CHAPTER 3

CONCEPTUAL FRAMEWORK
SECTION I

INFORMATION RETRIEVAL (IR)
3.1.1) MULTIMEDIA INFORMATION RETRIEVAL (MIR) AND COMMUNICATION

Yu (2004) indicated that the research on Information Retrieval started in the 1960s and initially focused on textual document retrieval. Now it has become a very wide research field that studies the representation, storage, organization, and access to various information items (e.g., documents, image, video, audio etc.). Hence, it has been named Multimedia Information Retrieval (MIR). Smeulders et al. (2000) denoted the primary goal of MIR, to retrieve relevant information items in response to user queries, while returning as few irrelevant ones as possible. One example is content-based image retrieval (CBIR). In fact, the term MIR is not limited to the retrieval of information. It covers a wide range of topics in information processing which are related to information retrieval including information translation, summarization and classification as well.

DUC A. Tran (1996) discussed that proliferation of Internet, the needs and methods for accessing and sharing information have reached a new height. Users have become more demanding with interests not only in traditional alphanumerical contents but also in multimedia contents. Requests for the richer data have amplified in parallel with the emergence of many important applications including digital libraries, distance learning, public information systems, e-commerce, movie on demand, and other corporate communications. Many of these applications allow the user to watch or listen to the stream of data as it arrives. Overlapping the layout of the data at the receiver with the transmission by the sender is referred to as multimedia streaming. ChengFu Chou (2002) indicated that technological advances in digital signal processing, data compression techniques, and high speed communication networks have made multimedia
applications (such as large-scale continuous media servers) and wide-area upload applications (such as distance education over Internet) an active research area. Rao et al. (2008) discussed that a new infrastructure of digital audio, image and video recorders and players, online services, electronic commerce and education is rapidly being deployed. Digital media offers several distinct advantages over analog media. The quality of digital audio, image and video signals is higher than that of their analog computers. Editing is easy because one can access the exact discrete locations that need to be changed. Copying is simple with no loss of fidelity. A copy of digital media is identical to the original. Digital audio, image and video are easily transmitted across networked information systems. These advantages have opened up many new possibilities. Multimedia consists of Multimedia data set of interactions. Multimedia data is informally considered as the collection of three Ms: multi-source, multi-type and multi-format data. The interactions among the multimedia components consist of complex relationships without which multimedia would be a simple set of visual, audio and other data. Multimedia and multimedia communication can be globally viewed as a hierarchical system. The multimedia software and applications provide a direct interactive environment for users. When a computer requires information from remote computers or servers, multimedia information must travel through computer networks. Because the amount of information involved in the transmission of video and audio can be substantial, the multimedia information must be compressed before it can be sent through the network in order to reduce the communication delay. Constraints, such as limited delay and jitter, are used to ensure a reasonable video and audio effect at the receiving end. Therefore, communication networks are undergoing constant
improvements in order to provide for multimedia communication capabilities. LANs are used to connect local computers and other equipment and Wide Area Networks (WANs) and the Internet connects the LANs together. Better standards are constantly being developed, in order to provide a global information superhighway across which multimedia information will travel. Multimedia communication is the field referring to the representation, storage, retrieval and dissemination of machine-processable information expressed in multiple media, such as text, image, graphics, speech, audio, video, animation, handwriting and data files. Multimedia best suits the human being's complex perception and communicating behavior, as well as the way of acting. Namely, it does not only provide communication capabilities and information sharing for people, irrespective of location and time, but it also provides easy and immediate access to widely distributed information banks and information processing centers. Applications in medicine, education, travel, real estate, banking, insurance, administration and publishing are emerging at a fast pace. These applications are characterized by large multimedia documents that must be communicated within very short delays. Computer-controlled cooperative work, whereby a group of users can jointly view, create, edit and discuss multimedia documents, is going to be characteristic of many transactions. Some glamorous applications in multimedia processing include distance learning, virtual library access and living books. In distance learning, students learn and interact with instructors remotely across a broadband communication network. Virtual library access means that they instantly have access to all of the published material in the world, in its original form and format, and that they can browse, display, print and even modify the material instantaneously. Living books supplement the written word and the associated pictures
with animations, and hyperlinks provide access to supplementary material. Applications that are enabled or enhanced by video are often seen as the primary justification for the development of multimedia networks.

3.1.1.1) MULTIMEDIA COMMUNICATION MODEL

(Roa et al., 2008) described multimedia communication model in the following words:

"A multimedia communication model is strongly influenced by the manufacture-dependent solutions for PCs, and workstations, including application software on the one hand and the intelligent network concept on the other." (p.30)

A layered model for future multimedia communication comprises five components, which are as follow:

- Partitioning of complex information objects into distinct information types for the purpose of easier communicating, storing and processing. This comprises data, video or audio and takes into account the integration of different information types.
- Standardization of service components per information type, possibly with several levels of quality per information type.
- Creation of platform as two levels: a network service platform and a multimedia communication platform.
- Definition of generic applications for multiple use in various multimedia environments and different branches meeting common widespread needs.
- Specific applications.

3.1.1.2) ELEMENTS OF MULTIMEDIA SYSTEMS

Multimedia systems generally use two key communications modes: person-to-person communications and person-to-machine communications. Figure (3.1.1.2) presents key elements of multimedia systems.
3.1.1.2.1) In the person-to-person mode shown in Figure 3.1.1.2 (a), there is a user interface that provides the mechanisms for all users to interact with each other, and there is a transport layer that moves the multimedia signals from one user location to some or all other user locations associated with the communications. The user interface creates the multimedia signal and allows users to interact with the multimedia signal in easy-to-use manner. The transport layer preserves the quality of the multimedia signals so that all users receive what they perceive to be high-quality signals at each user location. Examples of applications for the person-to-person mode are teleconferencing, videophones, distance learning and shared workspace scenarios.

3.1.1.2.2) In the person-to-machine mode, shown in Figure 3.1.1.2 (b), there is again a user interface for interacting with the machine, along with a transport layer for moving the multimedia signal from the storage location to the user. Some mechanism for storage and retrieval of multimedia signals are either created by
the user or requested by the user. The storage and retrieval mechanisms involve browsing and searching to find existing multimedia data. In addition, these mechanisms involve storage and archiving in order to move user-created multimedia data to the appropriate place for access by others. Examples of applications for person-to-machine mode include creation and access of business meeting notes, access of broadcast video and document archives from a digital library or other repositories.

3.1.1.3) MEDIA INTERACTION

Media interaction is shown in Figure (3.1.1.3). As can be seen, media are categorized into three major classes. The first is textual information, the second is an audio, including speech and music, and third represents image and video. The goal of speech recognition is to transcribe spoken inputs literally into individual words, but the goal of spoken language understanding research is to extract meaning from whatever was recognized.

