Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 1993 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

December 1993

Medical Multi-media Information Management: A Research Framework

Hong-Mei Garcia
University of Hawaii

Olivia Sheng
University of Arizona

Medical Multi-media Information Management: A Research Framework

Hong-Mei Chen Garcia
Department of Decision Sciences
College of Business Administration
University of Hawaii
Honolulu, HI 96822, USA

Olivia R. Liu Sheng
Department of Management Information Systems
College of Business and Public Administration
University of Arizona
Tucson, AZ 85721, USA

Abstract

The multi-media nature and global scope of PACS give rise to a broad range of medical information management issues. Due to the complexity of each design issue, most existing research projects only work on a single facet of these issues. As novel medical information systems are being rapidly developed, there is a critical need for integrated data management and strategic planning of information system development in hospitals. This paper presents a research framework to relate proliferating medical information management issues, facilitate strategic PACS database development and allow research results to be easily integrated when they become available. The characteristics and requirements of medical multi-media information systems are summarized and design issues encompassed in the framework are examined. Partial research and development results at the University of Arizona based on the proposed framework are discussed.

1 Introduction

Picture Archiving and Communication Systems (PACS) have been actively researched and developed for more than a decade as a means to achieve total computerization of medical image acquisition, processing, storage, transmission, and manipulation [7]. A PACS not only replaces old manual image file management systems but also offers many advantages, such as vast image processing and display capability, concurrency data access capability, and innovative usage for remote consultation, research, and education purposes.

One of the most important aspects of PACS research is the design and implementation of a PACS database system (DBS). A PACS database stores and manages all patient image data and related graphical/textual/voice data in a hospital. It will support daily radiological transactions and also increase diagnostic accuracy by providing easy access to high-quality data and quick search of relevant data. Because of the large volume of image data involved and high system performance required in a clinical environment, design considerations for a PACS database are very different from those for traditional database designs. Moreover, a PACS database must communicate with several already existing databases in a hospital if it is to achieve diagnostic and administrative efficiency and effectiveness. Furthermore, advances in multimedia wideband communications technology encourage a distinct movement toward a Global PACS (GPACS) environment [14]. In a GPACS environment, manual transfer of image data between hospitals is obviated, scarce medical resources and expert knowledge can be shared, and regional/national/international medical teleconferencing/research/education can be facilitated [14].

The multi-media nature and global scope of PACS give rise to a broad range of medical information management issues [9]. Due to the complexity of each design issue, most existing research projects only work on a single facet of these issues. As novel medical information systems are being rapidly developed, a critical need has arisen for integrated data management and strategic planning for information system development in hospitals. Therefore, it is imperative that a research framework be provided to relate proliferating multi-media medical information management issues, facilitate strategic PACS database development and allow research results to be easily integrated when they become available. In the next section, we summarize the characteristics and requirements of medical multi-media information systems based on a 2-year intensive requirement analysis at the University Medical Center at the University of Arizona. In Section 3, we propose a research framework and examine four aspects of the framework: 1) Local PACS (LPACS) core design issues, 2) Global PACS (GPACS) design issues, 3) PACS DBS design enabling technology and 4) PACS DBS implementation enabling technology. In Section 4, we present research and development results addressing some issues encompassed in the proposed framework. Section 5 contains the conclusion of this paper.

2 Medical Multi-media Information: Characteristics and Requirements

As shown in Figure 1, a PACS database system (DBS) which manages integrated radiological data serves many users and handles numerous processes within as well as outside a hospital. The processes of traditional radiological services include reception, exam scheduling, exam taking, and exam reading (interpretation), report transcribing and distribution, as well as image filing and loan requests handling. Referring sections and other hospitals often request patient files for both diagnostic and research/education purposes. Many outside entities such as insurance companies, lawyers, medical quality assurance committees, etc., also request patient information from the radiology department, mostly for legal purposes. All of the above processes center on a radiological data management system (e.g., PACS DBS) and closely interact with Hospital Information Systems (HIS) and Radiology Information Systems (RIS). An HIS is typically a managerial system for centralized control of patient demographic, accounting, and history data. RIS manages cost accounting and various control/utilization statistics reporting for the radiology department. As depicted in Figure 2, many hospital functions (e.g., patient admission/discharge/transfer (ADT), etc.) are across boundaries of these information systems (e.g., PACS, HIS and RIS), while consulting, research, educational and conferencing activities are often conducted with other hospitals. The PACS DBS plays a crucial role in providing these health care functions; therefore, its design needs to address the requirements and considerations arising from these functions.

PACS DBS design requirements and considerations largely fall into the categories of data characteristics, transaction characteristics, communication requirements, performance requirements and implementation considerations. For instance, important data characteristics include: A) **large volume of image data:** a typical digi-

