UNIT - IV

Metadata standards

Standard meaning:

A standard is a document that provides requirements, specifications and guidelines or characteristics that can use consistency to ensure that materials, products, processes and services are fit for their purpose.

Uses of existing standards:

- Cost serving
- Access to help and advice
- Usability
- Resource discoverability
- Sustainability

Metadata:

o Data about data
o A set of data that describes and gives information about other data.
o "Metadata is structured information that describes, explains, locates or otherwise make it easier to retrieve, use and manage an information resource".
o Metadata describes how and when and by whom a particular set of data was collected, and how the data is formatted.
o Metadata is defined as the data providing information about one or more aspects of the data; it is used to summarize basic information about data which can make tracking and working with specific data easier.
o For example, Author, Date created and date modified and file size are examples of very basic document metadata.

Functions of metadata:

- Metadata serves following functions :

Resource discovery

- Allowing resources to be found by relevant criteria;
- Identifying resources;
- Bringing similar resources together;
- Distinguishing dissimilar resources;
- Giving location information.

Organizing e-resources

- Organizing links to resources based on audience or topic.

- Building these pages dynamically from metadata stored in databases.

Facilitating interoperability

Using defined metadata schemes, shared transfer protocols, and crosswalks between schemes, resources across the network can be searched more seamlessly.

  • Cross system search, e.g. using z39.50 protocol

  • Metadata harvesting e.g. using OAI protocol

Digital identification

- Elements for standard number, e.g. ISBN.
- The location of digital object may also given using :

  • A file name
  • A URL

Archiving and preservation
Challenges :
  • Digital information is fragile and can be corrupted or altered;
  • It may become unusable as storage technologies change.
Metadata is key to ensuring that resources will survive and continue to be accessible into the future.
Archiving and preservation require special elements:
  • to track the lineage of a digital object,

  • to detail its physical characteristics, and

  • to document its behaviour in order to emulate it in future technologies.

Need of metadata:
Metadata is created and collected because it enables and improves use of that data. Metadata ensures that we will be able find data, use data, and preserve and re-use data in the future.

- To enable discovery of your digitised material
- To help you organise your digitised material
- To support archiving and preservation

Purposes of Metadata

Types of metadata:

NISO distinguishes among three types of metadata:

Descriptive metadata, Structural metadata, and Administrative metadata.

Descriptive metadata is typically used for discovery and identification, as information to search and locate an object, such as title, author, subjects, keywords, publisher.

Metadata used to identify and describe collections and related information resources.

Examples:

- Cataloguing records
- Finding aids
- Differentiations between versions
- Hyperlinked relationships between resources

Structural metadata describes how the components of an object are organized. An example of structural metadata would be how pages are ordered to form chapters of a book.

Administrative metadata gives information to help manage the source. Administrative metadata refers to the technical information, including file type, or when and how the file was created and who can access it.

Metadata used in managing and administering collections and information resources.

Examples:

- Acquisition information
- Rights and reproduction tracking
- Documentation of legal access requirements
- Location information

Many distinct types of metadata exist, including descriptive metadata, structural metadata, administrative metadata, reference metadata and statistical metadata.

There are several subsets of administrative data; two that are sometimes listed as separate metadata types are:

- Rights management metadata, which deals with intellectual property rights,
- Preservation metadata, which contains information needed to archive and preserve a resource.

International Standard Bibliographic Description (ISBD)

- The International Standard Bibliographic Description (ISBD) is a set of rules produced by the International Federation of Library Associations and Institutions (IFLA) to create a bibliographic description in a standard, human-readable form, especially for use in a bibliography or a library catalog.
- A preliminary consolidated edition of the ISBD was published in 2007 and the consolidated edition was published in 2011, superseding earlier separate ISBDs for monographs, older monographic publications, cartographic materials, serials and other continuing resources, electronic resources, non-book materials, and printed music.
- IFLA's ISBD Review Group is responsible for maintaining the ISBD.
- One of the original purposes of the ISBD was to provide a standard form of bibliographic description that could be used to exchange records internationally. This would support IFLA's Universal Bibliographic Control program.

Structure of an ISBD record

The ISBD defines nine areas of description. Each area, except area 7, is composed of multiple elements with structured classifications.

Elements and areas that do not apply to a particular resource are omitted from the description. Standardized punctuation (colons, semicolons, slashes, dashes, commas, and periods) is used to identify and separate the elements and areas.

The order of elements and standardized punctuation make it easier to interpret bibliographic records when one does not understand the language of the description.

0: Content form and media type area

1: Title and statement of responsibility area, consisting of

◦1.1 Title proper

◦1.2 Parallel title

2: Edition area

3: Material or type of resource specific area (e.g., the scale of a map or the numbering of a periodical)

4: Publication, production, distribution, etc., area

5: Material description area (e.g., number of pages in a book or number of CDs issued as a unit)

6: Series area

7: Notes area

8: Resource identifier and terms of availability area (e.g., ISBN, ISSN)

Example:

A typical ISBD record looks like this:

Text : unmediated

A manual for writers of research papers, theses, and dissertations : Chicago style for students and researchers / Kate L. Turabian ; revised by Wayne C. Booth, Gregory G. Colomb, Joseph M. Williams, and University of Chicago Press editorial staff. — 7th ed. — Chicago : University of Chicago Press, 2007. — xviii, 466 p. : ill. ; 23 cm — (Chicago guides to writing, editing, and publishing). — Includes bibliographical references (p. 409-435) and index. — ISBN 978-0-226-82336-2 (cloth : alk. paper) : USD35.00. — ISBN 978-0-226-82337-9 (pbk. : alk. paper) : USD17.00

Dublin core (DC)

• The Dublin Core Schema is a small set of vocabulary terms that can be used to describe web resources (video, images, web pages,etc.), as well as physical resources such as books or CDs, and objects like art works.