Figure (3.1.1.3): Media Interaction [2.1]. © 1990 IEEE (cited in Rao et al., 2008)
3.1.2) INFORMATION RETRIEVAL (IR) & DOCUMENTATION

Maristella Agosti (2007) referred to the term information retrieval which identifies that a person - the user - has to conduct to choose from a collection of documents, those that can be of interest to him to satisfy a specific and contingent information used. It allows that the aim of the area of information retrieval is to help and support the user in choosing, among the available documents, those that, with higher probability are more suitable to satisfy his information need. Figure (3.1.2.1) sketches the situation: the user has the possibility of choosing the documents of his interest from an available collection, but he needs to have a tool that can help him in choosing the subset of documents, which are of his interest without needing to invest a lot of time inspecting all the document of the collection. Figure (3.1.2.1) also shows the three main actors and aspects that information retrieval needs to address:

- User,
- Collection of documents,
- Retrieval, which means a function or model used in retrieval and accessing information.

Figure (3.1.2.1): Information Retrieval Aim according to A. Maristella (2007)
Doyle, L. B. (1975) and Salton, G. (1963) indicated that the study of hypertext methods was at the basis of a new sort of network-based and associative information retrieval entitled EXPLICIT. Associative information retrieval methods are those retrieval methods that have been proposed and experimented since the early days of information retrieval. The EXPLICIT model was based on a two-level architecture initially proposed in the hypertext as an effective information retrieval tool for the final user that introduced by Agosti et al. (1989) and refined in an approach for the conceptual modeling of IR auxiliary data introduced by Agosti et al. (1990). The two-level architecture holds the two main parts of the information resource managed by an information retrieval tool. On the one hand the collections of content objects (e.g. a single collection of documents, different collection of different types of digital content objects). Moreover, on the other the term structure, which is a scheme of concepts that can be composed of either one single indexing structure or some cooperative content representation structures such as those depicted in Figure (3.1.2.2) in a sort of "semantic network".

![Figure (3.1.2.2): the EXPLICIT architecture and model according to Agosti et al. (1990)](image-url)
The system manages this network to retrieve information of use for the final user, but also to present the representation of the contents of the collections to the user, who can use them for browsing and becoming acquainted with the information richness of the managed collections. Multimedia retrieval is the other relevant area that researchers started to face in a systematic way and for different media during the decade. The complexity of the management of collections of multimedia digital documents can be faced in particular for information retrieval purposes, but also from a general architectural point of view, that is, the area of digital libraries and digital library systems. The IMS (Information Multimedia System) research group is participating in the search in Audio-visual content using Peer-to-Peer (P2P) information retrieval (SAPIR) project (http://www.sapiir.eu). Fabio Crestani discussed that information retrieval techniques have been used for a long time to identify links between textual items for automatic construction of hypertexts and electronic books where sought information can be accessed by browsing. Information retrieval (IR) has been created with the purpose of the enabling a user to access information relevant to the user information need expressed in a query, where the information is contained in large archives of textual data. The term "information retrieval" is described as "searching and retrieval of information from storage according to specific subject." The word retrieval means to discover and bring to the notice of the users, the documents in which information is embedded. Vickery (1970) described the word retrieval as "retrieval is essential concerned with the structure of the operation of the device to select documentary information from the store of information in response to several questions". Salton (1979) and Salton and Fox (1985) defined an information retrieval system as a "system used to store items of
information that need to be processed, searched, retrieved, and disseminated to various user population. Nowadays multimedia indexing and retrieval techniques are being developed to access image, video, and sound database without text descriptions. Any information retrieval system is based on some theory such as classification theory, linguistic theories in the context of automatic indexing, psychological approaches and early structural models of Fairthone and others (Fairthone, 1961). Therefore, the different models of information retrieval systems are as follows: Models Based on Input and Output such as Data Retrieval Model, Information Retrieval Model, and Knowledge Retrieval Model. Models Based on Theories and Tools such as Linguistic Models (Ponte & Croft, 1998), Mathematical Models, Psychological Model Brooks, 1968; Oddy, 1977. Moreover, the Economic Model, User Model in Information Retrieval, Logical Models, Vector Processing Models, Vector Space Model (Salton, Yang & Wong, 1975), Probabilistic Models (Fuhr, 1992; Robertson, 1977), Cognitive Models (Ingwersen, 1992), the Fuzzy Set Model (Radecki, 1979; Lee, 1999), Cluster Model (Ellis, 1990; Sowa, 2000; Deewester, Dumais, Furnas, Landauer, & Harshman, 1990). Oard & Dorr (1996) defined that an information retrieval is a process by which users seek to locate documents that contain information about the subject of their query. The information retrieval process is distinguished from the conventional database access paradigm by the user's desire to find documents about a subject rather than data that directly answers the query. Lancaster (1968) defined that an information retrieval system does not inform (i.e. change the knowledge of) the user on the subject of his inquiry but it merely informs on the existence (or nonexistence) and whereabouts of
documents rating to this request. Figure (3.1.2.3) depicts a typical information retrieval system.

Figure (3.1.2.3): A typical information retrieval system according to Oard & Dorr (1996)

With the advanced of technologies in World Wide Web and digital libraries, the traditional discipline of information retrieval has been extended to multimedia including image, graphics, video, and speech, and multi-languages, including European, Asian and other languages. Users expect to submit a textual query to the digital library in one language and obtain all the relevant information in all media and languages. For example, the following are some scenarios of the application CLIR in digital libraries (Cross-Language Information Retrieval):

1. The digital library has a collection of monolingual documents but it supports users who speak different languages. In this case, the queries may be of different languages. The queries must be translated into document languages before retrieval.

2. The digital library has a collection of parallel documents. Users who know only own language may search in the digital library using his language and
obtain the relevant documents in the same language. The corresponding documents in other languages that are parallel to retrieved documents are also extracted. In this case, the digital library must be able to align the parallel documents if the documents do not come in pairs from the source.

3. The digital library has a collection of multi-lingual documents. Multiple languages may exist in individual document.

Cross-lingual information retrieval refers to the ability to process a query for information in one language, search a collection of objects, including text, images, audio files, etc. and return the most relevant objects, translated into the user's language of necessary (Klavans et al., 1999; Oard & Dorr, 1996). An information retrieval model has two major components, (i) representation of queries and documents, and (ii) comparisons of these representations. The objectives of an information retrieval system are to automate the process of examining documents by computing the comparisons between the representations of queries and documents (Turtle & Croft, 1991; Van Rijsbergen, 1979). See Figure (3.1.2.4) below that depicts the representation and comparison processes in an information retrieval model.
3.1.3) CONTENT-BASED IMAGE RETRIEVAL (CBIR)

