

<b>DIGITAL LIBRARY: DEFINATION TO IMPLEMENTATION.....</b>	<b>1</b>
Lecture Delivered at Ranganathan Research Circle, 26 <sup>th</sup> July 2003, Delhi .....	1
By Sukhdev Singh, esukhdev@yahoo.com .....	1
<b>WHAT IS A DIGITAL LIBRARY: .....</b>	<b>1</b>
<i>Major Characteristics of Digital Library: .....</i>	<i>3</i>
<i>Digital Libraries Vs Conventional Libraries:.....</i>	<i>4</i>
<i>Digital Library Myths: .....</i>	<i>5</i>
Myth 1: The Internet is the digital library .....	5
Myth 2: The myth of a single digital library or one-window view of digital library collections.....	5
Myth 3: Digital libraries will provide more equitable access, anywhere, any time. ....	6
Myth 4: Digital libraries will be cheaper than print libraries. ....	6
<b>BUILDING A DIGITAL LIBRARY:.....</b>	<b>7</b>
CHALLENGES: .....	7
DIGITAL LIBRARY DESIGN:.....	8
DIGITAL LIBRARY STANDARDS:.....	11
COLLECTION MANAGEMENT: .....	12
<i>Digitization: .....</i>	<i>12</i>
<i>SCANNERS AND SCANNING.....</i>	<i>12</i>
<i>OCR: OPTICAL CHARACTER RECOGNITION.....</i>	<i>13</i>
INFORMATION RESOURCES ORGANISATION:.....	15
<i>Classical Wisdom: Classification and Cataloguing.....</i>	<i>15</i>
<i>Problems of Information Organisation in Digital Libraries: .....</i>	<i>15</i>
<i>Metadata: What is it? .....</i>	<i>16</i>
INFORMATION ACCESS AND USER INTERFACE: .....	17
DIGITAL ARCHIVING AND PRESERVATION:.....	17
DIGITAL LIBRARY SERVICES: .....	18
SOCIAL, ECONOMIC AND LEGAL ISSUE:.....	18

# DIGITAL LIBRARY: DEFINATION TO IMPLEMENTATION.

Lecture Delivered at Ranganathan Research Circle, 26<sup>th</sup> July 2003, Delhi.

By Sukhdev Singh, [esukhdev@yahoo.com](mailto:esukhdev@yahoo.com)

## WHAT IS A DIGITAL LIBRARY:

What is a Digital Library? There are many buzz-words for it, sometimes even referring to remotely related activities but not limited to: multimedia database, information mining, information warehouse, information retrieval, on-line information repositories, electronic library, operational image applications, imaging, World Wide Web (WWW) and Wide Area Information Services (WAIS) (Gladney, HM *et. al*, 1994)<sup>1</sup>. There are too many topics with much overlap of related activities. Then what is Digital Library and how it can be distinguished from related activities. The answer is not straightforward. The concept of Digital Libraries is evolving over time. Moreover different communities are active in the area of Digital Libraries with competing visions. **Borgman, CL** (1999)<sup>2</sup> point out the term digital library is used in at least in two senses:

- i. In the computer science research community digital libraries are viewed as *content collected on behalf of users*.
- ii. In the library practitioner community digital libraries are seen as *institutions providing a range of services in a digital environment*.

While most of digital library projects falls into the first category, the speculation about the future developments concentrates on versions of the second.

The NSF/ARPA/NASA *Digital Library Initiative*, FY 1994 states<sup>3</sup>:

*“Information sources accessed via the Internet are ingredients of a digital library. Today, the network connects some information sources that are a mixture of publicly available (with or without charge) information and private information shared by collaborators. They include reference volumes, books, journals, newspapers,*

---

<sup>1</sup> Gladney H.M, et. al. (1994) Digital library: Gross structure and requirements: Report from a workshop. *IBM Research Report, RJ 9840, May 1994*.

<sup>2</sup> Borgman C.L (1999) What are digital libraries? Competing visions. *Information Processing and Management*, 35 (3), 227-43.

<sup>3</sup> Digital Library Initiative, FY 1994 (1993) A Joint Initiative of the National Science Foundation, the Advanced Research Projects Agency, and the National Aeronautics and Space Administration. *U.S. Government document NSF 93-141*.