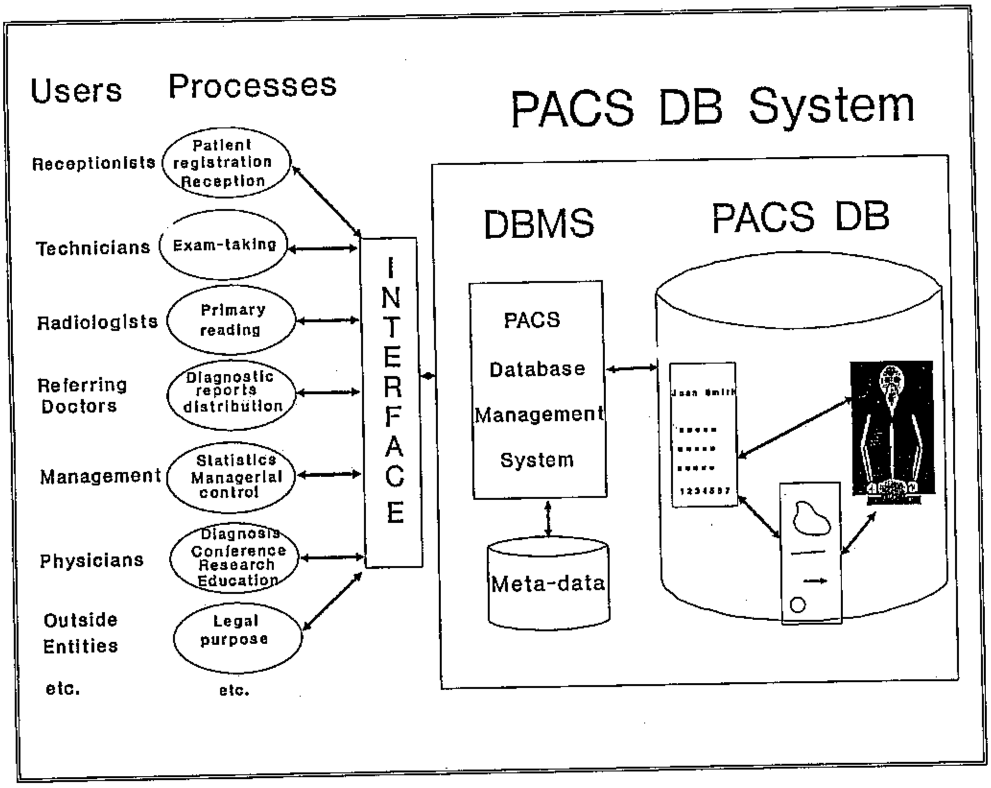

Figure 1: The Overview of PACS DB System

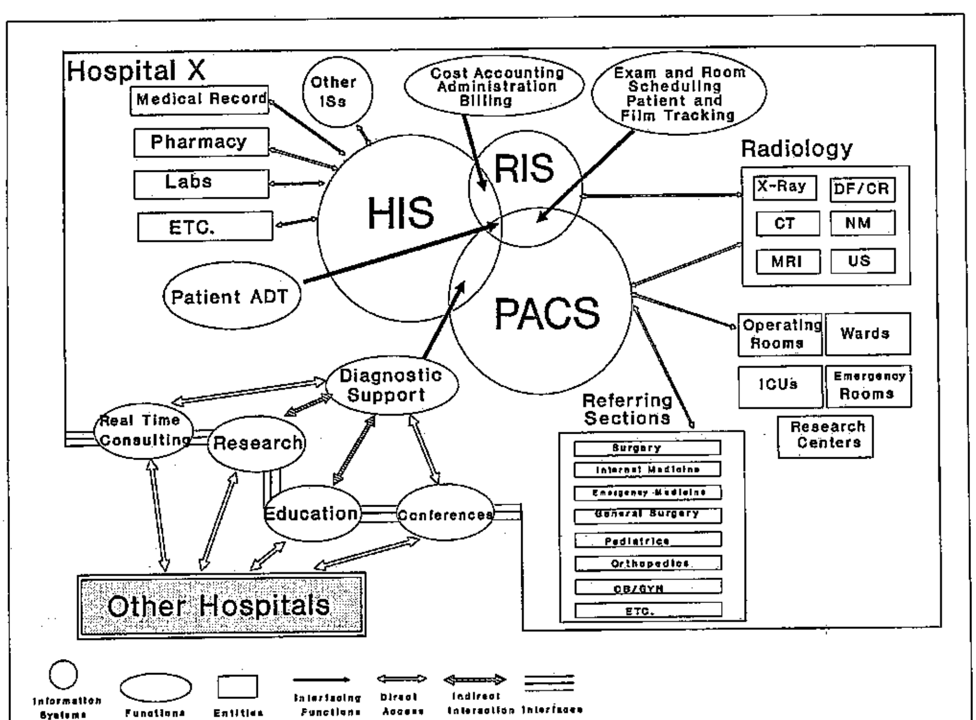

Figure 2: The Schematic View of Totally Digitized Radiology with PACS

tized X-ray image requires from 10 megabits (1024 × 1024 × 12) to more than 200 megabits (4096 × 4096 × 12) of storage space, while the amount of image data generated in a hospital each day can be two, ten or hundred gigabytes. In addition, each state has its own legal requirement for long-term (2 to 30 years) archiving of patient data, amounting to terabytes of data to be handled by a medical image management system; B) **diverse types of image data** produced by different types of imaging equipment (modalities): in general, images contain two-dimensional information of varied sizes. There are three-dimensional spatially-reconstructed images from Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) and dynamic or time-serial images from Cine-angiogram (CINE) or Ultrasound US; C) **close relationships among image, textual and graphic data:** PACS image data retrievals are usually accompanied by textual data or voice-activated input (patient demographics, exam data and diagnostic reports) and graphic annotation (marks such as arrows, circles, lines on films). The specific considerations and requirements in all categories are summarized in Table 1 and detailed in [14].