Nowadays, the needs of electronic imaging are growing rapidly. The digital libraries are adding up visual data such as photographs, videos, motion pictures, maps, manuscripts, satellite pictures and other special forms of data. Images of these materials
are a significant component of digital libraries for the most demanding in terms of conversion, storage and retrieval (B. P. Prakash, 1996\textsuperscript{5}). Recent years have witnessed a rapid increase of the volume of digital retrieval. Early research of image retrieval is searching by manually annotating every image in a database. However, these text-based techniques are impractical for two reasons: large size of image databases and subjective meanings of images (Wu et al., 2000\textsuperscript{11}). To avoid manual annotating, an alternative approach is content-based image retrieval (CBIR), which automatically retrieves images of user interest from large image databases based on the visual contents such as color, texture, shape, etc. It has been an active and fast advancing research area since the 1990s (Smeulders et al., 2000\textsuperscript{60}). The task of image retrieval is to find as many as possible “similar” images to the query images in a given database. The retrieval system acts as a classifier to divide the images in the database into two classes, either relevant or irrelevant. Many supervised learning approaches have been employed to approach this classification problem. Successful examples of learning approaches in content-based image retrieval can be found in the literature (Tieu and Viola, 2000\textsuperscript{64}; Cox et al., 2000\textsuperscript{11}; Tian et al., 2000\textsuperscript{62}; Rui et al., 1999\textsuperscript{52}; Vasconcelos & Lippman, 2000\textsuperscript{58}; Tong and Chang, 2001\textsuperscript{65}; Yang, 2003\textsuperscript{73}; Laaksonen, 1999\textsuperscript{32}). Furthermore, concept-based image retrieval has been proposed, which is a combination of text-based image retrieval and content-based image retrieval. To overcome the limitations of text-based image retrieval, content-based image retrieval (CBIR) systems emerged in the early 1990s (Kato, 1982\textsuperscript{30}; Faloutsos et al., 1994\textsuperscript{22}). In CBIR, features directly derived from their visual content by using image-processing techniques automatically or semi-automatically indexed the images. The common functionalities in CBIR can be summarized as follows (Smeulder et
Image processing and pattern recognition techniques are used to extract low-level features, such as color, texture, shape, etc. from image.

For a given feature, a representation of the feature in a vector form and a notion of similarity are determined, and image is represented as a collection of features.

Finally, image retrievals are performed on computing similarity in feature spaces and results are ranked based on the similarity values computed.

Eakins John (2002) identified three distinct levels of abstraction of search requirement with increasing complexity. Level 1 comprises retrieval by primitive features such as color, texture, shape or the spatial location of image elements. At level 2, some degree of object and scene recognition as well as inference about the image content is required. Queries at this level may contain specific objects and scenes. Level 3 comprises retrieval by abstract attributes, involving a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted. This includes retrieval of named events of pictures with emotional or religious significance, etc. Searches in CBIR can also be distinguished into three major categories (Cox et al., 1996): target search category-specific search and open-ended search or browsing. The ten-level visual structure presented according to Jaime and Chang (2000) that provides an elaborate and systematic way of abstracting images based on syntax and semantics. Syntax refers to the way visual elements are arranged without considering the meaning of such arrangements (e.g., color, texture, etc.). Semantics, on the other hand, deals with the meaning of those elements and of their arrangements (e.g., objects, events, etc.).
1. Color is the most widely used feature in CBIR, since it is an important dimension of human visual perception and it is invariant with respect to image scaling, translation and rotation and above all it is computationally least intensive (Rui et al., 1999; Del Bimbo, 1999).

2. Texture, although there is no strict definition of the image texture, it is easily perceived by human and is believed to be a rich source of visual information. The existing texture descriptors are classified based on three different approaches as statistical, Model-based and Transform-based (Haralick, 1979; Tamura, Mori, and Yamawaki, 1978; Picard and Minka; Manjunath and Ma, 1996). The texture descriptor in MPEG-7 facilitates browsing and similarity retrieval in image and video databases. There are three texture descriptors as homogeneous texture, edge histogram, and texture browsing (ISO/IEC JTC1, 2000).

3. Shape, in many situations, people can recognize an object only by its shape and it is probably the most important property that is perceived about objects. Generally, there are groups of shapes descriptors (Mehrotra and Gary, 1995): boundary or contour-based shape descriptors and region based shape descriptors. Boundary representations emphasize the closed curve that surrounds the shape. This curve has been described by numerous models, including chain codes, polygons, circular arcs, splines, explicit and implicit polynomials, and boundary Fourier descriptors. Region bases shape descriptor on the other hand, emphasizes the material within the closed boundary or based on the entire shape region.

Image retrieval based on multi-modal information sources has been recently gaining popularity due the huge amount of multi-modal information available on the Web (i.e., images with collateral texts in image captions, headers, titles and other places in HTML OR XML documents) (Chen, 2006; Santini, 2002; Westerveld, 2000; Chen et al., 2001; Lu et al., 2003; Sclaroff et al., 1995; Rong and William, 2002). There are two main combination techniques currently investigated (Chen, 2006): the text and
image modalities are sequentially used; and the text and image modalities are simultaneously used, combined either linearly or nonlinearly.

3.1.4) HYPERTEXT AND HYPERMEDIA CONCEPTUALIZATION

Hypertext may simply be defined, as an electronic system to manage a collection of information that can be accessed none sequentially. It consists of nodes or ‘chunks’ of information and logical links between them. The variety of nodes and links that can be defined make hypertext a very flexible structure in which information is provided both by what is stored in each node and by the way these nodes are linked to each other. The importance of hypertext systems lies in their potential capacity to augment and amplify human intellect (Marchionini & Shneidman, 1988). Hypertext with its powerful, interactive nodes and associated information retrieval capabilities offers greater potential and flexibilities in the ways of presenting teaching materials. It offers great many possibilities for teaching and learning to both the teacher and the taught. The hypertext-based system not only presents information on a topic but also encourages thinking analytically and critically about the information from multiple perspectives. It further enables the learners to participate in the learning/teaching program actively. Another positive contribution of hypertext-based system is to enable the student to develop authoring skills and add his/her views as notes, new thoughts, arguments, comments etc. Findings of several students have provided evidence of the potential of hypertext-based learning systems. For example, a study at Broun University found that the students were
image modalities are sequentially used; and the text and image modalities are simultaneously used, combined either linearly or nonlinearly.