• Dublin Core is an initiative to create a digital "library card catalogue" for the Web.

• Dublin Core is made up of 15 metadata (data that describes data) elements that offer expanded cataloguing information and improved document indexing for search engine programs.

• The Dublin Core Metadata Element Set is a general-purpose scheme for resource description originally intended to facilitate discovery of information objects on the Web.

• The development of official specifications related to the Dublin Core is managed by the Dublin Core Metadata Initiative (DCMI), which consists of a small, paid directorate advised by a board of trustees, and a large number of loosely organized volunteers.

• Dublin Core Metadata may be used for multiple purposes, from simple resource description, to combining metadata vocabularies of different metadata standards, to providing interoperability for metadata vocabularies in the Linked Data cloud and Semantic Web implementations.

• The Dublin Core Metadata Element Set is a standard for cross domain resource description.

Background

• "Dublin" refers to Dublin, Ohio, USA where the schema originated during the 1995 invitational OCLC/NCSA Metadata Workshop, hosted by the Online Computer Library Centre (OCLC), a library consortium based in Dublin, and the National Centre for Supercomputing Applications (NCSA).

• "Core" refers to the metadata terms as "broad and generic being usable for describing a wide range of resources".

• The semantics of Dublin Core were established and are maintained by an international, cross-disciplinary group of professionals from librarianship, computer science, text encoding, museums, and other related fields of scholarship and practice.

• DCMI(Dublin Core Metadata Initiative) traces its roots to Chicago at the 2nd International World Wide Web Conference, October 1994. This discussion on semantics and the Web revolved around the difficulty of finding resources.

• **Dublin Core (DC) I**

    In October 1994, informal discussion at second WWW Conference, Chicago : Identified 13 core elements (properties) essential for resource discovery on the Internet . 13 elements were :

1. Time 2. Author 3. Subject 4. Publisher 5. Date 6. Object 7. Form 8. Identifier 9. Relation 10. Source 11. Coverage 12.Right Management 13. Other Agent

DC II : April 1996, Warwick England

Established the Warwick Framework

Establish a syntax for the DC…HTML tags for embedding in web documents (HTML 2.0)

DC III : September, 1996, Dublin, Ohio

DC elements extended to include digital images, 15 properties .

DC IV : March 1997, Canberra, Australia

 Minimalists and structuralists, Canberrra qualifiers (meta-meta)

SchemesISO 8601 for date encoding (syntax encoding)

LCSH or AAT for subject metadata (vocabulary encoding)

DC-5: 1997, Helsinki, Finland

DC-6: 1998, Washington,

D.C. DC-7: 1999, Frankfurt, Germany

DC-8: 2000, Ottawa Canada

Dublin Core Metadata Elements(DCME)

A set 18 elements designed to enhance discovery and retrieval of resources.

Goals of DCME

- Simplicity of creation and maintenance
- Commonly understood semantics
- Conformance to existing and emerging standards
- International scope and applicability
- Extensibility
- Interoperability among collections and indexing system.

Why use Dublin Core ?

 "The scope of Dublin core is specially designed to provide a metadata vocabulary of core properties able to provide basic description about any kind of resources… regardless of any format of media specialization or cultural origin. It is important that a semantic model used for resource discovery is not dependant on the medium of the source it means to describe…"

 The Dublin Core metadata vocabulary is the result of many years of collaborative research to determine a common set of properties universal for describing any type of resource. The use of a standardized general classifications system also enables metadata of such collections to be combined and for knowledge contained within each collection to be shared.

Levels of Standards

• The Dublin Core standard originally includes two levels: Simple and Qualified.

- Simple Dublin Core comprised 15 elements i.e. Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format Identifier, Source ,Language, Relation ,Coverage ,Right.
- Qualified Dublin Core included 3 additional elements i.e. Audience, Provenance and Rights Holder.

Elements of Dublin Core

1. Identifier: Title

Definition: A name given to the resource.

2. Identifier: Creator

Definition: An entity primarily responsible for making the content of the resource.

3. Identifier: Subject

Definition: The topic of the content of the resource.

4. Identifier: Description

Definition: An account of the content of the resource.

5. Identifier: Publisher

Definition: An entity responsible for making the resource available.

6. Identifier: Contributor

Definition: An entity responsible for making contributions to the content of the resource.

7. Identifier: Date

Definition: A date associated with an event in the life cycle of the resource.

8. Identifier: Type

Definition: The nature or genre of the content of the resource.

9. Identifier: Format

Definition: The physical or digital manifestation of the resource.

10. Identifier: Identifier

Definition: An unambiguous reference to the resource within a given context.

11. Identifier: Source

Definition: A reference to a resource from which the present resource is derived.