Category	Item #	Description
Data Characteristics	A	Large volume of image data
	B	Diverse types of image data
	C	Close relationship among image, textual graphic data
Transaction Characteristics	A	Frequent image/text generation and (concurrent) retrievals
	B	Image retrieval by set based on temporal, anatomical and modality relationship
	C	Local interest of image usage
	D	Few deletes and no updates on image data
	E	Frequent ad hoc retrievals for conference, research and education
Communication Requirements	A	Interdependence among PACS, HIS and RIS
	B	Support for diagnosis of complicated cases
	C	Real time consultation
	D	Data/resources sharing with other hospitals
Performance Requirements	A	Real time response and high system throughput
	B	Patient confidentiality, data availability system reliability and security
Implementation Considerations	A	Sectional hospital budget structure
	B	Projected growth

Table 1: PACS DBS Characteristics and Requirements

3 The Proposed Research Framework and PACS DBS Design Issues

PACS DBS characteristics and requirements summarized above present a range of issues and considerations to be taken into account when adopting an approach to PACS DBS design. Table 2 outlines the PACS design issues and their underlying requirements and considerations.

PACS DBS Design Issues	Characteristics and Requirements (see Table 1)															
	Data			Transaction					Comm.				Perf.		Implem.	
	A	B	C	A	B	C	D	E	A	B	C	D	A	B	A	B
LPACS Core DB Design Issues:																
Distributed PACS DB	X			X	X	X	X				X		X	X	X	X
PACS/RIS/HIS interoperability			X						X	X	X					
Physical image data storage	X	X	X	X	X		X	X					X	X		
AI for PACS	X			X	X			X	X	X			X			
DB machine	X			X	X								X			
Data security														X		
GPACS Data Management										X	X	X		X		

Table 2: Design Issues Driven by PACS DBS Characteristics and Requirements

While a few research projects have been undertaken to address PACS DBS design issues, these studies have primarily worked on a single facet of these issues due to the complexity of each design issue. From a strategic standpoint, we formulated a research framework so that multi-faceted PACS DBS design issues, either with a local or global scope, can be addressed in parallel and research results can be easily integrated when available.

Figure 3 depicts a layered research framework in which the inner layers are critical research issues central to the development of technologies that will address the design issues in the outer layers. The four layers are: PACS DBS Design Enabling Technology, LPACS Core Design Issues, GPACS Design Issues and PACS DBS Implementation Enabling Technology. The initial focus on local PACS (LPACS) design issues is appropriate since results obtained there will provide ways to extend PACS technology onto a global environment. However, the development of effective data models and efficient query support for PACS data management will be essential to allow fruitful development and easy integration of technologies that respond to the individual LPACS DBS core design issues. Under this framework, the technologies developed for PACS design issues are then transferred to PACS implementation support mechanisms, including specialized PACS DBS development (DBMS) environments, tools for generating customized designs and standards that enable fast modular implementations and industry-wide interoperability. While PACS DBS R&D should be conducted in the outward direction, feedback mechanisms also are needed between layers to allow insights into technological improvements to flow inward so that the framework will improve along with the PACS R&D.

In the following, we will first discuss the LPACS and GPACS core design issues and then discuss the PACS DBS design and implementation enabling technology isssuses.

3.1 LPACS Core Design Issues

A local PACS (LPACS) is situated in a single hospital and its core DB design issues are concerned with efficiently and effectively providing all the information services required by the radiology service in the hospital. The following discussion examines LPACS design issues and justifies the design approaches adopted based on the data characteristics and performance requirements identified in Section 2.

3.1.1 Distributed PACS DB design

The heavy workload associated with handling large amounts of image and related data in a centralized PACS DBS tends to result in severe system bottlenecks. By capitalizing on the simple concepts of resource sharing and load balancing, distributed PACS DB design, which is characterized by physical separation/replication and logical integration of data among multiple PACS archive and storage facilities, offers system performance and dependability improvement. In particular, by distributing and possibly replicating image and related data at several locations, most retrievals can be handled locally without experiencing communication delay and overhead, and thus system bottleneck can be alleviated. Under such a distributed design, a PACS DBS is also more dependable in that it can continue information service in the face of failures of individual sites or partial communication links.

In addition to performance considerations, a distributed design promises to meet other PACS DBS requirements. First, the hospital budget's sectional structure implies that a modular DB design is needed for the PACS DBS. Modularity is the essence of a

