held or accessed remotely via computer networks.

In libraries, the process of digitization began with the catalogue, moved to periodical indexes and abstracting services, then to periodicals and large reference works, and finally to book publishing. (ODLIS, 2015).

Digital Library Federation (DLF) <http://old.diglib.org>

A consortium of major libraries and library-related agencies dedicated to promoting the use of electronic technologies to extend collections and services, DLF is committed to identifying standards and best practices for digital collections and network access, coordinating research and development in the use of information technology by libraries, and assisting in the launch of projects and services that individual libraries lack the means to develop on their own.

World Digital Library (WDL)

<http://www.wdl.org>

In November 2006, the [Library of Congress](#) announced receipt of a \$3-million gift from Google to begin [digitizing](#) significant [primary source](#) materials from national and other major [research libraries](#) around the world, to make them freely available via the [Internet](#) in multilingual format.

Institutional Repositories (IR)?

1. “A **university-based institutional repository (IR)** 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” (ARL, 2012).

2. Institutional repositories (IR) also have roles beyond **disseminating and managing** the works of **individual scholars** that are part of the dialog of scholarly communications.

Institutional Repositories??

3. “Institutional Repository (IR) is a **digital collections** that capture and preserve the **intellectual output** of university communities” (Raym Crow, 2002).

4. “Institutional repositories (IRs) are critical to **developing, managing, and leveraging** digital contents and bringing greater value to **institutional output**” (Erv Blythe and Vinod Chachra, 2005).

Institutional Repositories??

5. “A set of services offered by a **university or group of universities** to members of its community for the management and dissemination of **scholarly materials** in digital format created by the **institution** and its community members, such as technical reports, theses and dissertations, data sets, and teaching materials” (ODLIS, 2015).

Continue.....

The responsibility of such materials needs their organization in a collective, [openly_accessible database](#) and a commitment to long-term [preservation](#) when appropriate.

Some IRs are also used as electronic presses to [publish e-journals](#) and [e-books](#).

IRs are part of a growing effort to reform [scholarly communication](#) and break the monopoly of [journal publishers](#) by repealing institutional control over the results of scholarship.

An IR may also serve as an indicator of the scope and extent of the university's [research](#) activities.

Establishing an Institutional Repository

Establishing an institutional repository is not a **simple, quick, or inexpensive** task.

To many, an institutional repository is **just technology**--a combination of a web server, relational database, and data storage accompanied by a wide spectrum of functionalities.

But the **building or purchasing of the technology** is just **one part** of an institutional repository project, and perhaps the **easiest part**.

Establishing an Institutional Repository

As universities implement institutional repositories (IR), not only do they have to consider **technical and training matters**, but they must also think about how to **develop policies** for **submitting items** to the IR, **digital rights management**, and similar issues that affect management of this database.

Scope of Digital Libraries

Scope of Digital Libraries

- ☐ Digital libraries offer unique ways of recording, preserving, and propagating culture in multimedia form.
- ☐ Examples:
 - ☐ Support human development
 - ☐ Preserving a traditional culture
 - ☐ Exploring popular resources

Ideas are Being Replaced

- ☐ **Focused Collections**

- ☐ Rather than the Alexandrian model

- ☐ **Service to library users**

- ☐ Rather than massive storehouse of information

Types of DL

- ☐ Born digital libraries
- ☐ Digitized from physical medium
- ☐ Hybrid libraries
- ☐ Stand-alone Digital Library (SDL)
- ☐ Federated Digital Library (FDL)

Born Digital Libraries

Documents developed in digital form i.e. all computer files in different formats (doc, xml, ppt, html, pdf, etc.).

Digitized from physical medium

Print form of material converted into digital form by imaging/scanning.

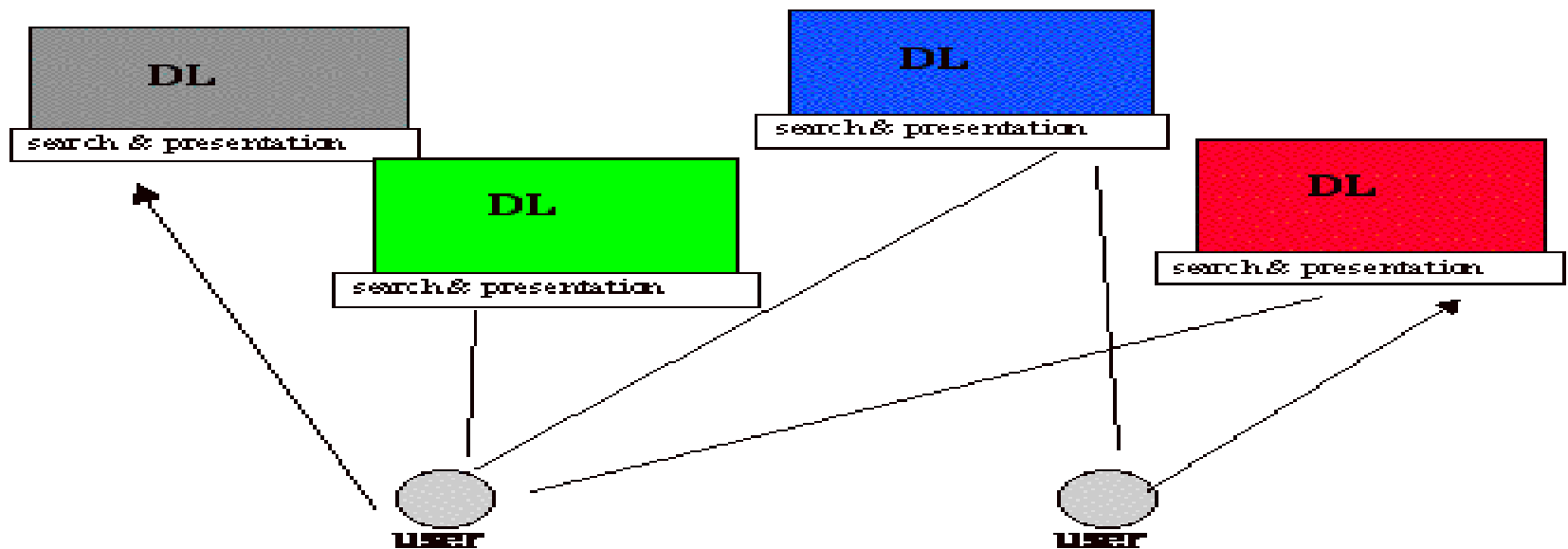


Hybrid libraries

- ✓ The hybrid library is a term used by **librarians** to describe libraries containing a mix of traditional print **library** resources and the growing number of **electronic resources**.
- ✓ Hybrid libraries are mixes of traditional print material such as books and magazines, as well as electronic material such as downloadable **audio books**, **electronic journals**, **e-books**, etc.
- ✓ Hybrid libraries are the new model in most public and **academic libraries**.