12. Identifier: Language

Definition: A language of the intellectual content of the resource.

13. Identifier: Relation

Definition: A reference to a related resource.

14. Identifier: Coverage

Definition: The extent or scope of the content of the resource.

15. Identifier: Rights

Definition: Information about rights held in and over the resource.

16. Identifier : Audience

Definition : A class of entity for whom the the resource is intended or useful.

17. Identifier: Provenance

Definition : A statement of any change in ownership and custody of the resource since its creation that are significant for its authenticity, integrity and interpretation.

18. Identifier : Right Holder

Definition : A person or organization owing and managing rights over the resource.

Machine-readable cataloguing (MARC21)

MARC (Machine-readable cataloguing) standards are a set of digital formats for the description of items catalogued by libraries, such as books.

Working with the Library of Congress, American computer scientist Henriette Avram developed MARC in the 1960s to create records that could be read by computers and shared among libraries.

By 1971, MARC formats had become the US national standard for dissemination of bibliographic data.

Two years later, they became the international standard. There are several versions of MARC in use around the world, the most predominant being MARC 21, created in 1999 as a result of the harmonization of U.S. and Canadian MARC formats, and UNIMARC.

MARC formats

Name	Description
Authority records	Provide information about individual names, subjects, and uniform titles. An authority record establishes an authorized form of each heading, with references as appropriate from other forms of the heading.
Bibliographic records	Describe the intellectual and physical characteristics of bibliographic resources (books, sound recordings, video recordings, and so forth).
Classification records	MARC records containing classification data. For example, the Library of Congress Classification has been encoded using the MARC 21 Classification format.
Community Information records	MARC records describing a service-providing agency, such as a local homeless shelter or tax assistance provider.
Holdings records	Provide copy-specific information on a library resource (call number, shelf location, volumes held, and so forth).

MARC 21

MARC 21 was designed to redefine the original MARC record format for the 21st century and to make it more accessible to the international community. MARC 21 has formats for the following five types of data: Bibliographic Format, Authority Format, Holdings Format, Community Format, and Classification Data Format. Currently MARC 21 has been implemented successfully by The British Library, the European Institutions and the major library institutions in the United States, and Canada.

MARC 21 is a result of the combination of the United States and Canadian MARC formats (USMARC and CAN/MARC). MARC 21 is based on the NISO/ANSI standard Z39.2, which allows users of different software products to communicate with each other and to exchange data.

MARC 21 allows the use of two character sets, either MARC-8 or Unicode encoded as UTF-8. MARC-8 is based on ISO 2022 and allows the use of Hebrew, Cyrillic, Arabic, Greek, and East Asian scripts. MARC 21 in UTF-8 format allows all the languages supported by Unicode

Protocols:

Protocols are a set of rules governing exchange of information in an easy, reliable and secure way in the network. The most common protocols used to transmit and receive over a network.

Z39.50 Protocol:

Z39.50 is an international standard client–server, application layer communications protocol for searching and retrieving information from a database over a TCP/IP computer network. It is covered by ANSI/NISO standard Z39.50, and ISO standard 23950. The standard's maintenance agency is the Library of Congress.

Z39.50 is widely used in library environments, often incorporated into integrated library systems and personal bibliographic reference software. Interlibrary catalogue searches for interlibrary loan are often implemented with Z39.50 queries.

Work on the Z39.50 protocol began in the 1970s, and led to successive versions in 1988, 1992, 1995 and 2003. The Contextual Query Language (formerly called the Common Query Language) is based on Z39.50 semantics.

Search syntax

The protocol supports search, retrieval, sort, and browse. Search queries contain attributes, typically from the bib-1 attribute set which defines **six attributes** to specify information searches on the server computer:

use, relation, position, structure, truncation, completeness.

The syntax of Z39.50 allows for very complex queries. The syntax of Z39.50 is abstracted from the underlying database structure.

 For example, if the client specifies an author search using attribute 1003, the server must determine how to map that search to the indexes it contains. This allows Z39.50 queries to be formulated without knowing anything about the target database, but it also means that results for the same query can vary widely among different servers. One server may have an author index and another may use its index of personal names, whether they are authors or not. A third may have no name index and fall back on its keyword index, and yet another may have no suitable index and return an error.

Modernization efforts

Z39.50 is a pre-Web technology, and various working groups are attempting to update it to fit better into the modern environment. These attempts fall under the designation ZING (Z39.50 International: Next Generation), and pursue various strategies

The successors to Z39.50 are the twin protocols SRU/SRW (Search/Retrieve via URL/Search/Retrieve Web service), which drop the Z39.50 communications protocol (replacing it with HTTP) while still attempting to preserve the benefits of the query syntax. SRU is REST-based, and enables queries to be expressed in URL query strings; SRW uses SOAP. Both expect search results to be returned as XML.

Firewall information

The registered internet IP port for Z39.50 is 210. Although the majority of servers use this port, there are dozens of other port numbers used worldwide by Z39.50 (e.g. 2100, 2200, 2210, 2213, 3520, or in one case, ports 2101 and higher for different databases).

File Transfer Protocol (FTP)

FTP stands for File transfer protocol. FTP is a standard internet protocol provided by TCP/IP used for transmitting the files from one host to another.

