

Chapter Five

Digital Collections and Content Organization

- ✓ Digital Collections & Content Organization
- ✓ Abstracting and indexing service
- ✓ Semantic web and ontology's
- ✓ Knowledge management (KM)
- ✓ Building digital collections

Digital Collections

- ✓ In the digital library, collections are transformed through the integration of new formats, licensed content and
- ✓ Third-party information over which the library has little or no direct curatorial control

Content Organization

- ✓ Knowledge Organization Systems (KOS) such as classification schemes, authority files, semantic networks, ontology's are mechanisms for content organization
- ✓ These systems are used to organize content for the purpose of retrieval and to manage a collection. This serves as a bridge between the user's information need and material in the collection
- ✓ Knowledge Organization Systems help the user to identify their need without prior knowledge of its existence

Content Organization

KOS guides the user through a discovery process
includes :-

- ✓ Index Terms
- ✓ knowledge organization systems,
- ✓ Abstracting/indexing service,
- ✓ Metadata
- ✓ Semantic web
- ✓ Ontology and
- ✓ knowledge management

Common characteristics of KOS

- ✓ It imposes a particular view of the world on a collection
- ✓ Same entity can be characterized in different ways
- ✓ Sufficient commonality between the concept expressed in the system and real world object

Types of KOS

- ✓ Term Lists
- ✓ Authority files
- ✓ Glossaries,
- ✓ Dictionaries
- ✓ Gazetteers'
- ✓ Classifications and Categories
- ✓ Subject headings
- ✓ Classification schemes
- ✓ Relationship lists
- ✓ Thesauri
- ✓ Semantic Networks
- ✓ Ontology

Origin and use of KOS

- ✓ **In the physical library classification schemes such as:-**
 - o *Universal Decimal Classification (UDC)*
 - o *Dewey Decimal Classification (DDC)*
 - o *Library of congress(LC)*

- ✓ Reflect the need to store a single item at a single location on a shelf

Abstracting and indexing service

- ✓ This is a very fruitful way of knowledge management
- ✓ The user can go through the abstract of the document and decide whether its relevant for further references
- ✓ In a physical library, professional librarian can go through the document and trace the major keywords for indexing purpose
- ✓ This helps to user to search his needs without wasting time and can find out exact information for his research purpose

Planning KOS

- ✓ When analyzing how the KOS might be used with a particular digital library it is essential to understand the environment of the user
- ✓ Librarian should analyze the user needs and its necessary to locate KOSs to meet the need
- ✓ Many users are aware of KOSs on the web within their discipline
- ✓ Developers may also turn to directories, librarians in the field, and reference sources, or they may perform a general search of the Internet
- ✓ If the **KOS** is stored in database, user would need logon information such as **user ID** and **password** from the digital library system to the external KOS, in order to provide access to the web-enabled database

Metadata

- ✓ Data about data or information about information
- ✓ In the other words, **metadata** is data that describe information resources
- ✓ A short descriptive note on a book, an informal description of search hits by search engines, a catalog and **MARC (*Machine Readable Cataloging*)** record, are data that describe an information resources and hence metadata.

Metadata

- ✓ **Metadata** is any data that supports the effective use of data, including information that can facilitate knowledge management, knowledge access and analysis
- ✓ The **data** that **metadata capture** to describe an information resource can be divided into **two** categories:
 - 1. Intrinsic data:-*** are characteristics extracted directly from the information resource such as title, author, and subject
 - It is useful for knowledge management and administrative purpose.

Metadata

II. *Extrinsic data:-* are those related to the administration and other **non-bibliographic** data such as:-

- ✓ Author e-mail
- ✓ Author department
- ✓ Password or digital signature
- Which is used to facilitates resource descriptions, identification and discovery

- ✓ The next generation of web, called **semantic web**
- ✓ It is based on the machine-process able semantics of the information, stored in the machine process able metadata
- ✓ This is not a separate web but an extension of the current web in which the information is given well-defined meaning, better enabling computers and humans to work in cooperation

Ontology

- ✓ The prerequisite of this web, as its definition implies, is metadata that explicitly represent semantics of data which called ontology
- ✓ Ontology that describe knowledge in specific area are often connected with systems for data mining and knowledge management
- ✓ Ontology as a new emerging form of metadata is revolutionizing the current classificatory approaches towards semantic metadata
- ✓ Consulting the traditional metadata systems with ontological view, such as thesauri, and card catalog systems as well as converting the controlled vocabularies into the ontology indicates this change

Knowledge management

✓ ***Knowledge management:-*** is a systematic approach that ranges from technology-driven methods of accessing, controlling, and delivering information to massive efforts in transforming corporate culture

✓

- ✓ Creating digital collections is an exciting way to reach beyond traditional boundaries and bring information to people no matter where they are
- ✓ Any person, group, or institution can begin digitization at some level
- ✓ Being aware of issues such as
 - the technology infrastructure
 - the pros and cons of major file formats, and
 - the intended uses of the collections can make the creation process easier and more efficient

Questions?