Stand-alone Digital Library (SDL)

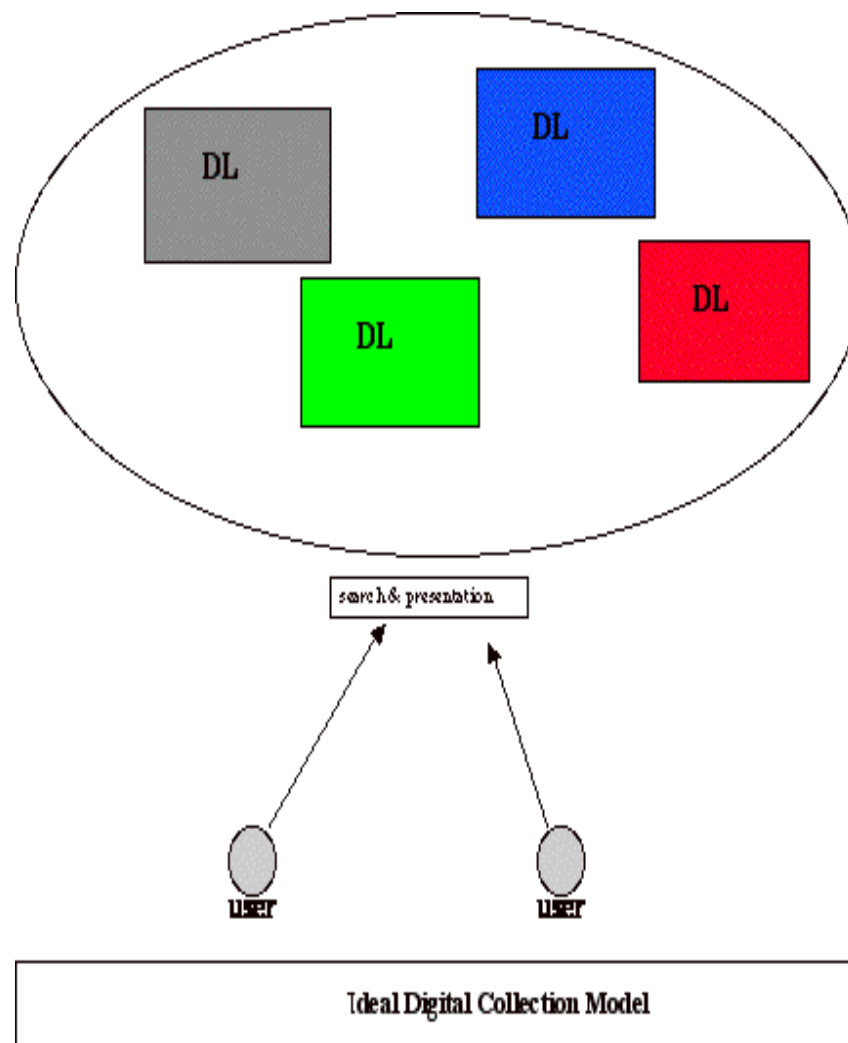
It is a classical library implemented in a fully computerized fashion. SDL is simply a library in which the holdings are digital (scanned or digitized). The SDL is self-contained - the material is localized and centralized.



"Traditional" Digital Collection Model

Federated Digital Library (FDL)

This is a federation of several independent SDLs in the network, organized around a common theme, and coupled together on the network. A FDL composes several autonomous SDLs that form a networked library with a transparent user interface. The different SDLs are connected via communication networks.



Harvested Digital Library (HDL)/Virtual DL

This is a virtual library providing summarized access to related material scattered over the network. Examples of HDLs are the Internet Public Library (IPL).

- 1.A HDL holds only metadata with pointers to the holdings that are "one click away" in cyberspace.
- 2.Developed by Library Professionals, or Computer Scientists

Advantages

- It saves and/or reduces the physical space taken up by library materials.
- It enhances searching capabilities in a digital format.
- The library materials are available at the user's desktop, regardless of where the user is physically located.
- It provides the user with the capability to download and manipulate text.
- It often allows for multiple users.
- It eliminates the problem of a book being missing or off the shelf.
- It requires less labour in searching the relevant material.
- Easy access.
- Access to multiple media, text, pictures, movies sounds etc.
- Time saving.
- Round the clock availability. Real time access.

Disadvantages of Digital Library

- High cost.
- Maintenance of technology
- Sustainability
- Copyright/intellectual property rights
- Long term planning.
- Electricity failure.
- The digital/virtual library relies on computer in order to be available for use.
- Users can't spread everything out in front of them and use it all at once.
- Every product has its own distinct user interface.
- It is not cost saving.
- Technology obsolescence.

Types of DR

- Subject based DR
- Research based DR
- National DR
- Archive based DR
- Magazine/Newspaper DR
- Etc

Pakistan Research Repository

- Developed by HEC in 2006
- Offers searching and browsing
- Search by Subject, Title, Authors etc
- M. Phil & PhD thesis Collection, Journals, Magazines, text books
- First developed by using E-Print & currently migrated data to DSpace
- Offering Open-access

Proposal Components

- ☐ Title page
- ☐ Table of Contents
- ☐ Summary/Abstract
- ☐ Introduction
- ☐ Statement of need
- ☐ Goals / outcomes
- ☐ Work plan
- ☐ Budget
- ☐ Evaluation / Assessment plans
- ☐ Sustainability Statement
- ☐ Promotion Strategies

Introduction to Proposal

- ☐ Describe the institution and its community
- ☐ What is the significance of the content you plan to digitize?
- ☐ Does your institution have a track record with digital projects?

Statement of Need

- ☐ What need will be addressed?
- ☐ What is the significance of the project?
- ☐ How the need matches funding institution's mission?

Goals / Objectives of Project

- ☐ How does project meet the mission of the institution?
- ☐ How does the project provide a solution to the need stated earlier?
- ☐ Who is involved?
- ☐ Who is being served?
- ☐ Is it realistic or overly ambitious?