*national phone directories, sound and voice recordings, images, video clips, scientific data (raw data streams from instruments and processed information), and private information services such as stock market reports and private newsletters. These information sources, when connected electronically through a network, represent important components of an emerging, universally accessible, digital library.”*

This definition asserts inclusion of all network accessible information in the digital libraries. Moreover digital libraries are emerging and universally accessible. However this definition does not tell components of the Digital Library.

According to **Gladney H.M, et. al.** (1994)<sup>4</sup>

*“A digital library service is an assemblage of digital computing, storage, and communications machinery together with the software needed to reproduce, emulate, and extend the services provided by conventional libraries based on paper and other material means of collecting, storing, cataloguing, finding, and disseminating information.”*

According to them a digital library is a machine-readable representation of materials, which might be found in conventional library. Along with this representation, organising information is also available to assist users in finding specific information.

**Association of Research Libraries** (1995)<sup>5</sup> has identified the following five elements in various definitions of the digital libraries:

1. *The digital library is not a single entity;*
2. *The digital library requires technology to link the resources of many*
3. *The linkages between the many digital libraries and information services are transparent to the end users;*
4. *Universal access to digital libraries and information services is a goal;*
5. *Digital library collections are not limited to document surrogates: they extend to digital artefacts that cannot be represented or distributed in printed formats.*

This definition introduces the concept of distributed and linked resources. The digital resources are the collections and information services. The digital collections are – digital surrogates, Non-printable objects and digital artefacts that cannot be distributed in print form.

---

<sup>4</sup> Gladney H.M, et. al. (1994) Digital library: Gross structure and requirements: Report from a workshop. *IBM Research Report, RJ 9840, May 1994.*

<sup>5</sup> Association of Research Libraries (1995) Definition and Purpose of a Digital Library. Available: <http://www.ifla.org/documents/libraries/net/arl-dlib.txt> [2003 June 24].

Paul Duguid (1997)<sup>6</sup> has defined the Digital Library as an environment to bring together in support of life cycle of information in addition to digital collection and information management tools.

*The concept of a "digital library" is not merely equivalent to a digitized collection with information management tools. It is rather an ENVIRONMENT to bring together collections, services, and people in support of the full life cycle of creation, dissemination, use, and preservation of data, information, and knowledge. (Duguid, Paul, 1997).*

More recent definition is of Deegan, M and Tanner, S (2002)<sup>7</sup>. They has given a set of defining principles that seem unarguable:

- *A digital library is a managed collection of digital objects.*
- *The digital objects are created or collected according to the principles of collection development.*
- *The digital objects are made available in a cohesive manner, supported by services necessary to allow users to retrieve and exploit the resources just as they would any other library materials.*
- *The digital objects are treated as long term stable resources and appropriate processes are applied to them to ensure their quality and survivability.*

### **Major Characteristics of Digital Library:**

Based on Chowdhury GG and Chowdhury S (2003)<sup>8</sup> following major characteristics have been jotted down:

- Variety of digital information resources
- Digital Libraries Reduce the need for physical space
- Users at remote
- Users may build their own personal collections by the facilities provided by Digital Library
- Provide access to distributed information resources
- Same information resource can be shared by many at the same time
- Paradigm shift both in use and ownership
- Collection development be based on potential usefulness and appropriate filtering mechanisms be followed to negotiate the problem of plenty
- Ability to handle multilingual content

<sup>6</sup> Duguid, Paul (1997) *Report of the Santa Fe Planning Workshop on Distributed Knowledge Work Environments*, [Online], Available: <http://www.si.umich.edu/SantaFe/>. [2003 June 24].

<sup>7</sup> Deegan, M and Tanner S. (2002) *Digital Futures: Strategies for the information age*, London: Library Association Publishing, 22.

<sup>8</sup> Chowdhury GG and Chowdhury S (2003) *Introduction to Digital Libraries*. London: Facet Publishing, 8-9.