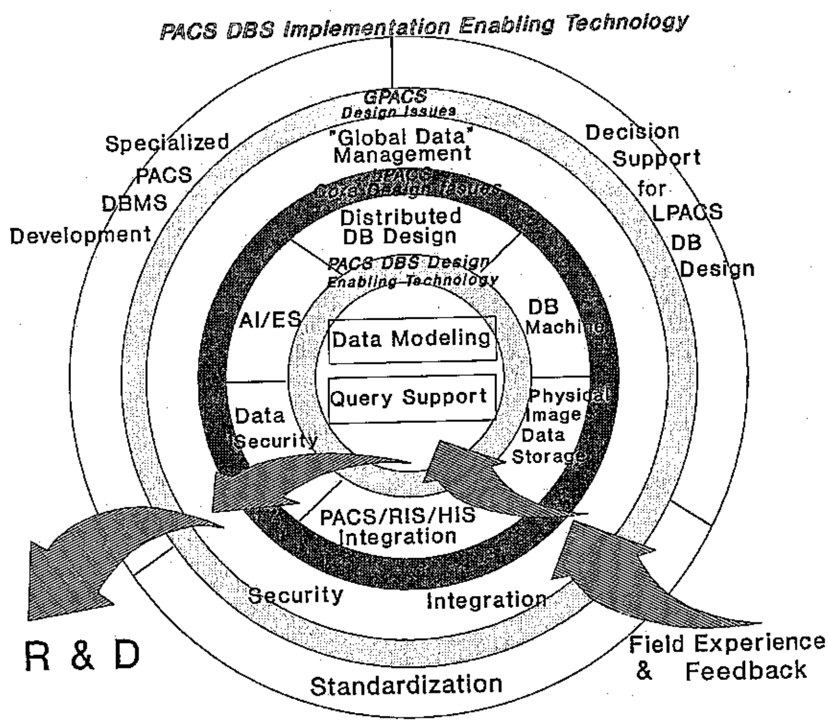

Figure 3: A Research Framework for PACS DBS R&D

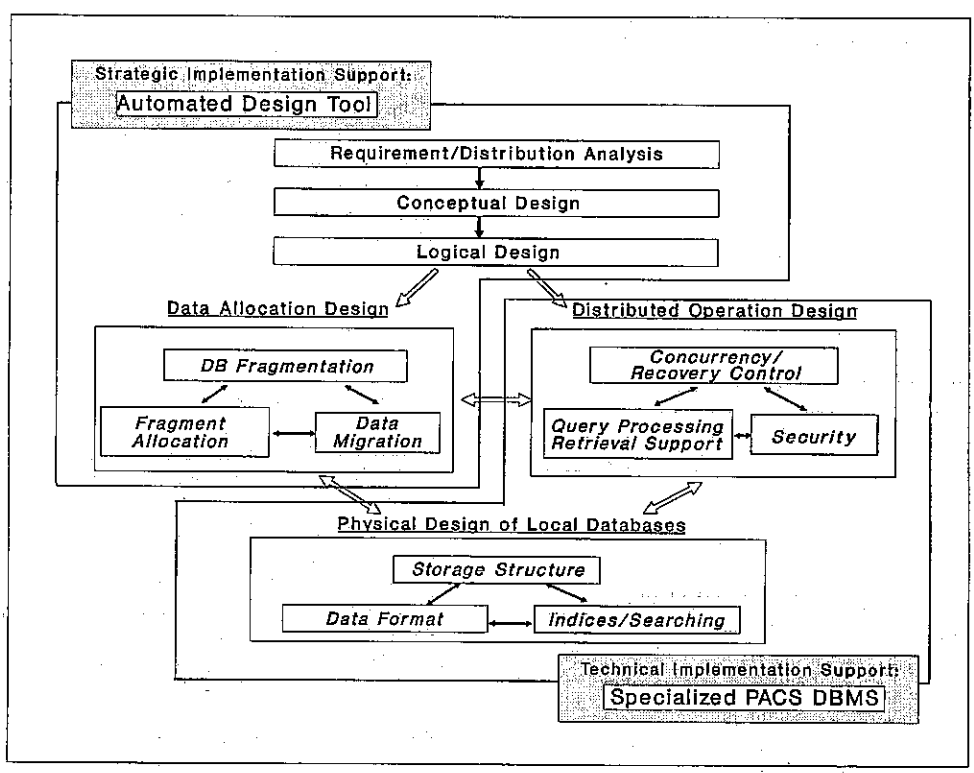

Figure 4: Local PACS DBS Development

distributed DB design. A DDB is composed of geographically separated sites which enjoy local autonomy but are coordinated to maintain system integrity. This feature allows new system components to be added on easily, meaning system-wide or local growth in capacity can be smoothly achieved. Moreover, cost economy is gained since the DDBS can make use of local workstations and their storage systems. Second, most of the image data generated are often later retrieved by the same requesting hospital section. This data usage pattern is ideal for a distributed design. Third, the fact that few update/delete data operations are required significantly reduces both the complexity of the distribution design and the distributed operation processing overhead [15, 11].

However, like other distributed systems, a distributed PACS DBS is confronted with complicated design problems such as allocation and capacity design of system resources (storage, data and processors) as well as database management (distribution transparency, concurrency/recovery control, query processing and security support) and transaction management (decomposition, routing and integration) problems. The difficulty in solving these design problems is further complicated by the size and the complexity of PACS [11].

3.1.2 PACS/HIS/RIS interoperability: heterogeneous DB integration

The need to integrate PACS, HIS and RIS exemplifies the decades-old information management problem in hospitals. In general, the existing heterogeneous databases (e.g. PACS, RIS, HIS, etc.) in the hospital need to be brought together for the following reasons: 1) solving the problems of data redundancy, data inconsistency, data loss and unavailability, data model incompatibility and time delay in providing medical services; 2) achieving operational efficiency, cost economy, and effective diagnostic decision making; and 3) increasing the hospital's competitive edge by minimizing legal risks, increasing education and research capability, and supporting strategic planning and managerial control [12].

The idea of integration is to bring multiple independent database systems together into one logically integrated system. Several approaches are available. The **physical integration** approach requires physically taking the information, data, and applications existing in each of the independent databases and incorporating them in one overall database system. The **interface integration approach** addresses the need of exchanging (or sharing) data located in different databases and allows the interconnection of large program components. The **virtual database integration** approach combines a number of existing databases into a single, logically integrated whole using a global data model [12].

As evaluated in [12], the virtual database integration (VDI) approach has advantages over the other two. It is less costly, more expandable and will not cause as much disruption to existing system operation as the physical integration approach. It is also more comprehensive and more adaptive to system growth and expansion than the interface integration approach. However, to adopt a virtual database integration approach, a number of complicated technical issues need to be solved, such as schema integration and conversion as well as query translation. Moreover, the administrative and behavioral difficulty of orchestrating the integration process among different departments in a hospital can never be overestimated. Subtle inter- and intra-departmental politics call for caution when removing any stumbling blocks to PACS/RIS/HIS integration [12].