Project Work Plan

- What is the quantifiable goal?
- What is the work plan to accomplish project?
 - Timeframe
 - Space
 - Equipment
 - Staffing
 - Software
 - Metadata

Project Actions Timeline

☐ **Who is responsible for what???**

Responsible staff could be:

Chief Librarian, Librarian, Assistant Librarian, catalogers, IT staff, etc. or hired

project staff and/or project technicians

For Example; Who will:

- ☐ scan items
- ☐ add scanned data
- ☐ create metadata
- ☐ Train the (hired) Project staff (i.e., who will train, when, where and how)
- ☐ purchase the equipment
- ☐ be responsible for promotion, (i.e., contact with Media & online groups; official opening; presentations organized for publicity etc.)

☐ **Timeline for all the steps taken for the Project**

Proposed Project Budget

- ☐ Salaries & Benefits
- ☐ Library materials
- ☐ Operation Equipment
- ☐ Indirect Costs
 - ☐ Institutions contribution, it can be a percentage of a permanent employees salary/benefit; library materials and equipment, building space and utilities, and other operating or overhead costs.
 - ☐ Under operation budget for training and travel that might be necessary

Evaluation / Assessment plans

□ Who will evaluate, how and when???

(Most effective Managements techniques can be used for continuous monitoring, and/or evaluation of a project/program, i.e., PERT)

Sustainability...

How the proposed system/program
will sustain???

Think about the new
avenues!!!

Promotion & Publicity for Completed Project: Few tips

- ☐ Opening day celebration with traditional programs
- ☐ Arrangement of dinner honoring all the partners, and the Project staff;
- ☐ Distribution of brochure;
- ☐ Advertisement through institutional magazine, local newspaper and television networks.
- ☐ Project presentations
- ☐ Write to your Govt. authorities, i.e., Parliamentarians, senators, Prime and chief minister, governor, etc. (to let them know about the successful project (If funded using government funds).

Infrastructure for Developing Digital Library

- **Infrastructure for Developing Digital Library**
- Development of digital libraries requires technological infrastructure, highly professionalized skilled staff and budget to develop and sustain digital information system.
- Following infrastructure is usually required for developing digital libraries:

- Hardware (main server, web server, file server, e-mail server, nodes, computers, laptops, printers, scanners, digital camera, video cameras, audio recording machine)
- Software (system software, Windows XP, Linux), (application software:, web browser,)
- OCR software: Adobe Fine Reader, Omani page, file format converter, Adobe distiller
- Web Server Software: Apache etc.
- Database Software: MySQL, ORACLE, MS access, MSSQL
- Anti Virus Software: Microsoft Essentials, AVG, Notion (Registered Mandatory)
- Digital Library Software: Eprint, Dspace, Greenstone
- LAN and Internet Infrastructure (Internet connection with high bandwidth, modem, switch, fiber optical cables)
- UPS 10 KV
- Air-condition computer lab.
- Well-trained library staff

Challenges

The staff of the National Digital Library Program at the Library of Congress (2003) have identified **ten challenges** that must be met in building large and effective digital libraries.

Building the Resource

- *Challenge One:* Develop improved technology for digitizing **analogue materials**.
- *Challenge Two:* Design search and retrieval tools for cataloguing or descriptive information.

Interoperability

Challenge Three: Establish protocols and standards to facilitate the assembly of distributed digital libraries.

Intellectual Property

Challenge Four: Address legal concerns associated with access, copying, and dissemination of physical and digital materials.

Effective Access

- *Challenge Five:* Develop approaches that can present **heterogeneous resources in a logical way.**
- *Challenge Six:* Make the Digital Library **useful to different communities of users** and for different purposes.
- *Challenge Seven:* Provide more efficient and more flexible tools for **transforming digital content** to fulfil the needs of end-users.

Sustaining the Resource

Challenge Eight: Develop economic models for the support of the Digital Library.

Designing Principles

Eight general principles by Arms, (2000) for the design of digital libraries

1. The **technical framework** exists within a **legal and social framework**.
2. Understanding of digital library concepts and **standard terminology**.
3. **Names and identifiers** are the basic building blocks for the digital library.

4. Digital libraries are more than **collections of bits**.
5. The digital library object that is used is **different from the stored object**
6. **Repositories must look after the**
information they hold
7. Users want **intellectual works, not digital objects**.

Digital Collection Development

Selection of Materials

- In conventional libraries the budget could be a limitation but here the task is much more complicated due to problem of plenty.
- Moreover, for identification of printed materials there are well-established book trade and selection tools like **Books-in-Print**, **National Bibliographies**, **Union Catalogues** etc. along with the users **suggestions and approvals**.
- In case of **digital contents** such tools are not yet available. Normally some **staff members** are assigned the task of finding new and relevant information resources from the **Internet using search engines**. **Vendors of e-journals and e-books** are also sources of discovery of digital contents.

Selection Criteria

Major factors to be considered for the **selection of digital information resources** includes the following:

- Appraisal of the subject and discipline in relationship to the institution's collection goals;
- Content, Quality, Currency and uniqueness of the content..
- Version of the product, Pdf. CD-ROMs.
- Price, licensing and copyright agreements.
- Information Retrieval system and user interface.
- Archiving procedures.

Digitization

Digital documents may be **born-digital**, created using digital publishing tools (e.g. Word, ppt., excel, etc.) or created by **converting** from an analogue format to digital format or **converted from one digital format to another** to suite the requirements of a particular Digital Library.

The process of capturing and converting from analogue to digital format is often called as **digitization**.

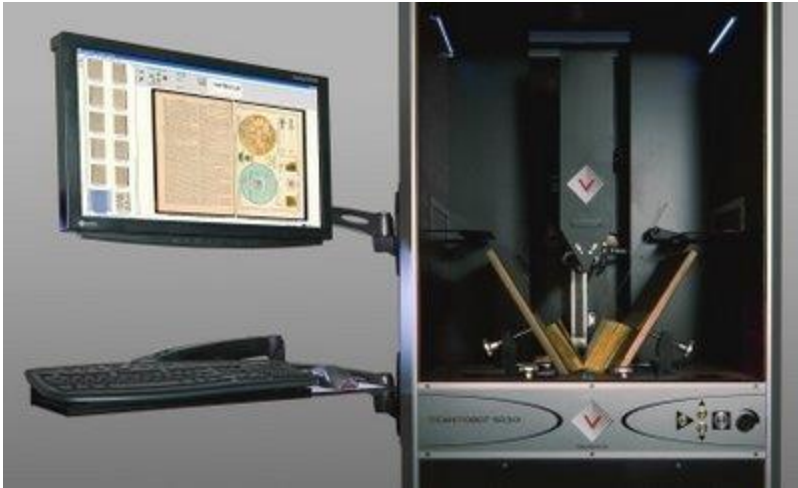
Scanners and Scanning

The first step in converting paper documents into a digital library collection is to obtain images of pages of publications in digital format.

