Hispanic Digital Archive

Feasibility Rationale

Team 5

Erin Ballew
Cyrus Fakharzadeh
Deepak Ganesan
Nikunj Mehta

Version control:

Date	Author	Changes	Version
10/27/98 7:30 PM	Didi Yao	Initial draft of document	0.1
10/31/1998 7:00 PM	Didi Yao	Incorporated comments of walkthrough	0.2
11/4/1998 1:30 PM	Didi Yao	Documentation standard, LCO Package	1.0
12/2/1998 1:00 PM	Didi Yao	Initial draft of LCA	1.1
12/14/1998 3:30 PM	Didi Yao	Documentation standard, LCA Package	2.0
2/15/1999 2:00 AM	Nikunj Mehta	Risk Assessment added	2.1
2/21/1999 12:00 PM	Nikunj Mehta	Review comments	2.2
2/21/1999 11:20 PM	Cyrus Fakharzadeh	Corrected for updated SSRD	2.3
2/22/1999 12:20 AM	Nikunj Mehta	RLCA package	3.0

Contents

Feasibility Rationale Description (FRD)

The Feasibility Rationale Description (FRD) provides insight into the feasibility and justifications of the Hispanic Digital Archive. This is achieved through several means including a cost analysis, an analysis of the requirements and risks, as well as discussion of various architectures and designs. This document attempts to convince and justify the objectives of the Hispanic Digital Archive (HDA) to stakeholders, in particular, the Customer, Sponsors, and Users. Although the discussions are on a higher level than other project documents, references to other documents are made for details and technical points.

The architecture designed for the Hispanic Digital Archive covers all current and evolution requirements set forth by the OCD, SSRD, and SSAD. These requirements have been thoroughly analyzed to ensure all possible issues and options. We have conferred on these requirements with key stakeholders, namely the Customer and Users, and these agreements are depicted in the prototype and the architecture. Because the implementation of our objectives and requirements was time critical, we concentrate our efforts on a system that can be built in a 12-week time frame. Another constraint of the Hispanic Digital Archive is the cost of development and deployment. There are an unlimited number of options that would work for our system, but many of these options were eliminated in satisfying the above constraints.

Risk assessment is described in this document. Risk management plans are described in the Life Cycle Plan. Handling the trouble arising from unmitigated risks, during the construction phase, is much more difficult than catching the errors in the design phase.

1 Introduction

1.1 Purpose

The Boeckmann Center at the University of Southern California (USC) owns a vast collection of Hispanic-related material that is focused on Hispanic culture and heritage originating from the United States, Latin America, and Spain. The scope of the material is currently limited to the post-1700 era. The main objective of the Hispanic Digital Archive (HDA) is to make the Hispanic-related material, owned by the Boeckmann Center, widely available and effectively accessible to the public. The public in this sense refers to students from any university, researchers, and general public users. Because part of the objective is to present the material in the most accessible manner, a digital archive which can be accessed over the Internet is best suited. Traditional methods include electronic catalogs such as the example of HOMER, which houses library books and other material at the University of Southern California. HOMER uses a predominantly textual-based format for it is media presentation. Electronic catalogs differ from digital archives in that electronic catalogs do not store the content of items. Only the location for that item is given in electronic catalogs. Digital archives actually store item content such as images, table of contents, and abstracts. The heterogeneous collection of the material in the Hispanic Digital Archive requires a system which is highly multimedia oriented because the nature of the archive materials ranges from rich graphics of book covers and posters, to sound from CD-ROMs and video, and motion pictures from video tapes. These inherent qualities of the collection necessitate the need for a Web-based multimedia presentation environment.