- Presupposes the absence of human intermediaries
- Should provide better searching and retrieval facilities
- Digital information can be used and viewed differently by different people
- Digital Library breaks the time, space and language barrier

### **Digital Libraries Vs Conventional Libraries:**

Are Digital Libraries different from conventional libraries? If so, then why the word “Libraries” was tacked onto “ Digital” at all. On the other hand, if they are the same then why we need the term “Digital” at all. In fact there are people like R M Braud who feel, that using the term “digital” is a redundancy. *“The product that we manage in libraries, information, and the familiar container for that product, the codex book. These containers have influenced library architecture, but they do not themselves define what a library is. We do not bother to qualify our libraries by calling them Clay Libraries or Papyrus Role Libraries, why do we have to call the digital libraries”* (Braund, R M, 1999)<sup>9</sup>. However there is a distinction to be made between conventional libraries and digital libraries. Physical containers for information (for example books) are capable of direct access and can be managed physically. On the other hand, Digital data is made of electronic signals that rely on an interpreting machine before there can be any human interaction with it.

Let us re-look at the basic functions of conventional libraries.

- i. **Collection:**  
Includes techniques for evaluation of information resources directed towards target users. Cost-effective storage and preservation of such resources.
- ii. **Organisation and Representation:**  
Classification and cataloguing of information resources which is relevant to the potential users.
- iii. **Access and Retrieval:**  
Access considerations include design of physical space and organisation of materials within such space to respond effectively to user needs and expectations. Information retrieval has been addressed in the design of systems specific to that task.
- iv. **Analysis, synthesis and dissemination functions:**  
Circulation and value added service like reference services, producing evaluation reviews and devising communality out-reach programs.

Librarians and Information Scientists have developed techniques, procedures and systems for addressing each of these functions for many kinds of data and presentation. Digital Libraries are unlikely to omit or add any of the roles played by conventional libraries. Although the implementations would depend on the

---

<sup>9</sup> Braud, R.M (1999) Virtual or actual: the term library is enough, Bulletin of the Medical Librarians Association, 87(1): 85-7

local context and technologies chosen, there must be with in commonly accepted constrains which must satisfy the users.

Digital Libraries research has been focused on automating the activities carried out by librarians, such as automatic indexing and classification and expert systems for reference desks. Digital catalogues can support long keywords along with deferential weights, long user queries, ranked retrieval etc. Information search via hypertext illustrates that indices can be implicit rather than explicit, giving users a seamless blend of primary and secondary works. Further, some current library activities may become irrelevant. For example, circulation problems originating in a fixed number of copies of each work simply disappears. We might redefine and redesign library services to achieve the basic aims more effectively than is possible now. Thus digital libraries not only involves automation of each traditional library activity and service, but also calls for redefinition of services, new groupings of services or replacements of groups of services with other solutions.

### **Digital Library Myths:**

There is lot of expectations and confusions about the Digital Libraries. Terry Kuny and Gary Cleveland (1998)<sup>10</sup> have tried to explode some common myths associated with Digital Libraries.

#### **Myth 1: The Internet is the digital library.**

Many different groups to signify simply a collection of digital objects that people can access from their desktops have appropriated the word "library". A global information network, of which the Internet is the seed, has the illusion of promising fingertip access to the world's information. A fairly spectacular example of what many people consider to be a digital library today is the World Wide Web. But is this a "digital library"? For many common library requests, locating information on the Internet remains highly inefficient compared to traditional library sources, especially for unfamiliar users. Finding information is difficult, the quality of the information is quite variable, and reliable, professional assistance for the confused and lost is lacking.

#### **Myth 2: The myth of a single digital library or one-window view of digital library collections.**

Even modest moves towards increasing digital collections and services will be strongly affected by future copyright and licensing regimes, as well as prohibitive costs for digitization and support of technical infrastructure. But more importantly, the digital future will be an unruly one composed of multiplicity of competing information providers. Libraries will be only one source of information. "Prime" information resources will probably be locked into proprietary Collections essentially "private digital libraries" which are accessible on a subscription or pay-per use basis. Developing interoperability standards for locating and retrieving information in this

---

<sup>10</sup> Kuny, Terry & Cleveland, Gary (1998) The Digital Library: Myths and Challenges. *IFLA Journal*, 24(2), pp 107-113, Available: <http://www.ifla.org/IV/ifla62/62-kuny.pdf> [2003 June 24].

highly distributed and heterogeneous environment will be a considerable challenge in their own right.