3.1.4) HYPERTEXT AND HYPERMEDIA CONCEPTUALIZATION

Hypertext may simply be defined, as an electronic system to manage a collection of information that can be accessed none sequentially. It consists of nodes or 'chunks' of information and logical links between them. The variety of nodes and links that can be defined make hypertext a very flexible structure in which information is provided both by what is stored in each node and by the way these nodes are linked to each other. The importance of hypertext systems lies in their potential capacity to augment and amplify human intellect (Marchionini & Shneiderman, 1988\textsuperscript{15}). Hypertext with its powerful, interactive nodes and associated information retrieval capabilities offers greater potential and flexibilities in the ways of presenting teaching materials. It offers great many possibilities for teaching and learning to both the teacher and the taught. The hypertext-based system not only presents information on a topic but also encourages thinking analytically and critically about the information from multiple perspectives. It further enables the learners to participate in the learning/teaching program actively. Another positive contribution of hypertext-based system is to enable the student to develop authoring skills and add his/her views as notes, new thoughts, arguments, comments etc. Findings of several students have provided evidence of the potential of hypertext-based learning systems. For example, a study at Broun University found that the students were
not only able to grasp the subject quickly but also developed analytically and designing skills (Beeman, W. O. et al., 1987). One of the first results of research in the automatic construction of hypertext is the recognition of the need of a conceptual modeling tool (Agosti & Crestani, 1993). The two-level EXPLICIT model is one of the conceptual models, which is based on two types of entity: data and auxiliary data. Data are the elementary objects of interest of the final user. Data are indexed and retrieved, and can consist of multimedia documents or fragments. Auxiliary data are the objects, which describe the semantic content of data. Thus the main elements of the EXPLICIT model are nodes and links, where links are used to express semantic relations between data and auxiliary data nodes. There are four different types of links: PT, TP, PP, and TT. PT and TP links represent relations between a page and terms and a term and pages respectively. A page is linked to terms describing the page content. PP links relate a page to other pages. Structural PP links, such as "next", "previous" or "go to the Table Contents" links are easily detectable. Finally, TT links are relate a term to other terms. The Hypertext Books project was carried out during 1997-2002 at the University of Padua. The objective of the project was to develop a methodology to enhance the paper-based version of a text book by the automatic insertion of semantic links. Effectively the Hypertext Book project aimed to developing a tool for the automatic construction of hyper-textbooks from text books (Landoni, M., 1997). In particular, the book pages are the unit for reference for their hyper-textbook. The user can browse the book sequentially, going from one page to the next or to the previous one. The relevance relationships from Subject Index terms to pages have to be preserved because the Subject Index is the set of terms the author used to index the textbook, and therefore is the most
important structure for browsing. The Table of Contents has to be presented and augmented, since it is one of the most used features of the paper textbook (Egan et al., 1989). Digital annotations: it is a Formal Model and its application. Marshall, C. C., (1997) studied personal annotative practices of American College students to point out the form the annotations have in the text books and function of the annotations, which is derived from their form. Marshall (1997) (pp. 237-238) discovered that: First, annotations are procedural signals, cluing in the student to where an assignment starts, what material is important, and what material might require a second (or successive readings). Second, annotations are place-marks: they hold the quotes that are being reserved for the paper that the student will write at the end of the term, the chemical reactions and term definitions the student must memorize for the final, the theorem that is key to the proof in the homework assignment. Third, they are in situ way of working problems. Fourth, annotations record interpretive activity, either from another reader (e.g. a professors explanation), or as the result of careful reading (the student has interpreted him/her). Fifth, and most elusively, these markings act as a visible trace of a reader’s attention, a focus on the passing words, and a marker of all that has already been read (as if these words are now processed). Finally, the markings may just be incidental reflecting the material circumstance of reading. Marshall, C. C., (1998) carried on her research work and categorized annotations along several dimensions which reflect the form annotations may take on: formal versus informal annotations, explicit versus tacit annotations as writing versus annotations as reading, hyper-extensive versus intensive annotations, permanent versus transient annotations, published versus private annotations. Finally, Marshall (2002, 2004), and Shipman et al. (2003) investigated
the relationship between private, shared and public annotations and how they can be exploited to find useful passages in the text. Recently, Hwang et al. (2007) conducted a study on the impact of annotations in improving the learning achievements of students. Four-month experimentation was performed, where the learning achievements of students who did not use annotations and students who used a Web-based annotation system for learning material were compared. By using a questionnaire, they found that most of the students agree that the annotation system improved their online reading performances and was easy to use; furthermore, using the annotation system improved the interaction between learners and the provided materials, by increasing students' interest in learning; finally, students reported that the possibility of sharing annotations both between groups and publicly improved their motivation to learn.
REFERENCES


CHAPTER 3 Sec2
SECTION 2

INFORMATION SEEKING
3.2.1) INFORMATION SEEKING MECHANISM

Marchionini (1989a) defined information seeking as a special case of problem solving where learners recognize and interpret in an information problem, establish a plan of search, conduct the search, evaluate the results, and if necessary, iterate through the process again. Kuhlthau (1993) suggested a conceptualization beyond merely seeking and gathering information to a more rigorous constructive process of using information to solve the problem with initiated the information need. Fine (1984) suggested individual engage in information seeking for a number of reasons: to reduce ambiguity; increase their ability to cope with situations, to make decisions, to locate information that will lessen their anxiety, or more toward a desired goal. In a sense, the search event involves problem solving, not simply information finding. It is comparing student activity to the research behaviors of scientists. McNally & Kuhlthau (1994) described how information seeking consists of both undirected and highly directed activities - undirected searching that leads to unexpected links, or discrepant events related to their topics, and highly directed searching for the purpose of finding specific information. These activities formed the basis for students move through predictable stages as they are engaged in these information-seeking activities. They claimed students’ progress from ambiguity to clarity, from seeking general information to seeking specific information. Kuhlthau (1993) described how learners commonly experience a series of phases as they seek information over extended periods. An affective, cognitive, and physical description of these phases includes:

1. Initially, the learner first becomes aware of a lack of knowledge or understanding, and recognizes a need for information. Learners often
experience feelings of uncertainty, and thoughts are often vague and ambiguous regarding the information problem.

2. Selection: The learners' task is to identify and select a general topic for investigating and develop an approach for pursuing the information. After a topic is selected, learners often experience a brief sense of optimism replacing the previous feelings of uncertainty.

3. Exploration: This phase is often the most difficult for learners, as the task is to investigate information on a specific problem or question in order to extend personal understanding. Learner thoughts center on becoming oriented and sufficient informed about the topic in an effort to begin formation of a focus or personal viewpoint. Learners often experience feelings of confusion, uncertainty, and doubt during this phase. Specifically, learners may experience two forms of uncertainty: conceptual – what information the learner is looking for, and one related to the technical process of information retrieval – how to make use of the search instruments (i.e. databases, search tools, interface options, commands).

4. Formulation: The task in this phase is to form a clear focus from information encountered during exploration. During this phase, feelings of uncertainty diminish and confidence increases, as learners perceive this stage as a turning point in their process of information seeking.

5. Collection: In this phase, the learner and system interact most effectively and efficiently.
6. **Presentation:** The central task at this phase is to complete the search and resolve the problem or question. Learners complete the search with a gain in personal understanding.

Given the fact that digital objects carry information in digital libraries and their respective metadata specifications, the proposed dimensions of quality for these two concepts can be connected to the life cycle of information in digital libraries (Borgman, 1996). Such connections can be used to determine when and where quality indicators can be measured, accessed, and improved – as well, as how possible quality problems can be prevented, detected, and eliminated. The connections are shown in Figure (3.2.1.1). The life cycle (see inner portion) has four major phases: information creation, distribution, seeking, and utilization. The outer arrows show in which stage information is active, semi-active, or inactive with regard to the phases. Each major phase (see inner ring) is connected to a number of activities or services. Finally, (see outer ring), each dimension of quality is associated with a corresponding set of activities/services. Similarity to other information, digital objects, or versions can be accessed at time of creation and modification. Preservability and timeliness (in relation to modifications) also are related to this phase.
Figure (3.2.1.1). Information life cycle (Borgman (1996)) with respective dimensions of quality added for each major phase and related activities.