An explanation of some of the terms in this document must be explained. The User subsystem refers to the system which a public User accesses the Archive through a Web browser. This subsystem, named by the Customer, is called Zaguán meaning “Entrance”. The Administrator subsystem refers to the subsystem which Administrators, such as Managers or Operators, use to add, modify, or delete items. This subsystem, named by the Customer, is called Mayordomo meaning “administrator of funds of the Hacienda”. The actual archive where the data is stored is called Archivo meaning “Archive”. For clearer understanding, this document will not use official names, but instead will refer to the components as User subsystem, Administrator subsystem, and Archive.

1.2 References

Operational Concept Description version 2.1
System and Software Requirements Description version 2.1
System and Software Architecture Description version 2.1
Life Cycle Plan version 2.2
COCOMO Manual
HDA Documentation Standard
Revised Guidelines for the LCO and the LCA deliverables for MBASE
User Interface Demonstration and Feedback Analysis Report
Proposed Dublin Core Standard

2 Product Rationale

The Product Rationale explains how the Hispanic Digital Archive would satisfy the requirements set forth by the Customer. Cost estimates are explained to explain where money would be spent and why some system components have little or no costs associated with them. Some scenarios of the system usage provide insight on system functionality. Stakeholders' conclusions on requirements and agreements are also provided. For analysis of the software architecture attributes, we used the Parametric Modeling technique, specifically COCOMO II. COCOMO II allowed us to perform effective cost and schedule analysis. Some pitfalls of COCOMO were the subjection of certain attributes and parameter input validation and verification. The following sections of the Product Rationale will be highly dependent on and refer to several sections in the Operational Concept Description, Software and System Requirements Definition, Software and System Architecture Description, and Life Cycle Plan documents.

2.1 Business Case Analysis

This section explains the costs to be incurred and the value gained in return. These costs are only estimates since both the price of hardware and human wages are constantly changing and are also documented in [LCP 5].

2.1.1 Development Cost Analysis

Initial development costs are negligible since students of CSCI 577b implement designs developed by students in CSCI 577a. This schedule is sufficient to satisfy the system requirements since data entry can begin after the HDA is constructed by students of CSCI 577b during the Spring semester of 1999. Data entry can begin in the Summer of 1999; the system will go live when 2% items i.e. 1,000 items are digitized. This is expected to occur roughly six months after the start of data entry. The Customer has determined this percentage feeling that Users can benefit from the Hispanic Digital Archive once approximately 5,000 items are entered into the archive. So, the Hispanic Digital Archive can be expected to be online in September 1999.

Two weeks of system testing is needed toward the end of the Spring 1999 semester. This testing would involve two Operators for entering data and giving feedback on the procedures and the data entry interface. System designers and developers will observe the ability of Operators to learn how to operate the Administrator subsystem and how efficiently they use the subsystem after they become familiarized with it. These Operators would be paid $11/hour for 20 hours a week. So the costs of Operators for testing will be $880. Also, the purchase of two PCs will be needed at this time for the testing Operators and for data entry scheduled to begin towards the end of the Spring 1999 semester. These two PCs have been estimated to cost $1,850 each. [LCP 5.2]

Finally, the costs of image conversion software and scanner TWAIN interface software are considered. We suggest purchasing AccuSoft ImageGear 98 ActiveX-LT 16/32 since it supports TWAIN interface and the conversion of many image types such as TIFF, GIF, JPEG, and BMP. ImageGear costs $300. [LCP 5.2]

2.1.2 Transition Cost Estimate

The costs of implementing the Hispanic Digital Archive will consist mainly of hardware currently not provided by ISD. The tools provided by ISD are free for USC student use and COTS software licenses would be obtained by ISD. The HDA will require server hardware to hold the DBMS and disk storage space and this is expected to be provided by ISD as a part of the IBM Digital Library (IBM DL). The only hardware costs will include two PCs for administrative use to manage the system. The cost of a PC for administrative purposes after academic discount would be $1,850 each as of December 1998. This price is for a Dell Dimension XPS R with a 400 MHz Pentium II processor, 12.9 GB of disk space, 96 MB of SDRAM, and a 17-inch monitor. This particular type of desktop PC was chosen because of the need for rapid image processing of a large volume of images, but since it is not a state-of-the-art system, it is moderately priced. This PC would be depreciated over 4 years after which very little administration work is expected to be required and this could be accomplished by the Boeckmann Center's existing equipment.