3.1.3 Physical image DB design

The physical design of a DB specifies an appropriate data storage structure and data access schemes. It is concerned with specific devices in which the data are to be physically processed and stored. **1) Data structures and compression.** There are two major methods of storing image data. One is the storage of the original data as is, and the other is the storage of data that are made to resemble the original image by an encoding process. In the former case, the most commonly used data structure would be: two dimensional array, one-dimensional array (or list), tree or hierarchical (variable resolution image) structure such as quadtrees, octrees and pyramids, based on the principle of recursive decomposition [2]. In the latter case, a close resemblance of the original is created by a presentation method that transforms a raw image into symbolic information using basic structural elements of the picture and applies or fits them into curvature or curved surfaces and/or smoothed images.

Because of the large volume of image data, data compression is essential for saving physical storage space. Popular encoding schemes include Run-Length and Huffman coding which are error-free compression and transform coding, and two-dimensional Discrete Cosine Transform (DCT) which is an irreversible form of compression. Hierarchical/tree data structures can also achieve varying degrees of error-free compression. Irreversible compression is more efficient in saving storage (e.g., 20:1 as opposed to 2:1 on the average by error-free compression), but its effect on diagnostic accuracy and legal implications are often attacked by the health care community. Line-drawings (annotations on images) can be broken down to chain codes or series of coordinates of points. In both cases, the original picture must be decoded and reconstructed. The compression ratio and the time for decoding are occasional trade-offs.

2) Hierarchical storage architecture. Because of legal requirements for long-term archiving, the image data to be stored are voluminous. Although a PACS DBS needs to meet the real-time response time requirement, storing the entire collection of PACS image data in high-speed storage devices is cost-ineffective and space-inefficient. Since data retrievals exhibit a temporal correlation such that newly generated images are retrieved more frequently and older ones less frequently, it is sensible to have a hierarchical storage architecture where older images are stored in a lower-cost but slower storage system (e.g., optical disks) and more recent images on a more expensive but faster storage system (e.g., magnetic disks). An important design issue, however, is to determine the data migration scheme among the different levels of a storage hierarchy.

3) Data access support. Special considerations for access support in a PACS DBS are necessary due to frequent "addings" and "archiving" and few deletions of image and related data, coupled with random access with real-time responses required for image retrievals. Common data access support schemes include indexing, such as that based on the B-trees family, and hashing. Both of these suffer some types of deficiency when handling large amounts of data additions. However, since newly generated data are periodically archived into slower storage devices in a PACS DBS, data reorganization literally takes place at every archiving, which can prevent a hashing scheme from performance degradation. Furthermore, a different data access support scheme can be employed for archived data since retrievals are the only activities on archived data.

3.1.4 AI technologies for data management and diagnosis support

AI technologies have been found to have many applications to medical problems [8]. In particular, clinical diagnosis support, radiographic image analysis, interpretation and consultation are typical functions that can use solutions based on expert systems (ES) and neural networks (NN) approaches [24, 22]. Although these examples are notable AI applications in radiology and medicine, they rarely manifest implications for a PACS DBS design beyond its data content, organization and representation and manipulating language. On the other hand, AI application in image data management, exemplified by expert systems for image retrieval [16], is an important issue in PACS DBS design.

The need for image retrieval expert systems arises from the cost-performance tradeoff in computer storage technologies. New radiological exam interpretations often require that new images be compared to previous images, which are typically located in low-cost, low performance storage media such as optical disks or magnetic tapes. According to the current storage technology, on-line retrievals of old images from these low-performance data storage and access devices cannot meet the real-time response time requirement. Therefore, it is essential to raise the retrieval performance of old images to a real-time response level while at the same time maintaining minimum high-speed memory space requirement by pre-fetching to high-speed memory space only the *relevant* old images for comparison with new images [20]. An expert system which embodies radiologists' expert knowledge in image selection has been shown to serve this purpose adequately [16]. Since such an expert system manages data locations in a storage hierarchy as well as in a network environment, its design and development has become an integral part of PACS DBS development.

3.1.5 Database machine for supporting routine reading

As mentioned, PACS application involves many types of retrieval operation (data search). DBMSs rely heavily on pointers to retrieve data and, following the pointers, the retrievals will then involve numerous accesses to the disk at the hardware level. Moreover, if there are operations on data (e.g. join, etc.), the DBMS will have to perform additional processing. This database access (I/O bound) activity can easily tie up a single DBMS and slow down the performance of the whole database system. Therefore, it would be desirable to offload some or all of the DBMS activity (in particular the associative addressing function) from the host machine to some backend system - that is a "database machine." This backend machine can be a dedicated system running DBMS software, functionalized processors or specialized hardware such as high-speed associative memory systems. Although details vary from device to device, each unit in the associative memory system has the ability to select or reject on the basis of arbitrary restriction predicates. For instance, if a data transaction can search based on several predicates, (*Patient Name* = Smith, *Image Equipment* = X-Ray, *Image Age* not less than 1 month), each predicate can be searched in parallel. The link between the host and the database machine would normally be an I/O channel or a communication line. The inclusion of a database machine in a database has features of a distributed system. The database machine approach can improve the system's reliability through complexity reduction and enhance the system's performance through specialization, increased parallelism, and processing power. In particular, a substantially large percentage (e.g., 95%) of radiographic reading activities is "routine reading" in which rules for retrieving image data are simpler and speed is the most important concern of radiologists. The database machine approach (e.g., particularly the parallel searching capability) has great promise in supporting data retrieval for such "routine reading."