The next sub phase in the cycle deals with metadata creation and information organization and description; therefore, all quality dimensions associated with metadata specifications are located here. Special metadata quality processes such as enforcing filling out of mandatory fields and use of specific formats (e.g., for dates) as well schema validations, should be applied to related activities to guarantee quality (namely, accuracy, completeness, and conformance). In the next phase of archiving and distribution, the aspects of accessibility (e.g., means of storage, format, position in an organizational scheme, etc.) should be taken into consideration. In the seeking (e.g., searching) phase,
relevance of information as returned by the several information satisfaction services can be measured. Finally, most of the dimensions associated with the perceived value of the information (pertinence, significance) can be assessed during the utilization phase.

3.2.2) RELATIONSHIP BETWEEN LEARNING AND INFORMATION SEEKING

Information seeking and learning processes cannot be separated (Cole, 1999; Marchionini, 1995). Information seeking considered a response to some problematic tasks or situations characterized by uncertainty. The resolution of this uncertainty constitutes a change in the learner's knowledge state (Ford, 2003) that leads to learning. Under different information seeking scenarios, the change in knowledge state can be attributed to different modes of searching and browsing (Choo, Detlor, & Turnbull, 2000). In contrast, different conceptions of learning exist. Learning can be seen as problem solving, as inquiry, as sense-making, as intellectual socialization, as design and as constructivist activity by which the learner builds his/her own knowledge (Smith, 1993). Inherent in all these conceptions of learning are changes in the learner's state of knowledge, which results in learning. The interactive process of information gathering and processing, during an inquiry process, comes to an end when the learner can explain a phenomenon or when the learner achieves a state of certainty about some posed question. A careful comparison of the goals and mechanisms of information seeking and learning reveals their similarities and inseparability. Limberg (1999) had also reported a significant interaction between the process of information seeking and learning. Under
the light of the basic mechanisms involved and the evidence provided by the literature, we can see that both information seeking and learning alter the learner's knowledge state and that clearly, information seeking is encompassed by learning, as also is shown in Figure (3.2.2.1.) bellow:

![Learning and Information Seeking](image)

Figure (3.2.2.1). Relationship between information seeking and learning (Faisal Ahmed, 2004)

This is also evident from proposed information seeking and learning models (e.g., (Kuhlthau, 1993), (Choo et al., 2000), and (Ellis, 1993)). It is clear from the discussion presented in this section that information seeking addressed several aspects of learning at a fine granularity. The learning process is implicit in searching, but a learner can greatly benefit by explicit support for learning according to different learning conceptualizations (Faisal Ahmed, 2004).

3.2.3) RELATIONSHIP BETWEEN DIGITAL LIBRARY, INFORMATION SEEKING, AND LEARNING

Learning is the process by which we acquire and retain attitudes, knowledge, understanding, skills and capabilities that are not a part of inherited behavior patterns or physical growth. It deals with an act, process or experience of gaining knowledge or skill through exposure to a variety of resources during study. Learning is a way of interacting
with the world. Kolb (1984) argued that defining learning in terms of the change of behavior is not enough to define the learning process. Kolb defined learning as an adaptation process. "It is a process whereby knowledge is created through transformation of experience". Sue Roberts (2006), discussed the deep learning approaches in four directions that are first, a learning as active, situated, and social; second, constructivist perspectives; third, learner-focused pedagogy; and finally, developing metacognitive & other skills. He also described the role developments with three elements academic, learning technologist, and academic librarian. The important question is that how do libraries support teaching and learning?

To answer this question we should discuss the roles of digital library in learning. According to Cory Marchionini and Hermann Maurer (1995) who indicated that libraries serve at least three roles in learning: first, they play a practical role in sharing expensive resources. A community of users shares physical resources such as books and periodicals, films and videos, and software and electronic databases, and specialized tools, such as projects, graphics equipment, and cameras. Human resources – librarians (also called media specialists or information specialists) – support instructional program by responding to the requests of teachers and students (respective services) and by initiating activities for teachers and students (proactive services). Responsive services include maintaining reserve materials, answering reference questions, providing bibliographic instruction, developing media packages, recommending books or films, and teaching users how to use materials. Proactive services include selectively disseminating information to faculty and students, initiating thematic events, collaborating with instructional to plan instruction, and introducing new instructional methods and tools. In
these ways, libraries serve to allow instructors and students to share expensive materials and expertise. Second, libraries serve a cultural role in preserving and organizing artifacts and ideas. Great works of literature, art, and science must be presented and made accessible to future learners. Although libraries have traditionally been viewed as facilities for printed artifacts, primary and secondary school libraries often also serve as museums and laboratories. Libraries preserve objects through careful storage procedures, policies of borrowing and use, and repair and maintenance as needed. In addition to preservation, libraries ensure access to materials through indexes, catalogs, and other aids that allow learners to locate items appropriate to their needs. Third, libraries serve social and intellectual roles by bringing together people and ideas. This is distinct from the practice role of sharing resources in that libraries provide a physical place for teachers and libraries to meet outside the structure of the classroom, thus allowing people with different perspective to interact in a "knowledge space" that is both larger and more general than that shared by any single discipline or affinity group. Browsing through a catalog in a library provides a global view for people engaged in specialized study and offers opportunities for serendipitous insights alternative views. In many respects, libraries serve a centers of interdisciplinary – place shared by learners from all disciplines. Digital libraries extend such interdisciplinary by making diverse information resources available beyond the physical space shared by groups of learners. Hence, Cory Marchionini and Hermann Maurer (1995) denoted to one of the greatest benefits of digital libraries is bringing together people with formal, informal, and professional learning missions as follows: First, formal learning is systematic and guided by instruction. It takes place in courses offered at schools of various kinds and training
courses or programs on the job. The important roles libraries serve in formal learning are illustrated by their physical prominence on university campuses and the number of courses that make direct use of library services and materials.