Roughly 50,000 items are to be digitized and stored in the Hispanic Digital Archive. This will require scanning of archive materials and loading the digitized information in the IBM Digital Library. Several Operators (2-6 Operators) would be required to be trained to capture images. The cost of digitization as detailed in [LCP 5.2] would be $140,000. This price figure falls within the amount in which the sponsor is expected to donate yearly, which is about $100,000 for three years for the design, implementation, deployment, and maintenance of the Hispanic Digital Archive.

The justification for the above estimates arises from the fact that each item is estimated to take 12 minutes to enter into the archive. Each item will require 12 minutes to enter for the first six months. This includes text and scanned images. With a 2-person group, 80 man-hours per week would result if each person works 40 hours per week. This would result in 400 items being entered in the archive per week. At this rate 5,200 items will be entered in the archive in 13 weeks, or about 3 months. Once the HDA contains around 5,000 items, it will be brought online. This will occur in September 1999 if data entry begins on schedule in June 1999. So, for the remainder of the 3 months in 1999, 5,200 more items will be entered if a similar data entry pace is kept.

Beginning in 2000, Operators are expected to become familiar with data entry since Administrators would have had 6 months of experience. So data entry can be expected to require 10 minutes per item from this point forth. Using this as the item entry rate, 24,960 items can be entered per year so the 50,000 items can be entered in just over 2 years. Because of unforeseen circumstances, we have allotted 3 years for the complete entry of data in to the Hispanic Digital Archive. Therefore, the projected completion of the HDA will occur during June 2002, the target date set forth by the Customer.

The Customer will be required to hire a qualified Project Manager to assist in approving data entered by Operators and the cost for hiring this person would be $40,000 per year for 3 years.

Costs of setting up and installing the two Dell personal computers are minimal because they will require no more than four hours to unpack and plug in. Dell computers are shipped with pre-installed software so time need not be spent for software installation. The Boeckmann Center already has working Internet connections, so no additional costs and time need to be incurred for network connections. During set-up of the personal computers, the scanner will also be installed. The Boeckmann Center already owns a new Hewlett Packard ScanJet 4c scanner that scans at 600 dots per inch. They will not need to purchase a new scanner for data entry purchases. However, in the event that the Boeckmann Center is retrofitted, the Customer would be responsible for the site preparation required for the system. The Customer also envisages the installation of an operator center at the East library for rapid digitization.

2.1.3 Operational and Maintenance Cost Estimate

Administrative and maintenance costs will be required once the system is set up. System administration consists of data entry, manager operations, and DBMS maintenance. Data entry includes scanning images for archival input, textual data entry, and updates to items in the DBMS. Manager operations include data entry beyond the initial bulk loading of data as well as approval for item updates. DBMS maintenance includes the day to day system checks such as monitoring disk capacity, the number of concurrent users,

and the operational state of the DBMS. Costs of DBMS maintenance are not considered since it is a part of the IBM Digital Library that is maintained by ISD.

Initial costs such as software and hardware costs are only one-time costs. The only exception is the case of upgrading to new software and hardware components due to failure, a change in system functionality, or scaling the system to handle more data or more users. [LCP 5.2]

Future CSCI 577 student teams can perform perfective and corrective maintenance. In fact some of the *Optional* features not completed in the current production can be deferred to the next year and built by another student team. The cost of this new development effort is expected to be zero since it involves voluntary effort from the students.

2.1.4 Estimate of Value Added and Relation to Cost

The main rationale in developing the Hispanic Digital Archive is to be able to present the Hispanic-related material at the Boeckmann Center in an efficient, easy-to-use manner. Our solution is to store the material in the form of a digital archive accessible from the Internet. Patrons can also borrow material they find from the Hispanic Digital Archive.