**Myth 3: Digital libraries will provide more equitable access, anywhere, any time.**

A great deal of work must be done to turn this myth into reality. We can assume that a global computer network or the Internet or some descendant will be the primary delivery mechanism for digital information. Equitable access is currently compromised by the fact that the Internet is not available to every one equally. Furthermore, the connections that do exist for most people are slow. For a digital library to provide equitable access to information, it is imperative that the same universal availability that is a characteristic of the telephone system is also a characteristic of the network. In the future, complex multimedia resources and services may have specialized hardware and software requirements such that only a limited number of workstations can actually access the information.

**Myth 4: Digital libraries will be cheaper than print libraries.**

A common assumption among technology reporters about the costs of "digital libraries" is that digital is cheaper than paper. This contention is far from established in fact or in practice. Although many libraries project savings, especially when substitution strategies are used which replace selected serials titles with document delivery services, the cost/benefit analysis of making this switch remains unclear. Many libraries now devote significant resources for hardware and software infrastructure.

There is an increasing unease among members of the library community that copyright changes will adversely affect the ability of libraries to provide digital collections and services. The discomfort librarians feel is justified. One has only to consider the statement of the International Publishers Copyright Council<sup>11</sup>. On digital library collections to sense the challenge that librarians face:

*"Many national and regional libraries contemplate digitizing their print collections to facilitate a virtual library that can provide service to patrons at remote locations and facilitate resource-sharing. Such a concept will destroy not only the incentive to create new copyrighted works, but the revenue from existing works that provides the investment in new works by authors and publishers.....No longer will libraries be the sole repository of published matter. No longer will libraries be the only means of obtaining archival information. In some areas, libraries will be able to fulfill their function by merely pointing to other electronic repositories and in others they will seek out more active roles."*(International Publishers Copyright Council, 1996).

It is no surprise Digital Content providers are resorting to Contract Agreements and Licensing Mechanisms instead of normal copyright provisions.

---

<sup>11</sup> International Publishers Copyright Council (1996). *IPCC Statement: Libraries, Copyright and the Electronic Environment*. April 22, 1996.

# BUILDING A DIGITAL LIBRARY:

## Challenges:

The staff of the National Digital Library Program at the Library of Congress (2003)<sup>12</sup> have identified ten challenges that must be met if large and effective digital libraries are to be created during the 21st century.

### Building the Resource

*Challenge One:* Develop improved technology for digitizing analog materials.

*Challenge Two:* Design search and retrieval tools that compensate for abbreviated or incomplete cataloging or descriptive information.

*Challenge Three:* Design tools that facilitate the enhancement of cataloging or descriptive information by incorporating the contributions of users.

### Interoperability

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

### Intellectual Property

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

### Effective Access

*Challenge Six:* Integrate access to both digital and physical materials.

*Challenge Seven:* Develop approaches that can present heterogeneous resources in a coherent way.

*Challenge Eight:* Make the National Digital Library useful to different communities of users and for different purposes.

---

<sup>12</sup> Library of Congress. National Digital Library Program (2003). Challenges to Building an Effective Digital Library. [Online]. Available at: <http://memory.loc.gov/ammem/dli2/html/cbedl.html> [24 Jul. 03].



**Challenge Nine:** Provide more efficient and more flexible tools for transforming digital content to suit the needs of end-users.

## **Sustaining the Resource**

**Challenge Ten:** Develop economic models for the support of the National Digital Library.

## **Digital Library Design:**

Design of a digital library should facilitate variety of information resources residing on variety of computer systems in different parts of the world to a number of users of differing notions and needs. It should be One-Stop-Shop.

Design Principles:

Kahn and Wilensky (1995)<sup>13</sup> describe the basic infrastructure of a digital library which is “open in its architecture and which supports a large and extensive class of distributed digital information services”. This architecture has been the subject of a series of useful discussion from which the following eight general principles have emerged (Arms, 1995)<sup>14</sup>. These may be considered as the guiding principles for the design of digital libraries.

1. The technical framework exists within a legal and social framework.
2. Understanding of digital library concepts needs standard terminology.
3. The underlying architecture should be separate from the content stored in the library.
4. Names and identifiers are the basic building blocks for the digital library.
5. Digital library are more than collections of bits.
6. A digital library object should not be tied to a particular technology.
7. Repositories must look after the information they hold.
8. Users want intellectual works, not digital objects.