Scanners

- Scanners are available in all price ranges, and all shapes and sizes. They range from Rs. 5000 for flat-bed scanners to upwards of Rs. 25,00,000 for large industrial scanners .
- The output format of a scanned page is a computer file that is usually stored in **TIFF or Bitmap** format.
- An average page scanned and converted to this format occupies only 50 Kb, compared to perhaps 2 Mb (2100 Kb) for the equivalent page in uncompressed Bitmap form.

Scanners



Robotic Scanner



Bookeye



Scanning Programs

Every scanner comes with its own software, which means, that the program must be installed on the computer that manages the scanner. Some have a computer card that needs to be installed in your computer to speed up the scanning operation.

Scanning Process

Using software provided with the scanner, a digital image of each paper page is scanned and transformed into a Bitmap or TIFF image.

OCR: Optical Character Recognition

The following steps are involved in converting paper documents to computer form:

1. Page layout analysis
2. Recognition
3. Scanning images and tables

- **Intellectual Property**

- “Asset value of an intellectual activity such as writing of novel, formulation of patents, compilation of glossary.
- It includes the rights provided by the laws of patents, copyrights, trademarks and trade secret, rights of publicity.

- **Types of Intellectual Property**

- Trademarks (It is protect by any word, name, symbol, divide)
- Patents (License)
- Industrial design
- Registration
- Trade secrets
- Copyright

- **World Intellectual Property Organization Copyright Treaty (WIPO)**
- It is an intellectual treaty on copyright law adopted by the member state of the WIPO in 1996. It provides additional protection for copyright.

- **IPO of Pakistan**

- The new concept of IP was started in Pakistan during 2005. The cyber act was passed in 2005 to prevent the cyber-crime.

Z39.50 Protocol

- The name "Z39.50" comes from the fact that National Information Standards Organization (NISO), the American National Standards Institute (ANSI)-accredited standards development organization serving libraries, publishing and information services, was once the Z39 committee of ANSI.

- NISO standards are numbered sequentially and Z39.50 is the **fiftieth** standard developed by NISO. The current version of Z39.50 was adopted in **1995**. It is sometimes referred to as **Z39.50 Version 3**.

Z39.50 Protocol

- **Z39.50** is generally defined as the information search and retrieve protocol standard used primarily by library and information related systems.
- **Z39.50** standard specifies a client/server-based protocol for searching and retrieving information from remote databases simultaneously using a single interface.

Z39.50

- **Z39.50** is an international standard defining protocol a for computer-to-retrieval computer information
- **Z39.50** makes likely for a user in one system to search and retrieve information from other computer systems that have also implemented Z39.50 (i.e. LOC gateway)
- The **Library of Congress** is designated as Maintenance Agency and Registration Authority for ANSI/NISO Standard Z39.50 and ISO 23950.

DIGITAL LIBRARY SOFTWARE

Greenstone Software

- ❑ Greenstone Digital Library software
- ❑ New Zealand Digital Library Project
- ❑ University of Waikato
 - ❑ Open Source
 - ❑ Multilingual software
 - ❑ Developed and distributed in cooperation with UNESCO and Human Info NGO Belgium
- ❑ Issued under the GNU General Public License (The **GNU General Public License (GNU GPL** or simply **GPL**) is the most widely used free software license, originally written by Richard Stallman for the GNU project)

Greenstone Digital Library Software

The screenshot shows a Windows Internet Explorer browser window displaying the Greenstone Digital Library Software website. The address bar shows the URL <http://www.greenstone.org/>. The browser's taskbar at the bottom shows several open applications: "Welcome :: Greenst...", "iGoogle - Windows I...", "Microsoft PowerPo...", "Assignment topics f...", and "web2-ata-assignme...".

The website has a green header with navigation links: [en](#), [fr](#), [ru](#), [es](#), [ro](#), [ar](#). Below the header is a secondary navigation bar with links: [Home](#), [Download](#), [Examples](#), [Wiki](#), [Support](#), and [Developers](#).

On the left side, there is a sidebar with a "Greenstone Blog" section. It includes links to "Planet Greenstone", "Calendar", and "Workshop Map". A recent blog post is displayed with the title "First operational Greenstone OAI-PMH collections", dated "Wednesday 23rd December, 2009", and categorized as "Other News". The post text describes the implementation of OAI-PMH collections for public data harvesting, mentioning the "Biblioteca Digital Científica de les Illes Balears" in Spain and the "Memoria Académica" of the Faculty of Humanities and Educational Sciences of the National University of La Plata in Argentina.

The main content area features a section titled "About Greenstone". The text describes Greenstone as a suite of software for building and distributing digital library collections, produced by the [New Zealand Digital Library Project](#) at the [University of Waikato](#), developed in cooperation with [UNESCO](#) and the [Human Info NGO](#). It is *open-source*, *multilingual* software, issued under the terms of the GNU General Public License. A link to the [Greenstone Factsheet](#) is provided for more information.

Below the "About Greenstone" section, the text states the aim of the software is to empower users, particularly in universities, libraries, and other public service institutions, to build their own digital libraries. It mentions that digital libraries are radically reforming how information is disseminated and acquired in UNESCO's partner communities and institutions in the fields of education, science and culture around the world, and particularly in developing countries. Further information can be found in the book [How to build a digital library](#), authored by two of the group's members.