The HDA adds value to the collection of material owned by the Boeckmann Center by making it available to the general public. The intended use of the collection is for people to explore and learn from it. The current system for doing so requires a patron to visit the Boeckmann Center and search through materials by hand. The patron can then view the material, but only within the Boeckmann Center since materials cannot be loaned out. The Hispanic Digital Archive is an improvement since it makes the material available from any computer with a Web browser and an Internet connection. Thus, contents of the archive can be viewed without visiting the Boeckmann Center. The HDA will maximize the availability of the collection by taking advantage of the Internet and archiving technology. Clearly, the community benefits of the Hispanic Digital Archive outweigh the monetary costs of developing and deploying it. The HDA, as mentioned above, will only incur development costs for future improvements and features as well as minimal software and hardware costs. The bulk spending happens during the data entry stage. Approximately 50,000 items to be entered by 8 data entry consultants will require about 3 years to complete.

The Paging Request feature adds value to the system because there does not exist any way for efficiently requesting material. Currently, patrons must physically visit the Boeckmann Center and verbally request for a specific item. The patrons must then return to the Boeckmann Center to obtain their requests because many of the items are not physically located there. With the Paging Request system, users can send an electronic message to the librarian at the Boeckmann Center so that an item can be retrieved and be available on a specified date when patrons arrive at the center.

Current cataloging procedures require a cataloging expert to correctly organize and sort materials in a collection. This requires procedures to ensure security from theft and vandalism since most cataloged materials are for loaning out to the public. Such preventive measures include applying magnetic security tape to materials so that they trigger alarms if the magnetic tape is not desensitized during checkout. In addition, protective covers and lamination must be applied to certain cataloged material so that borrowers do not accidentally or maliciously damage materials due to liquid spills, dirt accumulation, or tear and mutilation. These are all concerns that administrators must address when material is loaned to patrons for extended periods of time.

In addition to preservation methods, catalogers must fill out cards with pertinent information about the material. Often times, some information such as an author or a title is missing from a certain piece of material. In order to complete the cataloging of an item, catalogers sometimes must embark in labor-intensive research to uncover missing pieces of information. With all this taken into account, the Customer has estimated the cost of successfully cataloging one item to be between $15 and $20. In other words, catalogers receive $15-$20 for each item cataloged. With this in mind, entering one item into the Hispanic Digital Archive is relatively cheaper as explained below.

Data entry of 50,000 items, including two PCs and a Project Manager for 3 years and software, is estimated to cost $264,000 as explained in 2.1.2 Implementation Cost Estimate. This amounts to approximately $4.12 per item, which is approximately 4 times cheaper than cataloging the 50,000 items using traditional cataloging methods. The main reason of the cost savings is because of bulk loading and the shortcuts during data entry. These shortcuts save time by using drop-down boxes with selectable choices for certain fields such as country, type, language, and collection. Selecting a choice is much faster than having to type in an entry. Administrators can, of course, modify these choices. Also, the trimming of the data being entered into the HDA is a substantial time saver. The HDA requires many less fields to enter than traditional systems such as HOMER. This trimming of data entry means that minimal data is created and stored. Data entry in the HDA represents a substantial savings compared to card cataloging and also provides a more accessible means to view material more quickly. Theft or destruction of items is also not a concern since items are not loaned to the public, so preventive measures are not needed. However, a differentiating cost is the maintenance costs associated with the HDA. This may be included as software upgrades and/or future data entry beyond the initial 50,000 items. Future data entry exhibit similar costs as initial data entry, so the comparison to additional cataloging is that data entry is cheaper. Software upgrading is the varying cost since it depends on how extensive the software is to be upgraded or whether new hardware is needed. As with any system software upgrade, this could be relatively costly, but the Customer currently does not foresee the need for extensive upgrades. Also, administrative services for the IBM Digital Library will continue to be free of charge for the Boeckmann Center.