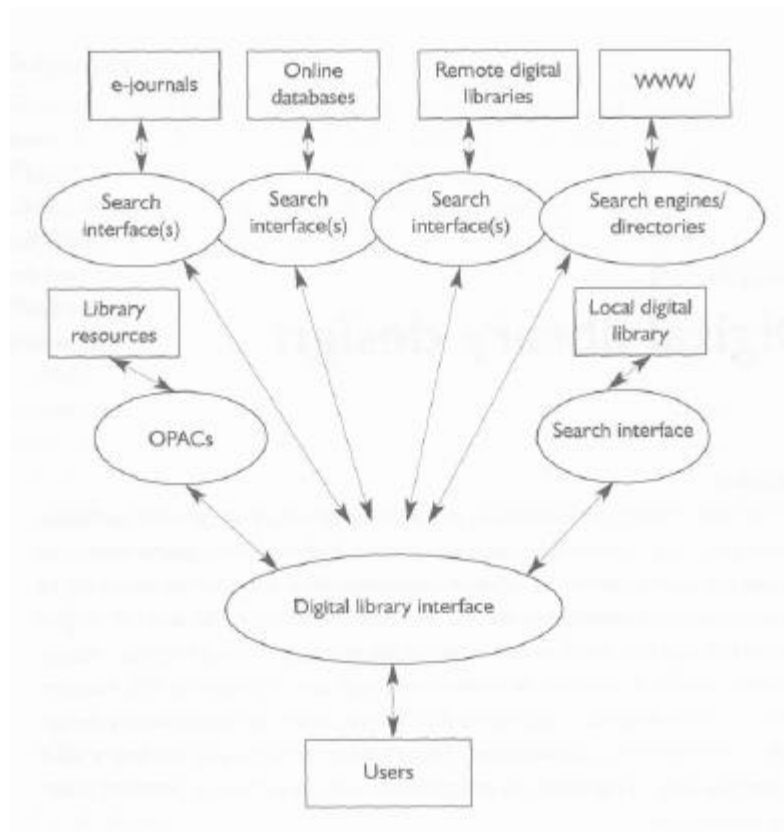
Conceptual design of a digital library:

The figure ahead gives a conceptual design of digital library.

---

<sup>13</sup> Kahn, R and Wilensky, R (1995) A framework for distributed digital object services. Available at <http://www.cnri.reston.va.us/home/cstr/arch/k-w.html>

<sup>14</sup> Arm, W.Y (1995) Key Concepts in the architecture of the digital library. D-Lib Magazine, 1(1), July.  
Available at <http://www.dlib.org/dlib/july95/07arms.html>



Conceptual design of a digital library

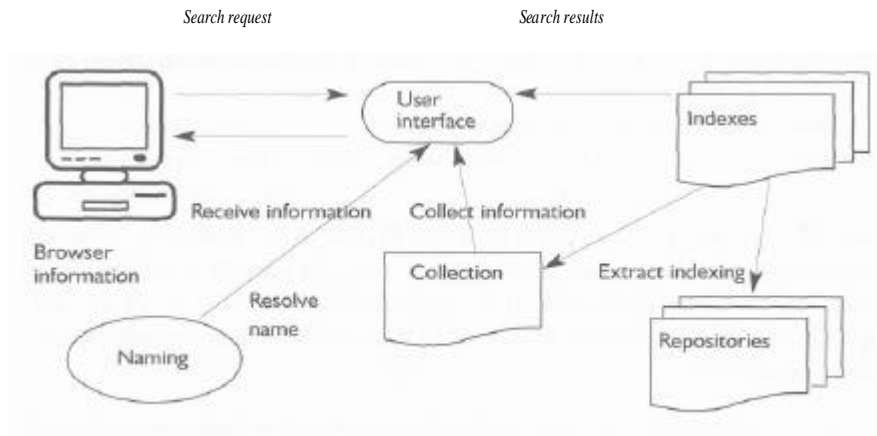
### Example of CRADDL

Cornell Reference Architecture for Distributed Digital Libraries (CRADDL) is component- or service-based digital library architecture. It defines the following five basic services (Lagoze and Fielding, 1998)<sup>15</sup>:

- the repository service, which provides the mechanisms for the deposit and storage of, and access to, digital objects
- the naming service, which identifies digital objects by unique names, URNs, then registered with the naming service
- the index service, which provides the mechanism for discovery of digital objects via query
- the collection service, which provides the mechanisms for the aggregation of access to sets of digital objects and services into meaningful collections
- user interface services or gateways, which provide human-centred entry points to the functionality of the digital library.