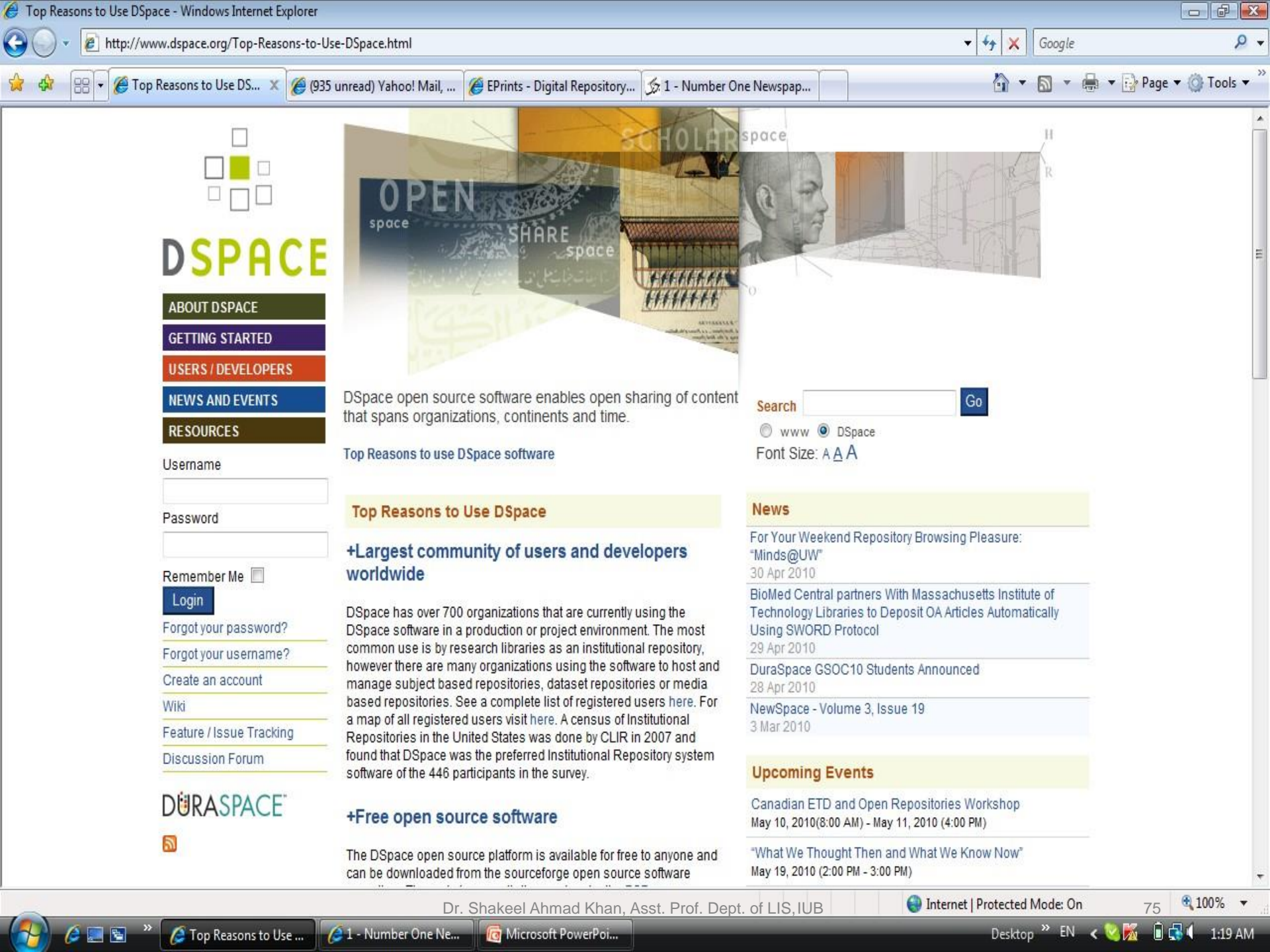
The text continues by stating that the complete Greenstone interface, and all documentation, is available in *English, French, Spanish, Russian and Kazakh*. Greenstone also has interfaces in many [other languages](#). They are looking for [volunteers](#) to add new language interfaces and help maintain existing ones.

Under the heading "International Links:", there are links to [Greenstone support for Southern Africa](#), [Greenstone support for South Asia](#), and [Grupo de Usuarios de Habla Hispana de Greenstone](#).

At the bottom of the page, a status bar shows "10 items remaining) Downloading picture http://greenstone.org/images/sitesbar/background.png...". The system tray at the bottom right shows the time as 5:26 PM and the date as 7/3.

Features of Greenstone Software

- Access through Web browser
- Windows or Unix
- **Latest Distribution (2.85)**
- Searching
- Browsing
- Easy to maintain
- Various metadata
- Plug-ins for new document types
- Multi languages
- Ten South-Asian Language Interfaces for Greenstone (Urdu)
- Text, pictures, audio, video
- Hierarchical phrase and key-
- Multi-gigabyte
- Compression
- Password Protection
- User logs
- Administrative functions
- Updates dynamically without bringing system down
- Publish to CD-ROM
- Uniform presentation across different computers
- Open-source software



- ABOUT DSPACE
- GETTING STARTED
- USERS / DEVELOPERS
- NEWS AND EVENTS
- RESOURCES

Username

Password

Remember Me ☐

Login

[Forgot your password?](#)

[Forgot your username?](#)

[Create an account](#)

[Wiki](#)

[Feature / Issue Tracking](#)

[Discussion Forum](#)



DSpace open source software enables open sharing of content that spans organizations, continents and time.

Top Reasons to use DSpace software

Top Reasons to Use DSpace

+Largest community of users and developers worldwide

DSpace has over 700 organizations that are currently using the DSpace software in a production or project environment. The most common use is by research libraries as an institutional repository, however there are many organizations using the software to host and manage subject based repositories, dataset repositories or media based repositories. See a complete list of registered users [here](#). A census of Institutional Repositories in the United States was done by CLIR in 2007 and found that DSpace was the preferred Institutional Repository system software of the 446 participants in the survey.

+Free open source software

The DSpace open source platform is available for free to anyone and can be downloaded from the [sourceforge](#) open source software

Search Go

☐ www ☒ DSpace

Font Size: A A A

News

- For Your Weekend Repository Browsing Pleasure: "Minds@UW" 30 Apr 2010
- BioMed Central partners With Massachusetts Institute of Technology Libraries to Deposit OA Articles Automatically Using SWORD Protocol 29 Apr 2010
- DuraSpace GSOC10 Students Announced 28 Apr 2010
- NewSpace - Volume 3, Issue 19 3 Mar 2010

Upcoming Events

- Canadian ETD and Open Repositories Workshop May 10, 2010(8:00 AM) - May 11, 2010 (4:00 PM)
- "What We Thought Then and What We Know Now" May 19, 2010 (2:00 PM - 3:00 PM)

DSpace Software

- DSpace Digital Library software
 - MIT/HP
 - DSpace Federation
 - Open Source
 - Extensible
 - Issued under the BSD General Public License
- (The phrase **BSD licenses** represents a family of [permissive free software licenses](#).
The original was used for the [Berkeley Software Distribution](#) (BSD), a [Unix-like](#) operating system after which the license is named.