3.1.6 Data security

Data security refers to the protection of data against unauthorized disclosure, alteration or destruction. Since patient information is highly confidential, security control is a major concern for PACS DDBS users. The general considerations of security control are physical controls, user access authorization schemes (e.g., passwords, etc.) hardware controls (e.g., CPU provides security features, storage protection keys and privileged operation mode), operating system security, and database security (e.g., "views" in relational databases, access authorization to data, etc.). In addition, an important security problem arising in a distributed database and in a GPACS environment is the initiation and protection of inter-site communication.

3.2 GPACS Data Management Issues

The advances in LPACS and networking technologies promise a future GPACS health care environment consisting of networked LPACS among regional, national or even international hospitals. The technical and behavioral considerations for data management issues in GPACS vary according to "the level of data sharing" required. The desired level of data sharing is determined by the managerial "protocol" among all the hospital entities in a GPACS and can be classified as follows.

1. **Individual files** in case of hospital transfer, seeking for second opinion, or remote consultation. This level only requires a network interface method and translation of format.

2. **A defined set of data** in case of conference preparation. The set of data to be shared has to be explicitly defined.

3. **A small set of data across patients with no predetermined searching criteria** in case of diagnostic conference. All the data should be allowed to be accessed on an ad-hoc basis.

4. **A large set of data across patients with various searching criteria** in case of research. The data should allow various types of comparison and analysis.

Each level of data sharing can be obtained real-time or off-line, depending on the performance requirements of different services. As noted, the first level of data sharing in this category is closely related to clinical service. It involves only one patient file for one retrieval. The second, third, and fourth level of data sharing will require a certain degree of "integration" to allow ad-hoc queries of different local PACS (LPACS) databases. In addition, the second, third, and fourth level will require the management of "Global Medical Data," including data in a "Conference Log" which keeps the cases discussed in medical conferences and a "Research Bank" which stores various research findings. These "Global Medical Data" should be well-managed so that research efforts are not repeated and medical expertise can be shared among medical professionals.

1) Intra-LPACS interoperability. It is clear that the degree of data sharing and forms of data manipulation are determined by managerial agreement among entities involved in a GPACS, and subsequently, shape the technical requirements for GPACS data management. However, even the lowest degree of data-sharing (moving one patient file from one hospital to another) involves many difficult technical problems. The difficulty lies in the diversity of Database Management Systems (DBMSs) utilized by LPACSs in different hospitals and the conflicts in the syntactical format and semantic usage of patient data managed by different DBMSs. The central issues are those of heterogeneous database integration, the same as for intra-hospital interoperability, and shared data directory management in a global environment, which needs to take into account the underlying networking technologies and hospital administration philosophies and their implications.

2) "Global Medical Data" management. The development and management of a "Global Medical Data" database is essential for achieving the full advantages of a GPACS environment. The design issues of this database are similar to those for information systems design in a geographical area, for which such design approaches as the distributed, the centralized or the hierarchical should be considered based on the functional, operational and performance requirements. Its management, however, is a tender issue that requires well- thought-out strategies.

3) Security issues in GPACS. Due to the large amount of image data and the need for utmost confidentiality of patient data, security control in a GPACS environment entails special consideration. Most security control (e.g., user access, database security) has been embedded in LPACSs. The concern in GPACS is the communication/processing cost for imposing certain security schemes (e.g., encryption and physical detection device) during long-distance transmission. Due to the large volume of image data, it appears that the security control for textual data and image data should be considered separately. For example, image data without any indication of patient identity may not be required to be encrypted. But the viable security control in a GPACS will largely depend on the legal requirement for handling image data in the future.

3.3 PACS DBS Design Enabling Technology

Although each of the PACS DBS design issues represents a specific aspect of PACS DBS design that requires special consideration, they relate to one another through the common technologies that enable the development of their solutions. These technologies include data modeling and query support.

1) Data modeling issues. A uniform and proper semantic data model of medical data can provide a foundation for tackling individual design issues as well as for integrating the results into one unified system design without duplicating effort. PACS data typically exhibit an aggregation hierarchy and specialization (subclasses, generalization hierarchy) structure. In our view, the object-based modeling approach [26] is deemed appropriate for this purpose because of its flexibility in supporting multi-media and knowledge-intensive applications as well as its natural fit in representing PACS data. Besides presenting the PACS data, what is more important is the modeling mechanisms needed to support distributed database design, heterogenous database integration, data and knowledge interaction and catalogue/metadata management. Metamodeling for CASE tools to support PACS design is another important data modeling issue.

2) Query support issues. All PACS DBS design issues require some sort of query support. Specifically, distributed database processing needs to address query decomposition and processing over distributed data and heterogeneous DB integration in LPACS or in GPACS depends on query translation and transformation support. Among the AI applications in medicine and radiology, image analysis, diagnosis and consultation often require query by content which needs to be facilitated using pictorial query languages. Other PACS DBS design issues also deal with data access support. It is important to note that PACS query support cannot be addressed separately from PACS data modeling because of the close relationship between how data are represented and how data usage should be supported.

3.4 PACS DBS Implementation Enabling Technology

While the development of PACS DBS design approaches is itself a long-term research challenge, making use of the research results for a PACS DBS implementation that needs front-end execution also requires strategic and technical approaches. Figure 4 depicts the development process of a local PACS (LPACS) DBS and the implementation support required.

1) Specialized PACS DBMS. DB applications have typically employed database management systems (DBMS) to provide the general mechanisms for operations (queries, updates, recoveries, security controls) support and physical database implementation based on an underlying data model. Efficient implementation of a PACS DBS can benefit from using a specialized PACS DBMS, which handles distributed operations and pictorial (image and graphic) data, provides a natural interface to AI applications and is capable of meeting the performance requirements by stripping down the overhead for unnecessary procedures and focusing only on the core operations in PACS. Current DBMS technologies, which have been geared toward general commercial applications, have not in their current forms fulfilled the requirements of a PACS DBMS [17]. Specialized PACS DBMS that facilitates easy implementation of efficient PACS DBS is critical if PACS technologies are to earn a place in the competitive health care environment.