<sup>15</sup> Lagoze, C. and Fielding, D (1998) Defining collections in distributed digital libraries, D-Lib Magazine, November, available at <http://www.dlib.org/dlib/november98/lagoze/11lagoze.html>

Of these five services, only the user interface is accessed directly by a human. The design of the user interface gateway can be customized for a specific community using mechanisms such as language, help facilities and graphical aids. The user interface provides users with access to one or more collections through interaction with the collection services corresponding to those collections. The following figure provides an outline of the CRADDL design.



Interaction among the digital library services in CRADDL

## **Digital Library Standards:**

There is a great need for adopting various standards and best practices to build interoperable digital libraries. Dempsey *et al* (1998)<sup>16</sup> propose the following standards for use in eLib Projects:

### **User Interface:**

Common Web Browser

### **Data Handling and Interchange:**

Graphic Formats – JPEG, TIFF, GIF, PNG, Group 4 Fax, CGM

Structured Documents – SGML, HTML, XML

Moving Pictures/3-D – MPEG, AVI, GIF89A, QuickTime, Real Video, ViviActive, VRML

### **Metadata:**

Resource Description – Dublin Core, WHOIS++ Templates, US-MARC, TEI Headers, Other Open Source and Domain Specific Standards.

Resource Identification – URN, PURL, DOI, SICI

### **Security, Authentication and payment services:**

Emerging e-Commerce Standards.

---

<sup>16</sup> Dempsey L. *et al* (1998) eLib Standards guidelines version 2.0. Available at : <http://www.ukoln.ac.uk/services/elib/papers/other/standards/version2/>

## **Collection Management:**

### **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 content 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 eJournals and eBooks are also sources of discovery of digital content.

### **Selection Criteria:**

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

- Content, Quality, Currency etc.
- Hardware, Software and Network requirements.
- Version of the product, Network or Stand Alone especially in case CD-ROMs.
- Number of concurrent users allowed.
- Access control – Passwords, Proxy Server Authorization.
- Price and licensing and copyright agreements.
- Database features, the retrieval engine and the user interface(s).
- Archiving procedures.

### **Digitization:**

Digital documents may be born-digital, created using digital publishing tools (e.g. Word, LaTeX, DTP), 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' or 'digitalization'.

### **SCANNERS AND SCANNING**

The first step in converting paper documents into a digital library collection is to obtain images of all pages of all 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 from manufacturers such as Bell & Howell. The output format of a scanned page is a computer file that is usually stored in TIFF or Bitmap format. Compressed TIFF IV is the best format to use. An average page scanned and converted to this format occupies only 50 Kb, compared to perhaps 2 Mb for the equivalent page in uncompressed Bitmap form.

- Low-cost flat-bed scanner: Low-cost flat-bed units are the cheapest and most widely available type of scanner. There are many brands: HP, Agfa, Acer, etc. Prices range from 5,000 to 15,000. Both black-and-white and color images can be scanned.
- Low-end scanner with sheet feeder: Low-end scanners with sheet feeders typically cost between Rs. 25,000 and Rs. 60,000.
- Professional duplex scanners: Professional scanners are reliable, heavy-duty machines capable of processing a large volume of pages—typically from 2000 pages to 10,000 pages per day. Prices range from Rs. 2,50,000 to Rs. 25,00,000. For example, the Canon DR-6020 duplex scanner costs Rs.2,50,000 and works with double-sided documents.

*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.