DSpace

- DSpace: “DSpace is a ground-breaking digital repository system that captures, stores, indexes, preserves, and redistributes an organization's research data. The DSpace software platform serves a variety of digital archiving needs.”
- An open source software
- Example use:
 - American Museum of Natural History Research Library
 - Chapel Hill, Theses & Dissertations
 - University of Cambridge – Academic & related content
 - Edinburgh Research Archive (ERA)

EPrints

- Eprints is a generic archive software under development by the University of Southampton. It is intended to create a highly configurable web-based archive.
- Open Source software
- Example uses:
 - HEC: Pakistan Research Repository
 - Southampton Crystal Structure Report Archive
 - Central European University – Preprint Archive
 - DLIST – Digital Library of Information Science & Technology

Fedora

- Fedora: “Open source software that gives organisations a flexible service-oriented architecture for managing and delivering their digital content.”
- Open source software
- Example uses:
 - Digital Case, Case Western Reserve University's electronic repository and archive: stores, disseminates, and preserves faculty research in digital formats (both born digital and digitised)
 - University of Queensland eSpace – research digital repository with published articles and conference papers, book chapters, theses and other forms of written research

Commercially Available DL software

CONTENTdm: Digital Collection Management Software

CONTENTdm [OCLC - Digital Collection Services] - Windows Internet Explorer

http://www.oclc.org/contentdm/

Search: Go

OCLC® The world's libraries. Connected.™

Worldwide (English) [Change](#) [About OCLC](#) [Public Purpose](#) [News and Events](#) [Contact us](#)

WorldCat Membership Products and Services Research

◀ Digital Collection Services

CONTENTdm

- At a glance
- Overview
- Options
- Frequently asked questions
- CONTENTdm in action
- Free evaluation
- Ordering
- Support, training and documentation

View a demo

Request an evaluation

Login to the User Support Center

View recent Web sessions

Subscribe to OCLC updates and offers

CONTENTdm®
Digital Collection Management Software

AT A GLANCE

CONTENTdm helps you to make your digital collections available on the Web, more quickly and easily. No matter the format—local history archives, newspapers, books, maps, slide libraries or audio/video—CONTENTdm can handle the storage, management and delivery of your collections to users across the Web.

[View complete at a glance >>](#)

[Download the brochure](#) [Contact us](#)

NEWS

- [North Carolina State Archives enters the 150,000th record](#)
- [CONTENTdm Featured Collections: May 2010](#)
- [Making the history of Seattle more accessible](#)

[More news >>](#)

Why do nearly 2,000 organizations use CONTENTdm?

- [A complete solution for digital collections](#)
- [Easy to use](#)
- [Discovery through WorldCat](#)
- [Unequaled range of options](#)

Featured collection

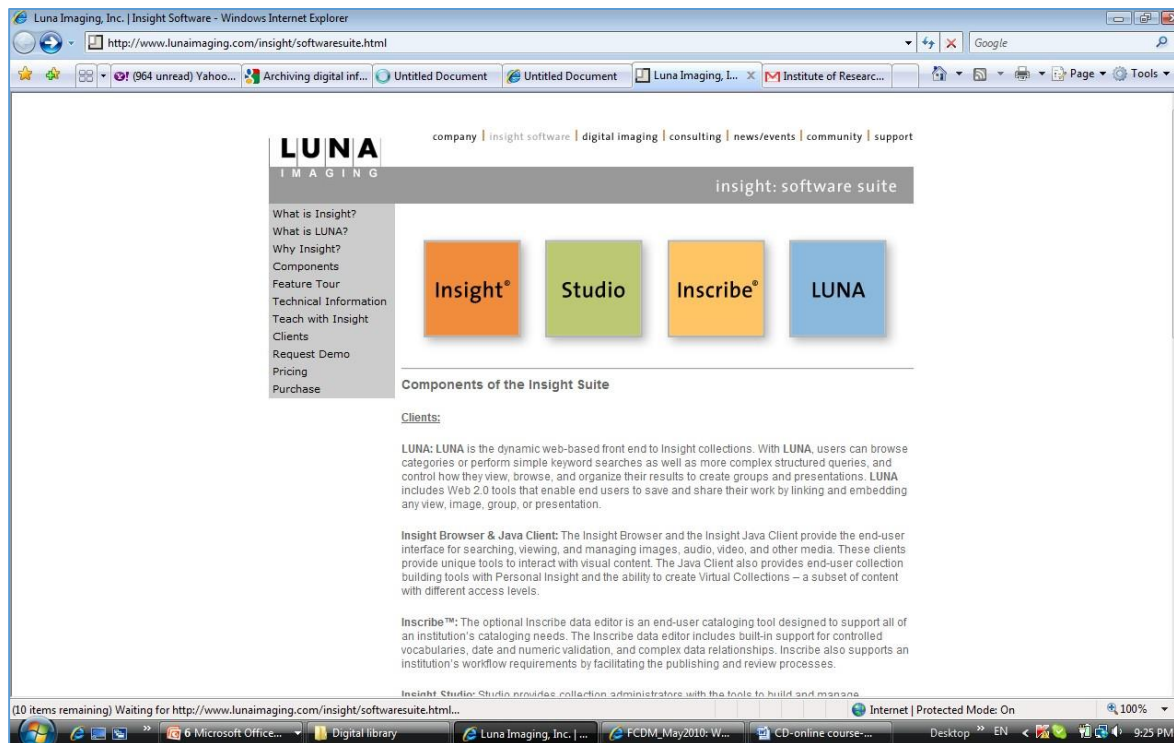
[Connecticut State Library Digital Collections](#)

Dr. Shakeel Ahmad Khan, Asst. Prof. Dept. of LIS, IUB

Internet | Protected Mode: On 81 100%

5 Microsoft O... Digital library CONTENTdm [...] FCDM_May2010... 78% of eaolug-... CD-online cours... Desktop » EN 9:17 PM

LUNA: the dynamic web-based front end to Insight collections



Digital Storage Devices

Storage

- ☐ Primary storage
 - ☐ Volatile storage
 - ☐ Temporary storage

RAM

- ☐ Secondary storage
 - ☐ Nonvolatile storage
 - ☐ Permanent storage
- ☐ Secondary storage characteristics
 - ☐ Media
 - ☐ Capacity
 - ☐ Storage devices
 - ☐ Access speed





Hard Disk Drive

Storage Devices



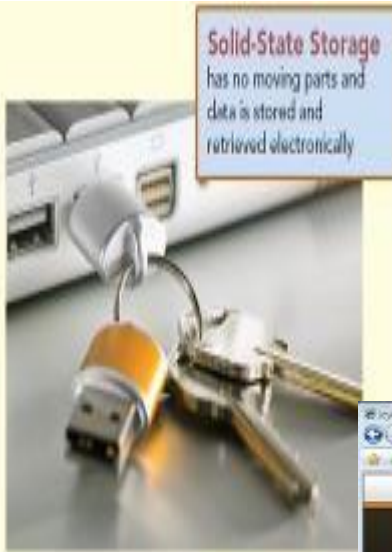
MP3 Player

Zip – 100 MB



**External Drive (Passport Drive)
320 GB**

Hard Disks
use magnetic charges to record data, have large storage capacities and fast retrieval times