2) Automated design support for PACS DB. Because of differences in hospital organization, managerial style, physical size and distribution of workload, each hospital needs to rely on the information gathered from requirement/distribution analysis of hospital operations to determine its own effective logical data and allocation design. Typically, this is a complex and time consuming process in which DB experts need to be joined by hospital administrators and staff in developing a customized design. Moreover, the process is often hampered by inaccessible knowledge and the design tends to be error-prone as a result of miscommunication.

For PACS DBS to be promptly and efficiently implemented in hospitals, hospital personnel must be provided with user-friendly computer-aided tools that will enable them to design a PACS DDB easily and effectively. Such tools need to incorporate system design expert knowledge and methods so that they can be used to help structure the analysis and design process that employs a semantic data model. This will make it possible for an automated PACS DB design to be derived, one that is provided with human input that includes hospital environment descriptions and can be modified easily when new information becomes available.

3) Standardization. The development of a PACS and its integration into a single or multiple hospitals requires dealing with heterogeneous computing, communication and software environments. Smooth introduction of a system into such environments necessitates not only technical solutions but also standardized interfaces. Although developing a set of standards for any technology is often a painstaking and not always a rewarding process, it cannot be overemphasized how important the industry-wide interface standards are in ensuring harmony in the fast-growing development of PACS.

4 Partial Research and Development Results at UA

This section presents partial research and development results of several research projects at the University of Arizona under the proposed PACS DBS R&D framework.

PACS Distributed DB Design and Modeling
The goal of this project is to develop data allocation designs for a distributed PACS DB, and to evaluate and compare their performance by simulation. Current research results include the followings: 1) three general image/textual DB fragmentation schemes:

fragmentation by modality only, by modality, type of patients and referring section, and by modality, type of patients, referring section and anatomical portion; 2) three general PACS DB fragments allocation schemes: random, cluster and total locality schemes; 3) a SIMSCRIPT II.5 simulation model of distributed PACS image/textual data generation/storage/retrieval on Toshiba S-PACS network using input parameters to specify, among other things, PACS configurations and data allocation specifications; 4) analysis of simulation evaluations of PACS response time and network and database utilization over a range of transaction loads for different design alternatives [17].

Specialized PACS DBMS

The research results include 1) the custom-design of a distributed DBMS for PACS DBS: the prototype was implemented in a Sybase client-server configuration environment. Additional distributed operation capabilities (coded in C language) was added to existing Sybase DBMS functions [17]; 2) a prototype of a distributed DB to evaluate the distributed DBMS design and to gain insights into PACS DBS implementation issues [19].

The Design of PACS, HIS and RIS Integration

We have thus far 1) identified problems among HIS, RIS and PACS; 2) evaluated different integration approaches; 3) examined design issues: schema integration, mapping definition, query/transaction processing and behavior aspects; and 4) investigated schema translation from MUMPS applications to Sybase relational schema using a reverse engineering approach and a semantic data model, Synthesized Extended Entity-Relationship (SEER) model. The use of Sybase Open Server to achieve heterogeneous database integration was also investigated [17].

Image Retrieval Expert System

The purpose of this project is 1) to specify a data migration (prefetch) policy for PACS using an expert system approach and 2) to present the architecture and prototype of IRES which will be coupled with the distributed database system and provide a friendly user interface (interactive processing, explanation capability), report capability, and performance required by radiologists. The architecture of IRES has been designed and a prototype has been implemented utilizing an expert system shell, EXSYS, and a DBMS, dBASEII Plus [16]. particular An expert system tool which can efficiently access a SQL based DBMS is being investigated to couple with the relational PACS DDBS.

Data Modeling and Automated Design Support

The research results include: 1) SEER-DTS (Synthesized Extended Entity Relationship model - Distributed Transaction Scheme)Methodology for integrated distributed database design has been developed and its effectiveness has been verified in an experiment; the methodology also facilitates distributed database design automation [5, 6]. A distributed database design tool, called Auto-DDB, has been prototyped using Visual Basic in a Microsoft Window environment [25]. 2) SOOER (Synthesized Object-Oriented Entity Relationship Model) methodology has been developed to couple IRES knowledge-base and PACS database [18]. A SOOER CASE tool has been prototyped in an X-Window environment. In addition, metadata management for catalogue management and metamodeling issues for CASE tools are being investigated and integration of SEER-DTS and SOOER methodologies is being considered.

5 Conclusion

Integrated information planning and management in hospitals is becoming increasingly important due to the recent rapid development of numerous novel medical information system applications. This paper has provided insights into technical solutions for critical information management issues in hospitals, considering the multi-media nature of PACS technology and the future health care environment (Global PACS) (which is still in infancy). Some of the issues are not unique to medical information management, but their importance becomes accentuated in the context of health care. Further research on technology/methodology for solving these design issues is called for. The research framework for PACS DB development proposed in this paper serves as a strategic plan for efficient and effective future research in this area.

References

[1] Ceri, Stefano and Pelagatti, Giuseppe. *Distributed Databases - Principles & Systems.* New York: McGraw-Hill Book Company, 1984.

[2] Chang, Shi-Kuo. *Principles of Pictorial Information System Design.* New Jersey: Prentice-Hall, Inc., 1989.