It is mainly used for transferring the web page files from their creator to the computer that acts as a server for other computers on the internet.

It is also used for downloading the files to computer from other servers.

Objectives of FTP

It provides the sharing of files.

It is used to encourage the use of remote computers.

It transfers the data more reliably and efficiently.

Why FTP?

Although transferring files from one system to another is very simple and straightforward, but sometimes it can cause problems.

 For example, two systems may have different file conventions. Two systems may have different ways to represent text and data. Two systems may have different directory structures. FTP protocol overcomes these problems by establishing two connections between hosts. One connection is used for data transfer, and another connection is used for the control connection.

How FTP works

FTP is a client-server protocol that relies on two communications channels between client and server: a command channel for controlling the conversation and a data channel for transmitting file content. Clients initiate conversations with servers by requesting to download a file. Using FTP, a client can upload, download, delete, rename, move and copy files on a server. A user typically needs to log on to the FTP server, although some servers make some or all of their content available without login, known as anonymous FTP.

FTP sessions work in passive or active modes. In active mode, after a client initiates a session via a command channel request, the server initiates a data connection back to the client and begins transferring data. In passive mode, the server instead uses the command channel to send the client the information it needs to open a data channel. Because passive mode has the client initiating all connections, it works well across firewalls and Network Address Translation (NAT) gateways.
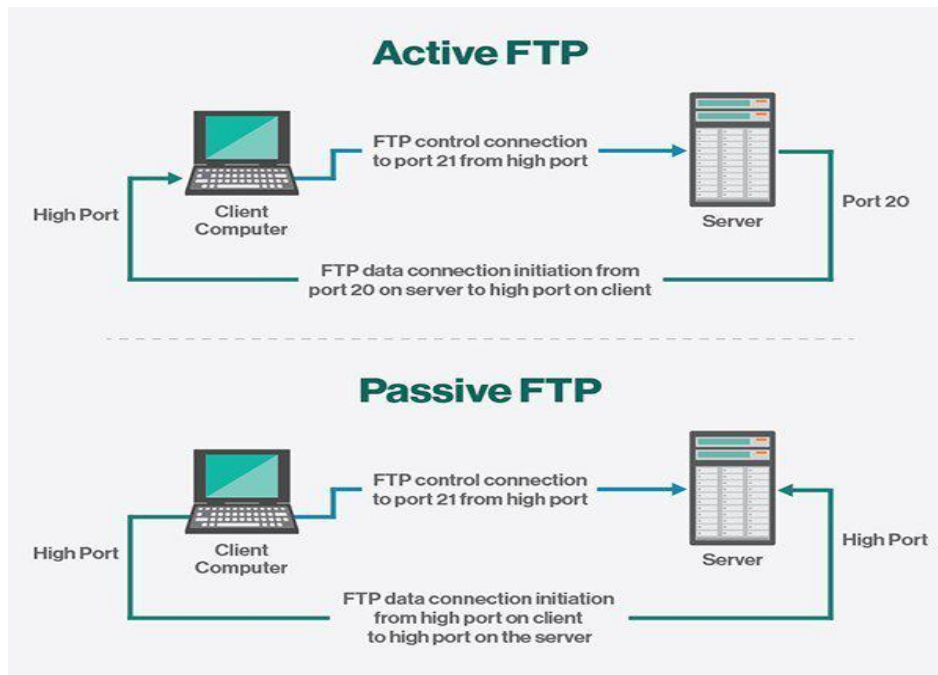
Users can work with FTP via a simple command line interface (for example, from a console or terminal window in Microsoft Windows, Apple OS X or Linux) or with a dedicated graphical user interface (GUI). Web browsers can also serve as FTP clients.

The FTP client has three components: the user interface, control process, and data transfer process.

The server has two components: the server control process and the server data transfer process.

There are two types of connections in FTP:

Control Connection: The control connection uses very simple rules for communication. Through control connection, we can transfer a line of command or line of response at a time. The control connection is made between the control processes. The control connection remains connected during the entire interactive FTP session.

Data Connection: The Data Connection uses very complex rules as data types may vary. The data connection is made between data transfer processes. The data connection opens when a command comes for transferring the files and closes when the file is transferred.

FTP Clients

FTP client is a program that implements a file transfer protocol which allows you to transfer files between two hosts on the internet.

It allows a user to connect to a remote host and upload or download the files.

It has a set of commands that we can use to connect to a host, transfer the files between you and your host and close the connection.

The FTP program is also available as a built-in component in a Web browser. This GUI based FTP client makes the file transfer very easy and also does not require to remember the FTP commands.

Advantages of FTP:

Speed: One of the biggest advantages of FTP is speed. The FTP is one of the fastest way to transfer the files from one computer to another computer.

Efficient: It is more efficient as we do not need to complete all the operations to get the entire file.

Security: To access the FTP server, we need to login with the username and password. Therefore, we can say that FTP is more secure.

Back & forth movement: FTP allows us to transfer the files back and forth. Suppose you are a manager of the company, you send some information to all the employees, and they all send information back on the same server.

Disadvantages of FTP:

The standard requirement of the industry is that all the FTP transmissions should be encrypted. However, not all the FTP providers are equal and not all the providers offer encryption. So, we will have to look out for the FTP providers that provides encryption.