Preparing the documents

*The 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:

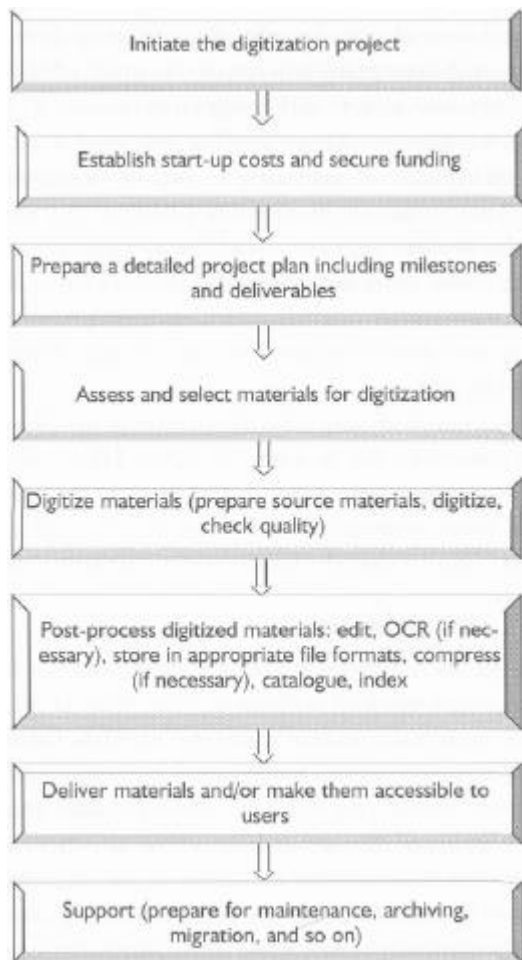
1. scanning;
2. page layout analysis;
3. recognition;
4. scanning images and tables.

On the market are many good OCR programs, with prices ranging from Rs. 5000 to Rs.20,000.

For example, among many others are:

- Read-Iris (<http://www.readiris.com/>)
- Omnipage (<http://www.omnipage.com/>)
- Fine-Reader (<http://www.finereader.com/>)

The process of digitization is given in the diagram shown ahead.



## Information Resources Organisation:

Librarians have used a number of tools for organising their information resources, including classification schemes, catalogue codes, bibliographic formats and vocabulary control tools. Many researchers have also used these tools for organising web information resources. However they noted that traditional cataloguing tools were inadequate for the creation of metadata for digital information resources to support resource discovery and information access. Hence new metadata standards had to be developed.

### Classical Wisdom: Classification and Cataloguing

Classification is the process of bringing related items together. Conventional libraries to stack books on related subjects together have used modern library classification. This facilitated the browsing approaching of the information seekers and the learners. Cataloguing created document surrogates. Which provided searching facility by Authors, Titles, Series and other elements. Since classification used notation symbols, Subject cataloguing / indexing was done to provide both browsing and searching by subjects.

### Problems of Information Organisation in Digital Libraries:

Should the same classical wisdom be applied to Digital Libraries? People are debating it. Digital libraries provide access to different types of information resources, which may be organised in varying ways by their producers or access providers. It is a challenge for a given digital library to build and /or adopt a simple and yet effective method for organising information. Some researchers argue that the traditional classification schemes may not be required in digital libraries (Jones, 2002<sup>17</sup>), though others don't agree (E. Hunter, 2002<sup>18</sup>; Pollitt, 2002<sup>19</sup>).

Traditional classification schemes are being used for organising web information resources according to the disciplines and specific subjects and topics within a discipline. Users find it comparatively easy to get information from such organised resources structures. For example:

BUBL LINK. <http://www.bubl.ac.uk/link>. DDC & LCSH  
CyberDewey. <http://www.anthus.com/cyberdewey/cyberdewey.html>  
SCORPION. <http://orc.rsch.oclc.org:6109/>.  
CyberStacks. <http://www.public.iastate.edu/~cyberstack/>. LC.  
INFOMINE. <http://infomine.ucr.edu>. LCSH.  
BIOME. <http://www.biome.ac.uk>.

---

<sup>17</sup> Jones, D (2002) Classification: are we overdoing it? Update, 1(1), 24.

<sup>18</sup> Hunter, E (2002) Classification: are we underdoing it? [M@il](#), Update, 1(2), 33

<sup>19</sup> Pollitt, A (2002) Still the name of the game. [M@il](#), Update, 1(2), 33.



However, digital libraries deal with many new and nascent subjects, Classification schemes are unable to update to that pace. Moreover it is expensive to classify and catalogue according to traditional schemes by human experts. Various new metadata schemes are being developed to organise digital resources.