[3] Chang, Shi-Kuo and Hsu, Arding. "Image Information Systems: Where do we go from here?" *IEEE Transactions on Knowledge and Data Engineering, Vol. 4, No. 5,* Oct. 1992, pp. 431-441.

[4] Chen Garcia, Hong-Mei and Liu Sheng, Olivia R. "A Semantics-Based Methodology for Integrated Computer-Aided Distributed Database Design." In the Proceedings of *25th Hawaii International Conference on System Science,* January 1992.

[5] Chen Garcia, Hong-Mei and Liu Sheng, Olivia R. "An Entity-Relationship-based Methodology for Distributed Database Design: An Integrated Approach Toward Combined Logical and Fragmentation Design." In Proceedings of the *11th International Conference on the Entity Relationship Approach (Karlsruhe, Germany, October 7-9, 1992).*

[6] Chen Garcia, Hong-Mei. "A Semantics-based Methodology for Integrated Distributed Database Design: Toward Combined Logical and Fragmentation Design and Design Automation." *Ph. D. Dissertation,* University of Arizona, 1992.

[7] Dallas, William J. "Digital Radiology." *IEEE Spectrum,* April 1990, pp. 33-36.

[8] De Vries, P. H. and P. F. De Vries Robbe. "An Overview of Medical Expert Systems." *Methods Inf Med 1985; 24(2):* 57-64.

[9] Goldberg, M., Georganas, N. d., Robertson, J., Mastronardi, J. and Reed, S. "A prototype multimedia radiology communication system." In *Proceedings of the 2nd IEEE International Workshop on Multimedia Communications (Ottawa, Canada, Apr. 1989).*

[10] Hsiao, D. K. and Kamel, M. N. "Heterogeneous Databases: Proliferations, Issues, and Solutions." *IEEE Transactions on Knowledge and Data Engineering, Vol. 1, No. 1,* March 1989, pp. 45-62.

[11] Liu Sheng, Olivia R. and Chen Garcia, H. M. "The Design of Medical Image Databases: A Distributed Approach." In *IEEE 1990 Conference Proceedings of International Pheonix Conference on Computers and Communications (Scottsdale, AZ, March 21-23, 1990),* pp. 288-295.

[12] Liu Sheng, Olivia R. and Chen Garcia, H. M. "Information Management in Hospitals: An Integrated Approach." In *IEEE 1990 Conference Proceedings of International Pheonix Conference on Computers and Communications (Scottsdale, AZ, March 21-23, 1990),* pp. 296 - 303.

[13] Liu Sheng, Olivia R., Chen Garcia, Hong-Mei, and Ozeki, Takeshi et. al. "Requirement Analysis for PACS Database Design." In *SPIE Proceedings of Medical Imaging Conference IV (Newport Beach, CA, February 4-9, 1990,)* 842-854.

[14] Liu Sheng, O. R., Chen Garcia, Hong-Mei and Ozeki, Takeshi. "Database System Development for Medical Image and Information Management." *Under review,* June, 1991.

[15] Liu Sheng, O. R., Chen Garcia, H. M., Wei, C. P. et.al. "Distributed Database Design and Modeling for PACS." In *SPIE Proceedings of Medical Imaging Conference IV (Newport Beach, CA, February 4-9,*

1990,) pp. 256-269.

[16] Liu Sheng, O. R., Wang, H. W., and Chen Garcia, H. M. "IRES - Image Retrieval Expert System" In *SPIE Proceedings of Medical Imaging Conference IV (Newport Beach, CA, February 4-9, 1990,)* pp. 832-841.

[17] Liu Sheng, O. R. and Chen Garcia, H. M. "PACS Database Research." *Technical Report No. 1 - No. 7*, Toshiba Corp, Japan, 1989-1992.

[18] Liu Sheng, O. R. and Wei, C. P. "Object-oriented modeling and design of knowledge-base/database system." In *Proceeding of eighth Int. Conf. on Data Engineering*, Tempe, AZ, 1992, pp. 92-105.

[19] Liu Sheng, Olivia R., Wei, C. P. and Chen Garcia, H. M. "Design and Implementation of a Distributed PACS Database System." In *SPIE Proceedings of Medical Imaging Conference VI (Newport Beach, CA, February 4-9, 1992).*

[20] Martinez, Ralph and Nemat Mohammad. "Image Data Base Archive Design Using Parallel Architectures and Expert Systems." In *Proceedings of SPIE Medical Imaging II, (Newport Beach, CA, Jan. 31 - Feb. 5, 1988).*

[21] Ozsu, M. T. and Valduriez, Patrick. *Principles of Distributed Database Systems.* New Jersey: Prentice Hall, Inc., 1991.

[22] Piraino, D., B. Richmond, M. Uetani, et. al. "Problems in Applying Expert System Technology to Radiographic Image Interpretation." *Journal of Digital Imaging 1989; 1(2)*: 21-26.

[23] Samet, Hanan. *The Design and Analysis of Spatial Data Structures.* Addison-Wesley Pub. Co., 1990.

[24] Swett, H. A., P.R. Fisher, A. I. Cohn, P. L. Miller, and P.G. Multalik "Expert System - Controlled Image Display." *Radiology 1989; 172*:487-493.

[25] Tai, Lin-Tao. "Auto-DDB: A CASE Tool for Distributed Database Design." *Master Project Report*, University of Arizona, 1992.

[26] Woelk, Darell, Won Kim and Willis Luther. "An Object-Oriented Approach to Multimedia Databases." In *Readings in Object-Oriented Database Systems.* Edited by S. B. Zdonik and D. Maier. 1990, pp 592 -606.