![Diagram of Learning Types](image)

Figure (3.2.3.1). Current model of technological support for different types of learning

(Cory Marchionini and Hermann Maurer (1995))

Second, digital libraries open new learning opportunities for global rather than just local communities. Learners take advantage of other people, mass media, and their immediate environment during informal learning. Third, professional learning refers to the ongoing learning adults engage in to do their work and improve their work-related knowledge and skills. The main information resources for the professional learning, however, are personal collections of books, reports, and files; subscriptions to journals, and the human networks of colleagues nurtured through professional meetings and various
communications. Many of the data sets and computational tools of digital libraries were originally developed to enhance professional learning. Digital libraries combine technology and information resources to allow remote access, breaking down the physical barriers between resources. Although these resources will remain specialized to meet the needs of specific communities of learners, digital libraries will allow teachers and students to take advantage of wider range of materials and communication with people outside the formal learning environment. This will allow more integration of different types of learning, as depicted in Figure (3.2.3.2) as shown below:

Figure (3.2.3.2). Digital libraries lead to integrated resources and types of learning (Cory Marchionini and Hermann Maurer (199515))
Digital libraries will allow learners of all types to share resources, time and energy, and expertise to their mutual benefit. For example:

a-) textual databases of classic words and image collections for important artistic exhibits or museums have been assembled by scholars and made available through the Internet (Santa and Calif, 1994).

b-) although electronic journals are becoming more common, they have not penetration as many expected (Oldzyko, 1994; Schaffner, A., 1994). Two common approaches to electronic journals are: to store files in LaTeX, PostScript or ASCII form in a fileserver and email the files or allow FTP access to them (generic approach); and to store documents in hypertext systems and allow on-line browsing and perusal (hypertext approach).

c-) new reading and filtering programs (Ahlberg et al., 1994; Stevens, C., 1993) and search tools such as Archie and Veronica (Deutsch, P., 1992) provide rudimentary aids for locating information in these electronic discussions.

d-) a variety of hypermedia materials are becoming available, and these collections are often served from a library rather than through dedicated machines in classroom (Crane, G., 1994; Mylonas, E., 1992).

Furthermore, learning and digital library have two distinct components: Learning related to accessing evaluating and using the information resources available in this environment; and learning related to mastering and building upon the ideas embodied within those individual resources (Della Neuman, 1997). Explorations of the relationship of cognitive theory to instructional technology have in fact co-existed with behaviorist ideas in the instructional technology community for years, and the field has a
long history of drawing on both traditions to explore the relationship of media and learning. For example, Kozma (1994) offered his conclusions and insights as argument of reframing instructional technology's longstanding debate on the contributions of media to learning:

"Perhaps we should ask, what are the actual and potential relationship between media and learning? Can we describe and understand this relationship? Moreover, can we create a strong and compelling influence of media on learning through improves theories, through improves research, and through improves instructional designs?" (Kozma, 1994, p.233)

Liebscher and Marchionini (1988), Neuman (1993, 1995) and Solomon (1993, 1994) all suggested important capabilities that students must acquire to use text-based electronic information resource successfully; Large et al. (1994a, 1994b), Perzylo and Oliver (1992), and Small and Ferreira (1994) performed the same function for multimedia materials.

One of the enduring works in instructional technology was Malcolm Fleming and Howard Levie's (1978) Instructional Message Design: Principles from the behavioral Science – re-issued in 1993 as Instructional Message Design: Principles from the Behavioral and Cognitive Sciences. The 1993 edition combined findings from both traditions of learning theory to specify over 300 research-based guidelines for designing instructional presentations in various print and non-print formats – and over 200 of these address issues that can be applied to creating components of digital library. Even such simple principles as "Purely decorative pictures should be used sparingly" (Fleming & Levie, 1993, p.89) and "Logically organized text is better remembered than poorly organized text" (p.208) provided useful insights about how components of the digital
library might be organized and presented to increase students' likelihood of learning in this environment. More complex principles – such as "The acquisition of unfamiliar content can be improved via familiar examples, analogies, and metaphors, while such strategies are less essential for familiar content" (Fleming & Levie, 1993, p.215) and "The presentation of visually richer and more realistic best examples leads to a richer and better consolidated prototype resulting in increased transfer" (p.244). "Macro-level environments" include both rich collections of resources and tools students use to explore them "to pursue interests or needs beyond the parameters typically provided in isolated lessons" (Hannafin, 1992, p.58); "Micro-level environments" offered similar arrays of materials but focus within more discrete domains. "Mathemagenic environments" supported access to various representations of content in a particular area (often through hypermedia links) and allowed students to "move rapidly among networks of concepts (and) to construct their own sets of relationships within the network" (p.59). Cora Busstra (2008) discussed five principles for the design of digital learning material for academic education in human nutrition. These principles were as follow:

**Principle 1: Motivate the students.**

Digital learning material provides sufficient opportunities to implement these principles and to motivate the student.

**Principle 2: Authentic learning content.**

Active construction of knowledge can be supported by providing meaningful, realistic and authentic learning contexts, which reflect the way knowledge is used in "real-life" (Brown JS, Collins A, Duguid P., 1998; Honebein PC, Duffy TM, Fishman BJ, 1993). Digital learning material provides many opportunities to provide an
authentic learning context. For example, digital learning material can be used to provide a simulated virtual laboratory environment in which the student can freely design and try out experiments (Wilmsen T., Hartog R., Biseeling T., 2002; Diederen J., Gruppen H., Hartog R., Voragen GJ., 2006).

Principle 3: Active Learning.

Active learning and practice is necessary for strengthening understanding, acquiring knowledge and retention knowledge (Anderson JR., 2000; Sweller J., Van Merriënboer JIG., Paas FGWC., 1998). Especially digital learning material provides many opportunities to engage each student individually in studying. With digital learning material a broad range of interactive exercises (like drag and drop questions, multiple-choice questions etc.) and other activities (e.g., interactive simulations, interactive practice possibilities for data analysis) can be provided that will stimulate the student to learn actively.

Principle 4: Visualization of important concepts.

An important benefit of digital learning material is the possibility to use dynamic visuals (e.g., interactive diagrams, animations and video clips) which are more elaborated than figures in textbooks (Lawalter D., 2003; Mayer RE., 2002). Mayer discussed that "the promise of multimedia learning is that students can learn more deeply from well-designed multimedia messages consisting of words and pictures than from more traditional modes of communication involving words alone" (Mayer RE., 2002). Furthermore, it is suggested that well-designed images or diagrams can improve understanding and retention of knowledge (Sweller J., Van Merriënboer JIG., Paas FGWC., 1998; Larkin JH., Simon HA., 1987; Schnotz W., 2002). With respect to
visualization, the challenge during the design of the digital learning material is to investigate which representations (such as animations, schemes, pictures etc.) are adequate to clarify a certain rule, concept or principle.

**Principle 5: Reduce unnecessary cognitive load.**

According to cognitive load theory, an individual's cognitive capacity is limited. There is a certain amount of information, which a student can process at a certain time (Sweller J., Van Merriënboer JJG, Paas FGWC, 1998; Baddeley AD, 1992; Kirschner PA, 2002). Digital learning material provides several opportunities to prevent cognitive overload. One principle is the use of Just-In-Time (JIT) information presentation (Kaster L, Kirschner PA, Van Merriënboer JJG, Baumer A, 2001). This means providing the student with information and feedback at exactly the moment he needs this information to perform a task.