FTP serves two operations, i.e., to send and receive large files on a network. However, the size limit of the file is 2GB that can be sent. It also doesn't allow you to run simultaneous transfers to multiple receivers.

Passwords and file contents are sent in clear text that allows unwanted eavesdropping. So, it is quite possible that attackers can carry out the brute force attack by trying to guess the FTP password.

It is not compatible with every system.

TCP/IP (Transmission Control Protocol/Internet Protocol)

The entire Internet Protocol suite -- a set of rules and procedures -- is commonly referred to as TCP/IP, though others are included in the suite. The TCP/IP protocol suite functions as an abstraction layer between internet applications and the routing/switching fabric.

TCP/IP specifies how data is exchanged over the internet by providing end-to-end communications that identify how it should be broken into packets, addressed, transmitted, routed and received at the destination.

TCP/IP requires little central management, and it is designed to make networks reliable, with the ability to recover automatically from the failure of any device on the network.

The two main protocols in the Internet Protocol suite serve specific functions.

1. **TCP** defines how applications can create channels of communication across a network. It also manages how a message is assembled into smaller packets before they are then transmitted over the internet and reassembled in the right order at the destination address.

2. **IP** defines how to address and route each packet to make sure it reaches the right destination. Each gateway computer on the network checks this IP address to determine where to forward the message.

Common types of TCP/IP include the following:

•**HTTP (Hyper Text Transfer Protocol)** handles the communication between a web server and a web browser.

•**HTTPS (Secure HTTP)** handles secure communication between a web server and a web browser.

•**FTP (File Transfer Protocol)** handles transmission of files between computers.

How TCP/IP works

TCP/IP uses the client-server model of communication in which a user or machine (a client) is provided a service (like sending a webpage) by another computer (a server) in the network.

Collectively, the TCP/IP suite of protocols is classified as stateless, which means each client request is considered new because it is unrelated to previous requests. Being stateless frees up network paths so they can be used continuously.

The transport layer itself, however, is stateful. It transmits a single message, and its connection remains in place until all the packets in a message have been received and reassembled at the destination.

The TCP/IP model differs slightly from the seven-layer Open Systems Interconnection (OSI) networking model designed after it. The OSI reference model defines how applications can communicate over a network.

Importance of TCP/IP

TCP/IP is non proprietary and, as a result, is not controlled by any single company. Therefore, the Internet Protocol suite can be modified easily. It is compatible with all operating systems,

so it can communicate with any other system. The Internet Protocol suite is also compatible with all types of computer hardware and networks

TCP/IP is highly scalable and, as a routable protocol, can determine the most efficient path through the network. It is widely used in current internet architecture.

Uses of TCP/IP

TCP/IP can be used to provide remote login over the network, for interactive file transfer, to deliver email, to deliver web pages over the network and to remotely access a server host's file system.

Pros and cons of TCP/IP

The advantages of using the TCP/IP model include the following:

•helps establish a connection between different types of computers;

•works independently of the operating system;

•supports many routing protocols;

•has client-server architecture that is highly scalable;

•can be operated independently;

•supports several routing protocols; and

•is lightweight and doesn't place unnecessary strain on a network or computer.

The disadvantages of TCP/IP include the following:

•complicated to set up and manage;

•transport layer doesn't guarantee delivery of packets;

•not easy to replace protocols in TCP/IP;

•doesn't clearly separate the concepts of services, interfaces and protocols so not good for describing new technologies in new networks; and

•especially vulnerable to a SYN (synchronization) attack, which is a type of denial-of-service attack in which a bad actor uses TCP/IP protocol.

Hyper Text Transfer Protocol (HTTP)

HTTP is a protocol which allows the fetching of resources, such as HTML documents. It is the foundation of any data exchange on the Web and it is a client-server protocol, which means requests are initiated by the recipient, usually the Web browser.

A complete document is reconstructed from the different sub-documents fetched, for instance text, layout description, images, videos, scripts, and more.
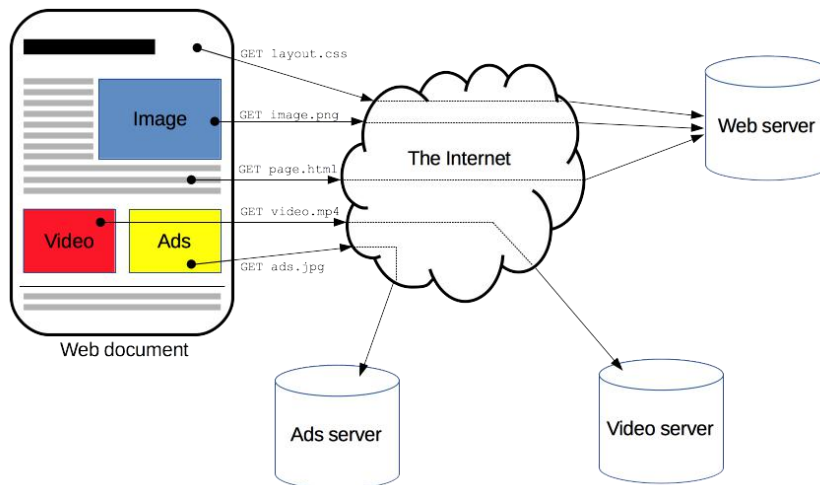

Clients and servers communicate by exchanging individual messages (as opposed to a stream of data). The messages sent by the client, usually a Web browser, are called requests and the messages sent by the server as an answer are called responses.