Solid-State Storage
has no moving parts and data is stored and retrieved electronically



Optical Discs
use reflected light to record data and have very large storage capacities



Internet Hard Drives
provide Web-based, low-cost storage

Mass Storage Devices
are specialized high-capacity storage devices for organizational data



Internal Hard Disk

- Usually mounted inside the computer's system unit.
- Can store billions of characters of data.
 - Stated in forms of bytes:
 - Megabytes or Gigabytes
- Designated as the C-drive
- Advantages over other removable media
 - Capacity
 - Access speed



External Hard Disks

- Removable hard disks
- Used to complement internal hard disk
- Capacities of 20 to 500 GB



Hard-disk cartridge



PC Card Hard Disks

Digital Versatile Disks

- DVD stands for digital versatile disk or digital video disk
- DVD-ROM
- DVD+R and DVD-R
- DVD+RW, DVD-RW, and DVD-RAM



Performance Enhancements???

- Disk caching
- Redundant arrays of inexpensive disks (RAID)
- File compression and decompression



Technique	Description
Disk caching	Uses cache and anticipates data needs
RAID	Linked, inexpensive hard-disk drives
File compression	Reduces file size
File decompression	Expands compressed files

Digital Audio Tape

- Method of storing large amounts of data on tape using helical scan technology to write data at high densities across the tape at an angle.

Optical Storage

- ☐ Means of recording data as light and dark spots on CD or DVD.
- ☐ Reading is done through a low-power laser light.

Optical Disks

- Hold over 50 gigabytes of data
- Three types
 - Compact Disc (CD)
 - Digital Versatile Disc (DVD)
 - Hi-Def Disc



Compact Disc: CD-ROM

- ❑ “CD – Read Only Memory”
- ❑ Also called CD-R
 - ❑ CD-Read
- ❑ Storage device that uses laser technology to read data that is permanently stored on compact disks, cannot be used to write data to a disk.
- ❑ Optical format
- ❑ From 650 MB to 1 GB capacity
- ❑ Rotation speeds vary
- ❑ Three basic types
 - ❑ Read only: CD-ROM
 - ❑ Write once: CD-R
 - ❑ Rewriteable: CD-RW



Digital Versatile Disc

- Digital Versatile Disc or Digital Video Disc (DVD)
- Similar to CDs, but can store more data
- Three basic types
 - Read only: DVD-ROM
 - Write once: DVD+R; DVD-R
 - Rewritable: DVD+RW; DVD-RW; DVD-RAM



High-Definition Disc

- Next generation of **optical disc**
- Far greater capacity than DVDs
- Three basic types
 - Read only
 - Write once
 - Rewriteable
- Two competing **hi def** formats
 - **HD DVD**
 - **Blu-Ray**

Format	Typical Capacity	Description
CD	650 MB to 1 GB	Once the standard optical disc
DVD	4.7 GB to 17 GB	Current standard
Blu-Ray	25 GB to 50 GB	Hi-def format, large capacity

Solid-State Storage

- **Flash memory cards**
 - Widely used in notebook computers
 - Used to record MP3 music files



Flash memory card

What is Metadata?

- The term metadata is used for description of digital data. It is nothing but data about data.
- “Metadata are structured information used to find, access and manage information resources primarily in digital format”. It is called:

- “A digital card catalogue”
 - “Data about data”
- “Structured information about digital contents”

- There are two major factors that influence the development of metadata:
- The need for systematic discovery and retrieval of networked resources.
- The ability to embed metadata in the digital object.

- **Role of Metadata in Digital Libraries**

- It manages digital resources.
- It provides administrative control.
- Used for rights management.
- Offers interoperability.
- Provides searching, browsing, advance search
- Discovery of resources.

- **Characteristics of Metadata**

-

- Followings are the basic characteristics of Metadata:

-

- **Semantic:**

- It refers to the number and type of metadata contents that include title, creator, date etc.

-

- **Syntax:**

- It refers to the way in which contents are structured by using specific grammar i.e SGML (Standard General Markup Language), XML (Extensible Markup Language), HTML (Hypertext Markup Language), RDF (Resource Description Framework).

- **Structure:**

- It provides a complete and over all information architecture that contains metadata elements and syntax.

- **Interoperability:**

- It refers to the ability to transfer metadata among different schemas and information system. It helps us in sharing metadata information of digital contents among various heterogeneous information systems.

- **Extensibility:**

- It gives the opportunity to extend the existing meta-data and develop the required elements as per meta-data requirements and offer more precise information retrieval.

- **Requirements for Creating Quality Metadata**
- Standards Metadata Sets: (Dubline Core, Z39. 50, AACR2, GILS (Government Locator Service), AGLS (Australian Government Information Locator Service).
- Syntax for Metadata: HTML, SGML, XML, RDF

METADATA AND OTHER STANDARDS

DUBLIN CORE

<http://dublincore.org>

DUBLIN CORE

- The name "Dublin" is due to its origin at a 1995 invitational workshop in Dublin, Ohio, USA; "core" because its elements are broad and generic, usable for describing a wide range of resources.
- The Dublin Core Metadata Initiative Limited is registered in Singapore.

DUBLIN CORE: Introduction

- The Dublin Core Metadata Initiative **Conference and Workshop Series** **gathers experts** from various areas in the private and public sectors, including the library world, the networking and digital library research communities, corporate knowledge management, e-government and public sector information.
- Initially organized as a series of invitational workshops with a main focus at the **definition and finalization of the Dublin Core metadata standards**, this has developed into an annual get-together of people involved in the work of DCMI and other metadata activities.
- **Since 2001**, the event includes a Tutorial track, a peer-reviewed Conference track, Special Sessions and Working Group meetings attended by hundreds of experts in the field from many countries around the world.

DUBLIN CORE: METADATA

- The Dublin Core Metadata standard provides core metadata vocabularies for **describing** a wide range of **networked resources**. The Dublin Core element set was originally developed in **English**, but versions are being created in many other languages
- The Dublin Core standard includes two levels: **Simple and Qualified**.
- Simple Dublin Core comprises **fifteen elements**.
- Qualified Dublin Core includes **three additional elements**.