Goodrum, Dorsey & Schwen (1993) described the conceptual and practical difficulties in designing an "enriched learning and information environment" that accommodates the difficulties that Perkins (1991) had identified for students operating with such a setting: high cognitive load, increased responsibility for managing their own learning, and need to adopt an unfamiliar learning process. Scardamalia and her colleagues (1989, 1992) had worked for years on the development and refinement of CSILE – “Computer-Supported International Learning Environment” – a shell that allows students to create their own knowledge base related to classroom instruction by working collaboratively in an electronic environment to generate hypotheses, ask questions, and revise their understanding of that information. Rieber (1990, 1996) had explored how both animation and elements of simulations and games can enhance students' abilities to
focus on and learn from multimedia "micro worlds". Hannafin, Hall, Land, and Hill (1994) noted the lack of compelling empirical evidence of how open-ended learning environment influence learning and, further, discussed the difficulty of obtaining such evidence. Hannafin and his group had compiled a set of empirically based guidelines for designing interactive multimedia that might provide some insights into how components of the digital library might be designed to enhance teaching (Park & Hannafin, 1993).

M. A. Gopinath (1996) indicated to the curriculum design that the organization of learning experiences for library and information professionals involves vertical and horizontal relationships. The vertical relationship among learning experiences is worth reference to time — the first semester, second semester etc. The horizontal relationship is in terms of one area to another, for instance, the first semester talks of information and its usage; the second semester on system technology and the third semester combines both in terms of human computer interface for information access and assimilation. Sridhar (1996) discussed the impact of electronic of libraries/media on education that aid all types of education — formal, informal, adult, vocational, distance and higher education (p.236). Natalie and San Pang (2005) discussed specifically at the process of learning between peers in a group and how digital libraries can lend themselves as a learning environment towards this purpose. Natalie and San Pang according to Tosey & Gregory (1998) identified five characteristics, providing a contextual design for peer learning. These five characteristics are Personal Development, Community Interaction, Facilitation, Formal Independence, and Boundary Management. Natalie and San Pang concluded that digital libraries could be implemented as learning environments —
involving resources, services, tools, and a profiled community of users. Digital libraries are potentially highly interactive environments encouraging electronic learning between its users. Logue (2003) discussed that the role of libraries in providing instructional supports has changed to one of a more dynamic nature in recent years – developing new digital resources and services for the purpose of distance learning, teaching aids, or support for the academic faculties in students' research and coursework. Microsoft (2003) in Singapore, many digital libraries have evolved in education. One of the most significant implementations is IVLE (Integrated Virtual Learning Environment); developed initially by the National University of Singapore. Jose et al. (2002), Kearsley (2000) with the proliferation of digital libraries in education, institutions and schools are realizing the benefits of digital libraries to provide new opportunities for learning activities. Marliom et al. (2002), Mayer and Moreno (2002) and Rapp et al. cognitive psychologists, instructional designers, and others coming from an educational research background often highlight the role that digital libraries can play as cognitive tools to support the learning and sense-making activities of individual users. This view focuses on how people can make use of multimedia resources in digital libraries to construct their own knowledge representation, and typically draws on cognitive theories such as constructivist learning or those taking a human-interaction processing perspective. MacDonald et al. (1999) denoted to that in distance learning literature, this is often referred to as 'resource-based learning' which strives to offer learners choices in their learning materials and to accommodate individual differences through the provision of a wide selection of typically related, multimedia learning resources. As such, resource-based learning requires learners to grapple with both the topic of study and information
seeking and information handling skills. Renninger & Shumar (2002) discussed how specific Math Forum services, and participation in the broader Math Forum Community, helped teachers to reconceptualize themselves as linking mathematics rather than being math phobic, and thus increasing their own personal skills in mathematics and changing the way they taught math in their classrooms. M. A. Goncalves (2004) discussed the 5S framework that Fox and his students at Virginia Tech have been developing a formal model of DLs based on Streams, Structures, Spaces, Scenarios, and Societies, hereafter related to as “5S”. “Streams” described all types of content as well as communications. “Structures” described organizational schemes, including data structures, database, and knowledge representations. “Spaces” cover 2D and 3D interfaces, GIS data, and presentations of documents and queries. “Scenarios” are specified as system states and events. “Societies” describe both software “service managers” and generic “actors” who may be human users or machine processes, or collaborations of one or more of both (see table 3.2.3.1).
### Table (3.2.3.1). 5S Framework (Goncalves, 2004)

<table>
<thead>
<tr>
<th><strong>5S</strong></th>
<th><strong>Examples</strong></th>
<th><strong>Formalization</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Streams</strong></td>
<td>Text, audio, video, image</td>
<td>Sequence (list)</td>
</tr>
<tr>
<td><strong>Structures</strong></td>
<td>Collection, catalog, hypertext, document, metadata, taxonomy</td>
<td>Graph, Function, Relation</td>
</tr>
<tr>
<td><strong>Spaces</strong></td>
<td>Used in indexing, browsing, and searching services- as well as interfaces</td>
<td>Set(vector, topological, measurable, measure, probability spaces)</td>
</tr>
<tr>
<td><strong>Scenarios</strong></td>
<td>Searching, browsing recommending</td>
<td>States, events, sequences (lists)</td>
</tr>
<tr>
<td><strong>Societies</strong></td>
<td>Service manages (software), Actors (learners, teachers etc.)</td>
<td>Tuple (relating events and actions)</td>
</tr>
</tbody>
</table>

Borgman et al. (2000) observed the use of the Alexandria Digital Library Prototype in college-level geography classrooms and labs by both teaching faculty and students. Janee & Frew (2002) the Alexandria Digital Library provides an innovative resource discovery service enabling users to retrieve geo-referenced maps and images based on geographic location. Abbas et al. (2002) at the K-12 level examined the use of the Artemis Digital Library by middle school students in 32 schools.
3.2.4) SCAFFOLDING AND PERSONALIZATION IN DIGITAL LIBRARIES