Components of HTTP-based systems

HTTP is a client-server protocol: requests are sent by one entity, the user-agent (or a proxy on behalf of it). Most of the time the user-agent is a Web browser, but it can be anything, for example a robot that crawls the Web to populate and maintain a search engine index.

Each individual request is sent to a server, which handles it and provides an answer, called the response. Between the client and the server there are numerous entities, collectively called proxies, which perform different operations and act as gateways or caches, for example.
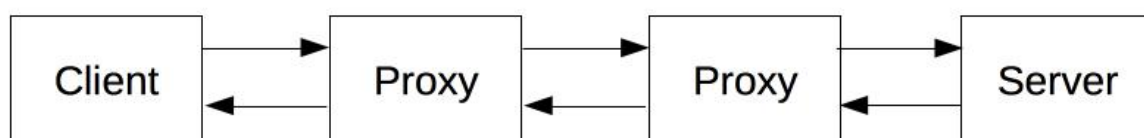
What can be controlled by HTTP

Here is a list of common features controllable with HTTP.

•Caching

How documents are cached can be controlled by HTTP. The server can instruct proxies and clients, about what to cache and for how long. The client can instruct intermediate cache proxies to ignore the stored document.

•Relaxing the origin constraint

To prevent snooping and other privacy invasions, Web browsers enforce strict separation between Web sites. Only pages from the same origin can access all the information of a Web page. Though such constraint is a burden to the server, HTTP headers can relax this strict separation on the server side, allowing a document to become a patchwork of information sourced from different domains; there could even be security-related reasons to do so.

•Authentication

Some pages may be protected so that only specific users can access them. Basic authentication may be provided by HTTP, either using the WWW-Authenticate and similar headers, or by setting a specific session using HTTP cookies.

•Proxy and tunneling

Servers or clients are often located on intranets and hide their true IP address from other computers. HTTP requests then go through proxies to cross this network barrier. Not all proxies are HTTP proxies. The SOCKS protocol, for example, operates at a lower level. Other protocols, like ftp, can be handled by these proxies.

•Sessions

Using HTTP cookies allows you to link requests with the state of the server. This creates sessions, despite basic HTTP being a state-less protocol. This is useful not only for e-commerce shopping baskets, but also for any site allowing user configuration of the output.

UNIT - V

Library automation software

Library Automation refers to the use of computer to keep track of all the books that are issued, returned and added in the library. Library Automation is another name of Library Management System. It is a user friendly system. It is created to ensure the proper management of books in a library.

Commercial software

"Commercial software" is software, which is available in market at a certain price. Which means, whosoever needs to use it has to pay for it to get a valid and authorized license; of course, the amount is too high because it is commercial and at certain level depends on market value and share. The commercial software is either developed by an individual or a group or an organization for an

individual, group or organization to meet a specific or chunk of requirements. Certain funds, time and measurable efforts are invested in order to meet these requirements and to achieve the final product hence, the cost to be retrieved/ regained is by means of selling it thus, are distributed commercially!

Open source software

"Open source software" is again a form of software developed either by an individual, group or an organization to meet certain requirements but here the case of its development is different that is, it's not developed for some individual, group or organization but is developed based on its developing body's interest and is published openly for general public. Open Source software also sometimes termed as "OSS" is computer software whose source code is available to the general public with relaxed or non-existent intellectual property restrictions (or arrangement such as the public domain) and is usually developed with the input of many contributors. The openly viewable nature of the source provides better flexibility and that if possible problems are found, they can be quickly addressed and altered to adapt, again with the supervision of more than one company / programming team located at various locations globally.

Pros and Cons of Commercial and Open source software:

Cost — As stated earlier the commercial software costs more and is ideally a big hurdle in selecting and investing funds over the open source software because, open source software is available freely; you hardly have to pay anything for it.

Security — It is again a big question. Commercial software is mostly developed with certain intension and under consideration of specific standards hence security constraints are taken care of in a better way. This may not be the case with Open source software — you may have to compromise over security (up to some extent)!

Reliability and Ownership — A software developed by a commercial software/ solution provider will always be highly reliable as it is produced with certain considerations and parameters. The vendor developing the software is solely responsible for all the failures (if at all) and also responsible for its long term (may be life time) support and maintenance till software is literally phased out. In Open source software reliability may not be confirmed, as everyone's requirements can't be fulfilled neither there is anyone responsible to own the failures.

Customization — Until few years customization in commercial software was not possible because, the software were distributed in packaged form (to be used as it is). But, now customization options are available for few commercial software. Ideally, open source software provides wide scope of customization, which is another great advantage that, one can mould the software as required.

Availability — This may be the case in both, what you need may not be available or cannot be made available either in commercial software or open source software. Thus, meeting exact requirements may be difficult. But, both the options up to certain level provide availability of making the software as required, especially the open source.

KOHA SOFTWARE

Koha is an open source Integrated Library System (ILS), used world-wide by public, school and special libraries. The name comes from a Māori term for a gift or donation.