DUBLIN CORE: Mission and Principles

- Providing **open access** to education, training and documentation resources related to innovative design and metadata best practices.
- Supporting a worldwide community of people working with metadata to share experiences and find common solutions through collaborative tools, publications and meetings.
- Promoting **co-operation**, and interoperability across standards and vocabularies, by engaging with other organizations and communities.
- Developing and maintaining the DCMI metadata vocabularies and **promoting their use** in conjunction with other vocabularies for describing resources.

DCMI Home: Dublin Core® Metadata Initiative (DCMI) - Windows Internet Explorer

http://dublincore.org/

Google

★

☰

✉ (941 unread) Yahoo! Mail...

UNESCO in the Spotlight...

World Digital Library Ho...

Using Dublin Core

DCMI Home: Dublin ...


🏠

🔍

🖨

📄 Page

⚙ Tools



Dublin Core® Metadata Initiative

Making it easier to find information.

Home

Metadata Basics

DCMI Specifications

Community and Events


About Us

Search

The Dublin Core® Metadata Initiative

The Dublin Core Metadata Initiative, or "DCMI", is an open organization engaged in the development of interoperable metadata standards that support a broad range of purposes and business models. DCMI's activities include [work on architecture and modeling](#), discussions and collaborative work in [DCMI Communities and DCMI Task Groups](#), annual [conferences and workshops](#), standards liaison, and educational efforts to promote widespread acceptance of metadata standards and practices.

You can learn more about metadata and DCMI by exploring the pages listed in the menu bar above: [the Home page](#) (this page), [Metadata Basics](#), [Specifications](#), [Community and Events](#), and [About Us](#).



DC-2010

International Conference on Dublin Core and Metadata Applications

Pittsburgh, 20-22 October 2010

DC-2010 Call for Papers now closed

2010-04-15, The deadline for submissions for [DC-2010](#), the tenth International Conference on Dublin Core and Metadata Applications, to be held in Pittsburgh, PA, USA, 20-22 October 2010, passed on 9 April. The Program Committee received Full Papers, Project Reports, and Posters from more than a dozen countries, from a variety of organizations including universities and research institutes, national libraries, museums, state archives, financial institutions, and governmental and international agencies. Author notification for the conference is set for 11 May 2010.

DCMI Status report March 2010 published

2010-03-29, The DCMI Managing Director has published the semestral [March 2010 DCMI Status Report](#).

New Advisory Board member

2010-03-29, We are welcoming Paul Walk, deputy director of UKOLN, University of Bath, as a member of the [DCMI Advisory Board](#).

DCMI co-sponsors Balisage


2010-03-29, DCMI has become a co-sponsor for [Balisage](#), the annual conference devoted to the theory and practice of descriptive markup and related technologies for structuring and managing information, to be held in Montréal, Canada, 3-6 August 2010. DCMI participants are eligible for discount registration.

Dr. Shakeel Ahmad Khan, Asst. Prof. Dept. of LIS, IUB

Internet | Protected Mode: On

111

100%



DCMI Home: Dublin... | YouTube - World Di... | (940 unread) Yahoo!... | Download Java for ...

Desktop » EN

4:07 PM

15 Dublin Core Elements

- | | |
|-----------------------|-----------------------|
| 1. Creator | 9. Format |
| 2. Title | 10. Coverage |
| 3. Subject | 11. Rights |
| 4. Contributor | 12. Relation |
| 5. Date | 13. Source |
| 6. Description | 14. Language |
| 7. Publisher | 15. Identifier |
| 8. Type | |

15 Dublin Core Elements

- **Creator:** An entity primarily responsible for making the content of the resource.
- **Title:** The name given to the resource.
- **Subject:** The topic of the content of the resource. Typically, a Subject will be expressed as keywords or key phrases.
- **Contributor:** An entity responsible for making contributions to the content of the resource. Example: a person, an organization.
- **Date:** A date associated with an event in the life cycle of the resource.
- **Description:** An account of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.
- **Publisher:** The entity responsible for making the resource available.
- **Type:** The nature or type of the content of the resource. Examples: Type="Image", "Sound", "Text", etc.

15 Dublin Core Elements

- **Format:** The physical or digital appearance of the resource. Typically, Format may include the media-type or dimensions of the resource. Examples: Format="image/gif/tif/jpg".
- **Coverage:** The scope of the content of the resource. Coverage will typically include location, period.
- **Rights:** Information about rights held in and over the resource. Rights information like Intellectual Property Rights (IPR), Copyright, and various Property Rights. Example: Rights="Access limited to members".
- **Relation:** A reference to a related resource.
- **Source:** A Reference to a resource from which the present resource is derived.
- **Language:** A language of the content.
- **Identifier:** An unambiguous reference to the resource within a given context.

3 Dublin Core Elements

Qualified Dublin Core includes **three additional elements**.

- **Audience:** A class of entity for whom the resource is intended or useful. A class of entity may be determined by the creator or the publisher or by a third party.

Examples:

Audience="elementary school students"

Audience="teachers"

Audience="deaf adults"

- **Provenance:** A statement of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity and interpretation. Example:

Provenance="Stolen in 1999; recovered by the Museum in 2003".

- **Rights Holder:** A person or organization owning or managing rights example:

Digital Skills for Developing Digital Libraries

Followings are the major skills required for developing digital libraries:

- Skills to select appropriate scanner & standards to digitize print document and develop digital contents
- Skills to use OCR software to convert image into digital document

- Skills to use digital library software for developing digital library
- Skills to assign metadata and related standards to develop searchable digital contents
- Skills to select appropriate indexing level in digital library software for developing digital contents

- Skills to use different storage devices to preserve digital contents (DVDs, CD-ROM, Hard-disk, library server etc.)
- Skills to tackle intellectual property rights/copy right of digital material and licensing issues
- Skills to design user-friendly interface of digital library according to your own searching preferences

Digital Skills for Managing Digital Libraries

Followings are the major skills required for managing digital libraries:

- Ability to manage over all digital library services & infrastructure
- Ability to plan cost for managing digital library
- Ability to define criteria for selecting new digital contents

- Ability to define policies and standards for digitizing print material
- Ability to define strategies to ensure the quality of digital through proper digitization of those contents
- Ability to cope all type of technical issues for managing digital infrastructure and software technology

- Ability to design system for obtaining digital library usage statistics and feedback from the users
- Ability to conduct digital library evaluation (local and external experts)
- Ability to plan long term funding to sustain digital library