Nowadays, research in information seeking has shed light on the complexity of the process, the different stages embedded within the process, and problems faced by the users during information seeking (Hirsh, 1997; Shenton & Dixon, 2004; Fidel et al., 1999; Ford, 2003; Choo et al., 2000; Shenton & Dixon, 2003; Wilson, 1999; Watson, 1998). Two prominent research directions in the digital library community have addressed this concern: providing task support, and personalizing interactions. Task support, to increase the effectiveness of information retrieval systems, researchers has realized the importance of providing task-specific support to users (Jarvelin & Ingwersen, 2004). Therefore, a number of studies have appeared such as, the work by Vakkari (2000, 2001) investigated the information needs associated with each stage in writing a master's research proposal. Vakkari found that during the three tasks the students used stages (pre-focus, focus formations, and post-focus) different types of information. Therefore, a number of digital library systems have emerged embed tools and processes to support task specific performances. ARTEMIS used a learner-centered design (Soloway, Guzdial, & Hay, 1994) to scaffold the inquiry-learning process via information seeking (Wallace et al., 1998; June Abbas & Soloway, 2002). In designing the ARTEMIS digital library the authors first identified the difficult-to-do task stages and then coupled these stages with appropriate tool support. In summary, ARTEMIS provides task structuring tools, guiding prompts, and reflection tools to scaffold the inquiry-learning process. A similar approach has been adopted by the science, mathematics, engineering and technology (SMETE) digital library (Andy & Alice M., 2001). In
designing SMETE, a deep understanding of the learner and educator's work is used to identify four key areas of support within a digital library interface, which are information organization, and information search. Personalization, the DELOS/NSF Working Group on Personalization and Recommender System for digital libraries has suggested that digital libraries lacking individual and community-based personalization capabilities might be considered as failing their constituencies (A. Smeaton & Callan, 2002). In a synthesis of discussions of this working group, Smeaton & Callan had proposed that the future of personalization in digital libraries includes the following research agenda challenges: richer user model, innovative personalization and recommender algorithms, more flexible user interactions, improved evaluation methods, and increase focus on the social effects enabled by digital libraries (A. F. Smeeton & Callan, 2005). Therefore, one of the important factors in this research is the interaction and interface personalization that it will discuss later in this section. Interaction modeling within digital library environments can provide a way to highlight usability issues and identify points of support where personalization can be most effective. For example, in the Corporate Digital Library (CDL) (Costabile, Esposito, Semeraro, & Fanizzi, 1999), user interaction logs are collected and classified into novice, expert, and teacher. DAFFODIL is another example of a system that makes recommendations based on analyzing user interaction paths (Fuhr et al., 2002). In addition to system development, empirical studies have been performed to define a generic interaction framework for digital libraries. Byran-Kinns et al. (2000) defined a five level interaction framework for analyzing and designing digital libraries, which are Properties, Causes, Symptoms, Traces, and Fundamentals. Interestingly this interaction framework can be used to
decompose the popular four levels of search activities in information seeking environment (Bates, 1990), which are level1 (move), level2 (tactic), level3 (stratagem), and level4 (strategy). Faisal Ahmed (2004) noted that exploratory learning environments reflect the conviction that students should be responsible for their own learning and should be empowered to take control of their learning processes. Under these philosophical commitments, learning environments and digital libraries approximate human scaffolding. Therefore, one of the basic issues with approaches to enabling learning with digital libraries is that learning is considered a process separate from information seeking. Due to this underlying assumption, Shallow user models that provide superficial personalization are mostly used in digital library personalization research. Recent literature has realized this problem and called for deeper user modeling for providing adaptive support in digital libraries (Fries-Martine, Magoulas, Chen, & Macredie, 2006). Therefore, Jayawardana et al. (2001) discussed that personalized information environment (PIE) in a digital library is a framework that provides a set of integrated tools based on individual user's requirements and interests with respect to his access to library materials. These tools can support active learning by integrating the user's personal library and a remote digital library. The user will be able to carry out learning activities when browsing the digital library. There are two schemas involved in PIE. Material personalization corresponds to facilities for learners to use library materials according to their individual requirements such as active consuming and information gathering. Collection personalization, on the other hand, captures the learner's content and interest from material personalization in order to provide a personalized view of the organization retrieving. In PIE, these two schemas of personalization benefit each other
by creating the cycle of interaction. Therefore, it is possible the three main active learning facilities in the PIE for the digital library.

In this architecture, there are two main components called DL Browser and Personal Document Editor. DL Browser supports the collection personalization and Personal Document Editor supports material personalization. Two information-seeking tools, which are personalized retrieving and personalized filtering, are provided in the DL Browser. On the other hand, tools for active consuming and information gathering are available in the Personal Document Editor. The interaction of both the DL Browser and the Personal Document Editor provide an interface for the digital library. Jayawardana et al. (2000, 77-84) indicated that the personal documents are categorized according to the following areas: Personalized Books, Personalized Articles, Personalized Video/Audio, and Notebooks. Personalized Books and Personalized Articles support the learner's active reading process, providing personalized view of static library materials. Personalized Video/Audio supports active watching and active listening processes and gives personalized views of audio-visual library materials. The Notebook, which represents the learner's information gathering, allows the user to integrate different segments that can be text, images, audio or video together with annotations. Manipulating of those personal documents is carried out by personal library functions, which include annotating, segmenting, formatting, modifying, organizing and integrating. Therefore, Shallow Copy is the main technique that is used to create Segments from the existing multimedia library materials. With Shallow Copy, only a pointer to the part to be copied is stored, and no physical copy of the material is made. In personal documents, Shallow copied segments are used instead of simple hyperlinks. Since the internal representation
varies with respect to media type, the design of Shallow Copy technique depends on the media type of the selected source. There are three categories, namely open formats corresponding to all text encoding such as HTML, XML and ASCII; print formats such as Postscript generally contains both text and images together; and Proprietary formats, for example Microsoft Word, usually come with an application programmers interface (API). A Shallow copied segment consists of three attributes: document identifier, being offset address and end offset address. Each segment contains two declarations called physical and logical. The physical declaration of document identifier would be URL or URL of the source object and it becomes invalid when the object is not retrieved or located. Therefore, the contents of the logical declaration will be used to identify it. The offset address and end offset address contain values that can be used to retrieve the segment from the documents. Image objects are also represented in a variety of formats such as JPEG, GIF, and BMP. The internal representation of these image objects vary significantly from each other and the platform (operating environment) also affects their representation. Since it is difficult to have a unique mechanism to present the Shallow copied image segment, we use a virtual coordinate system to map them in order to calculate the offset addresses uniquely and it is maintained in a device independent form. Furthermore, the virtual coordinate approach used to describe the Shallow segments taken from files or print formats. The data type of video/audio data is completely different from text and image types, since the appearance of video/audio data depends on the time. Although there are varieties of data formats for their representation, both the audio and video data are function of time when they are played. Shallow Copy uses time as the user specified parameter to identify segments of audio and video library materials.
Therefore, when opening a personal document using the browser, the dynamic video sources are generated based on the information in the Shallow video clips of the document. Such sources could stay in the temporary buffer during the activation time of the personal document but are not stored. Hewagamage et al. (2001) denoted to that the Shallow Copy technique demonstrates definite advantages in the case of multimedia data types such as audio/video data, which take a large amount of storage space.
REFERENCES


library for Earth system education: building community, building the library.


Reference Library, 44 (1), 53-60.


43(2), 192-201.

INFOSCI.

from instructional systems design. Educational Technology Research and

structures. Unpublished paper presented at the Annual Meeting of the American
Society for International Science, Chicago.

284-298.

traditional scholarly journal. Notice of the AMS.

interactive multimedia. Educational Technology Research and Development, 41
(3), 63-85.

Technology, 31 (9), 19-21.


156


ACM SIGIR Conference on Research and Development in Information Retrieval. Athens, Greece; ACM Press.