Features

Koha is a web-based ILS, with a SQL database (MariaDB or MySQL preferred) backend with cataloguing data stored in MARC and accessible via Z39.50 or SRU. The user interface is very configurable and adaptable and has been translated into many languages.[3] Koha has most of the features that would be expected in an ILS, including:

•Various Web 2.0 facilities like tagging, comment, Social sharing and RSS feeds

•Union catalog facility

•Customizable search

•Online Circulation

•Bar code printing

History

Koha was created in 1999 by Katipo Communications for the Horowhenua Library Trust in New Zealand, and the first installation went live in January 2000.

From 2000, companies started providing commercial support for Koha, building to more than 20 today.

2014 the Ministry of Culture (Turkey) started to use Koha -- Devinim version in 1,136 public libraries with more than 17,000,000 items and app. 2,000,000 active users. This is the biggest Koha installation for the moment.

Developers   :   Koha Community

Initial release   :   January 2000; 20 years ago

Stable release   :   19.11.04 / March 24, 2020; 8 days ago

Repository   :   git.koha-community.org/koha.git

Written in   :   Perl

Operating System: Linux

Type   :   Integrated library system

Licence   :   GNU General Public Licence v3

Website   :   koha-community.org

WINISIS

CDS/ISIS is a software package for generalised Information Storage and Retrieval systems developed, maintained and disseminated by UNESCO. It was first released in 1985 and since then over 20,000 licences has been issued by UNESCO and a worldwide network of distributors. It is particularly suited to bibliographical applications and is used for the catalogues of many small and medium-sized libraries.

Versions have been produced in Arabic, Chinese, English, French, German, Portuguese, Russian and Spanish amongst other languages. UNESCO makes the software available free for non-commercial purposes, though distributors are allowed to charge for their expenses.

CDS/ISIS is an acronym which stands for **Computerised Documentation Service / Integrated Set of Information Systems.** In 2003 it was stated that "This package is accepted by libraries in the developing countries as standard software for information system development".

Features

The major features of the CDS/ISIS software are:

The handling of variable length records, fields and sub fields, thus saving disk space and making it possible to store greater amounts of information;

The handling of repeatable fields;

A data base definition component allowing the user to define the data to be processed for a particular application;

A data entry component for entering and modifying data through user-created data base specific worksheets;

An information retrieval component using a powerful search language providing for field-level and proximity search operators, in addition to the traditional and/or/not operators, as well as free-text searching;

A powerful sort and report generation facility allowing the user to easily create any desired printed products, such as catalogues, indexes, directories, etc.;

A data interchange function based on the ISO 2709 international standard used by leading data base producers;

An integrated application programming language (CDS/ISIS Pascal and the ISIS_DLL), allowing the user to tailor the software to specific needs;

Functions allowing the user to build relational data bases, though CDS/ISIS is not based over a relational model;

Powerful hypertext functions allow to design complex user interfaces.

Latest official versions:

Winisis 1.5 build 3 (Standard)

Winisis 1.5 build 7 (Arabic enabled)

NewGenLib

NewGenLib is an integrated library management system developed by Verus Solutions Pvt Ltd. Domain expertise is provided by Kesavan Institute of Information and Knowledge Management in Hyderabad, India.

NewGenLib version 1.0 was released in March 2005. On 9 January 2008, NewGenLib was declared free and open-source under GNU GPL.The latest version of NewGenLib is 3.1.1 released on 16 April 2015. Many libraries across the globe (mainly from the developing countries) are using NewGenLib as their Primary integrated library management system as seen form the NewGenlib discussion forum.

Features

1.Android mobile and tablet capable

2.Integration with Twitter helping send messages of transactions directly to users' Twitter accounts.

3.Flexibility of defining own search field in OPAC.

4.Enhanced contents and interactive OPAC like Availability of Book jackets, Google preview, Comments/ Book review, Tagging, Favorite reading list, etc.

5.Zotero compliant OPAC

6.RSS Feeds in OPAC

7.Faceted Browsing (Refining search results)

8.Suggestion for other books in the rack

9.RFID supports

10.Provision for frequently used predefined templates along with freedom of defining own customized data entry template s in Cataloguing

11.Configurable SMS system - a proof of transaction.

12.Integration with Gmail or paid mailbox account. This enables automatic sending of email to patrons during issue /return.

13.Enhanced Report Module for generating in .csv format with a provision for wide customization.

14.Provision for integrating with Vufind SOPAC (Ex: OPAC of the Library of Bangalore University).

15.Catalogue can be harvested through Google site map, and thus the visibility of the library can be further improved.

Technologies Used & Standards Supported

NewGenLib uses a number of well supported and widely used, reliable and well tested open source components like PostgreSQL, Apache Tomcat, and Solr Lucene.,. It is entirely Java-based, platform-neutral, and uses the following major software technologies in its presentation, web server and database layers.

•Java SE

•Apache Tomcat server

•Spring framework

•Hibernate framework

•Lucene and Solr

•JDOM for XML messaging

•Java Servlets, JavaServer Pages

•Java Mail

•OpenOffice.org for form letters

•JasperReports

•FreeMarker template (from version 3.04 R1)

NewGenLib is platform independent and can be installed on Linux and Windows operating systems.

Types of libraries

NewGenLib can be used for any type of library. Presently, it is used by Libraries of following types.