### **Metadata: What is it?**

What is metadata after all? The term entered our vocabulary with rapid growth from around 1994. It is nothing else but DATA ABOUT DATA. Technically speaking it has been defined as “*Metadata are structured, encoded data that describe characteristics of information bearing entities to aid in the description, discovery, assessment and management of the described entities*” by the Association for Library Collection and Technical Services (ALCTS) Task Force on Metadata<sup>20</sup>. Vellucci (1998)<sup>21</sup> maintains that both Computer Science and Library and Information Community have adopted the term “metadata” for describing electronic data. However Michael Gorman (1999)<sup>22</sup> argues that the non-cataloguing world perceives metadata as being different from the traditional cataloguing, which according to him, have complex formats and expansive and stringent quality requirements. He believes that four approaches are available for the bibliographic control of electronic resources of varying control. a. Full MARC cataloguing for high quality resources that are likely to have continuing value. b. Enriched Dublin Core for next level. c. Minimal Dublin Core for Next level. d. Unstructured full-text keyword searching for the remainder.

Metadata is being used for Resource Identification and Rights Management. Resource Description and Discovery as well as Preservation of Digital Resources. The use of Metadata Standards facilitates interoperability.

---

<sup>20</sup> Association for Library Collection and Technical Services. Task Force on Metadata (2002). Final Report, June 2000. Available at <http://www.ala.org/alcts/organization/ccs/ccdata/tf-meta6.html>.

<sup>21</sup> Vellucci, S.L (1998) Metadata. In Williams, M.E. (ed.) Annual review of information science and technology (ARIST), Vol 33, Medford, NJ, Information Today Inc. on behalf of ASIS, 187-222.

<sup>22</sup> Gorman, M (1999) Metadata or Cataloguing? A false choice, Journal of Internet Cataloguing, 2(1), 5-22.

## **Information Access and User Interface:**

User Interface is very important aspect of Digital Libraries. It is because digital libraries are support and serve remote users globally. Since digital do not have any physical presence, users do not get a view of the collection and contents as they do in a printed library. Determining how much information to show is a major design choice in information access interfaces (Hearst, 1999)<sup>23</sup>.

Shneiderman (1999)<sup>24</sup> proposes eight areas of attention from the researchers to make future digital libraries useful for creative work:

- Searching and browsing digital libraries.
- Consulting with peers.
- Visualizing data and processes.
- Thinking by free association.
- Thinking by free association.
- Exploring solutions.
- Composing artefacts and performances.
- Reviewing and replaying session histories.
- Disseminating results.

## **Digital Archiving and Preservation:**

Digital information needs to be achieved and preserved for future use. Digital information may be lost for many reasons, for example:

- Changes in an organisation
- Content reorganisation
- Cession of sponsorship
- Technology obsolescence
- Content format obsolescence
- Hacking and sabotage
- Disaster, whether nature or man-made

The Issues involved in archiving and preservation of digital information are:

- Preservation Strategies
- Data creators and publishers
- Intellectual property rights
- Collection management
- Metadata
- Web archiving
- Staff expertise
- Collaboration

---

<sup>23</sup> Hearst, M (1999) User Interfaces and Visualization. In Baeza-Yates, R. and Ribeiro-Neto, B (eds), Modern Information Retrieval, New York, ACM Press, 257-323.

<sup>24</sup> Shneiderman, B. (1999) User interfaces for creativity support tools. In Proceedings of the Third Conference on Creativity and Cognition, Loughborough, New York, ACM Press, 15-22.

## **Digital Library Services:**

Reference service and information services are the integral to traditional services. Although digital library research and development so far has not concentrated on them, reference services should form an important part of digital libraries. A number of free and fee based reference and information services are now available through web, and many of them are offered by non-library organisations.

- Ask Jeeves. <http://www.AskJeeves.co.uk>.
- Ask A Librarian. <http://www.earl.org.uk/ask>.
- CDRS. <http://www.loc.gov/rr/digiref>.
- Automatic Reference Librarians for the World Wide Web.  
<http://www.fastlane.nsf.gov/servlet/showaward?award=9874759>.
- SIFTER. <http://sifter.indiana.edu>.
- The Virtual Reference Desk. <http://www.vrd.org>

## **Social, Economic and Legal Issue:**

Let us discuss.

Thank you for listening.

---

The opinion expressed here are of the author, the employer of the author may not endorse any of these. It is under consideration for publication in refined form.