•University libraries

•College/School libraries

•Public libraries

•Libraries in Research Institutes

•Church libraries

•Libraries in Offices/Corporates

Autolib

AutoLib software is a fully integrated, versatile, user-friendly, cost-effective and multi-user Library automation software. ... It is designed to automate various activities of Libraries in Universities,

Colleges, Schools, R&D institutions, Public Libraries and Corporate, Management and Special institutions, etc.

Software for University Libraries (SOUL)

Software for University Libraries (SOUL) is an state-of-the-art integrated library management software designed and developed by the INFLIBNET Centre based on requirements of college and university libraries.

It is a user-friendly software developed to work under client-server environment. The software is compliant to international standards for bibliographic formats, networking and circulation protocols. After a comprehensive study, discussions and deliberations with the senior professionals of the country, the software was designed to automate all house keeping operations in library.

The software is suitable not only for the academic libraries, but also for all types and sizes of libraries, even school libraries. The first version of software i.e. SOUL 1.0 was released during CALIBER 2000.

The latest version of the software i.e. SOUL 2.0 was released in January 2009. The database for new version of SOUL is designed for latest versions of MS-SQL and MySQL (or any other popular RDBMS). SOUL 2.0 is compliant to international standards such as MARC 21 bibliographic format, Unicode based Universal Character Sets for multilingual bibliographic records and NCIP 2.0 and SIP 2 based protocols for electronic surveillance and control.

Features and Functionalities

•UNICODE based multilingual support for Indian and foreign languages;

•Compliant to International Standards such as MARC21, AACR-2, MARCXML;

•Compliant to NCIP 2.0 protocol for RFID and other related applications especially for electronic surveillance and self check-out & check-in;

•Client-server based architecture, user-friendly interface that does not require extensive training;

•Supports multi-platform for bibliographic database such as My SQL, MS-SQL or any other RDBMS;

•Supports cataloguing of electronic resources such as e-journals, e-books, virtually any type of material;

•Supports requirements of digital library and facilitate link to full-text articles and other digital objects;

•Support online copy cataloguing from MARC21 supported bibliographic database;

•Provides default templates for data entry of different type of documents. User can also customize their own data entry templates for different type of documents;

•Provides freedom to users for generating reports of their choice and format along with template and query parameters;

•Supports ground-level practical requirements of the libraries such as stock verification, book bank, vigorous maintenance functions, transaction level enhanced security, etc.;

•Provides facility to send reports through e-mail, allows users to save the reports in various formats such as Word, PDF, Excel, MARCXML, etc.;

•Highly versatile and user-friendly OPAC with simple and advanced search. OPAC users can export their search results in to PDF, MS Excel, and MARCXML format;

•Supports authority files of personal name, corporate body, subject headings and series name;

•Supports data exchange through ISO-2709 standard;

•Provides simple budgeting system and single window operation for all major circulation functions;

•Strong region-wise support for maintenance through regional coordinators. Strong online and offline support by e-mail, chat and through dedicated telephone line during office hours; and

•Available at an affordable cost with strong institutional support.

SOUL Requirements Hardware and Software Requirements

Processor Type : PIV or Higher

Processor Speed : 1.6 GHz or Higher

RAM : Minimum 512 MB (1 GB Recommended)

Free Hard Disk Space : 400 MB (Minimum)

OS : Windows XP SP3 , Windows Vista , Windows 2003 Server, Windows 2008 Server

Libsys(Library System)

Libsys is a Delhi based software company that develops and deploy Library management systems, Library automation system, RFID, Digital Resource Management System, E-Commerce Platform software for Online stores and business etc.

The company was founded in 1984 by Mr. Anil Jain, Post Graduated from IIT Kanpur. He started his career with TCS and after working for 10 years, started Libsys which has approximate 100+ employees with its head office in Gurgaon, Haryana and additional offices in Kolkata and Bangalore.

Libsys was founded in 1984, and its continuous growth for over 3 decades, has made LIBSYS the most trusted brand for libraries in India and DIY E-commerce Platform for Online Stores.

In 1984 The Company Incorporated takes up software development projects for general insurance and reinsurance businesses.

In 1995 British Council adopts Libsys Ltd. For its libraries in India and South Asia by securing overseas contracts.

In 2003, the very first LSmart RFID system deployed in Bank of Baroda, Mumbai, LSPremia multi-site installed at GAIL, introduced LSDigital, MARC21 add on to LIBSYS4.

In 2010, new ventures initiated in other verticals through LSTech.

In 2012 LSAcademia and LSNet launched.

In 2015, LSNetX the DIY E-commerce platform software for online stores introduced.

Features

Libsys7 is a Library Management Solution which is designed to enhance the complete Library experience through value added features and services.

LSEase is based on client-server architecture and strengths to manage library with basic needs. It requires minimal data entry and easy data backup even for the large databases.

LSAcademia is a total ERP Solution to manage Academic campus, comes in 2 variants LSAcademia School Edition & LSAcademia Standard Edition

LSDigital(DRMS) is a document digitization software which is used to ease management, multi-access, storage space reduction and preservation.

LIBSYS is a solution based on RFID and Em Technologies.

An Online Bookstore with huge database based on eCommerce platform.

Do it yourself E-commerce platform for all offline business who are looking to be online. Very Easy and self explanatory.