2.2 Requirements Satisfaction

This section draws upon the architecture of the Hispanic Digital Archive to show how the HDA is able to achieve all the requirements. The requirements have been divided into four sections: capability, interface, quality, and evolution.

2.2.1 Capability Requirements

The specific architecture, which we have imposed, sufficiently covers all the crucial requirements. The requirements listed below are divided into Administrator subsystem requirements and User subsystem requirements [SSRD 2.3]. Criticality of capability requirements is described in one of three ways:

1. *Core Essential*
   - Required functionality to be provided at the earliest
2. *Core Priority*
   - Highly desired functionality to be mandatorily provided
3. *Optional*
   - Good to have feature that can be deferred to future releases

2.2.1.1 Administrator Subsystem

[Archive Administrative Subsystem (SSRD 2.3.1.1)]

Function	**The administrative subsystem allows administrator authentication.**
Criticality	*Core Priority* – This is to ensure that only authorized personnel can modify the items in the DBMS.
Technical Issues	System needs to recognize 2 levels of privileges, one for Operators and one for Managers.
Schedule	Increment II of the Construction phase
Risks	Initials are not a very safe way of authentication as they can be easily imitated.
Dependencies	RQ-01

Function	**The administrative subsystem adds, modifies, and deletes new archive content types.**
Criticality	*Core Priority* – To allow content types to be added, modified, and deleted.
Technical Issues	Must include all possible fields for items in case a new type requires new fields when adding types. If the type of an item is deleted, the system must handle data integrity so

	that item cannot be of non-existent types. The displayed fields of the modified type need to be adjusted according to what needs to be displayed for that type.
Schedule	Increment II of the Construction phase
Risks	The new type may not fit into the item templates currently in use. Templates may have to be modified. If data integrity is not reliable, some items may have dangling types when a type is deleted. Modified types may not display the correct fields so the content will not be accurately displayed to users.
Dependencies:	RQ-02, RQ-03, RQ-04

Function	**The administrative subsystem adds, modifies, and deletes new items to the HDA.**
Criticality	*Core Essential* – This allows the addition, modifications, and deletion of new items, which is vital to the system.
Technical Issues	The new item must be of at least one existing types. It must contain certain fields such as country, language, and type. Before deleting or modifying an item, a search has to be performed based on the defined fields. Two items cannot have identical locations in terms of box folder and item numbers.
Schedule	Increment I of Construction phase
Risks	Added items may be incorrectly inputted. Additions, modifications, and deletions could adversely affect the DBMS during a crash. The recovery mechanism of the DBMS should prevent any loss of data.
Dependencies	RQ-05, RQ-06, RQ-07

Function	**Maintain Archive item fields.**
Criticality	*Core Priority* – This allows certain item field options to be edited.
Technical Issues	Certain item fields provide options for reducing the number of key strokes. When existing options are removed, there could be dangling options in use by the item, but not defined in the interface. Options cannot be modified, they can only be deleted once added. Changes made by one administrator are not visible to other concurrent administrators.
Schedule	Increment II of Construction phase
Risks	Dangling options when deletions are made
Dependencies	RQ-08

Function	**The administrative subsystem can be used to add images associated with archive items.**
Criticality	*Core Essential* – This allows the administrator to present images associated with an item to help the user understand more about the nature of an item to be viewed.
Technical Issues	Images would need to have some form of watermarking applied to protect the copyrighted material. Scanned images have to be preconverted to GIF or JPG form. This increases the latency of the workflow.
Schedule	Increment I of Construction phase
Risks	Images may take up too much disk space if too many of them are used. The operator may not convert images to the correct format.
Dependencies	RQ-09

Function	**The administrative subsystem can be used to modify the order of images and delete images associated with archive items.**
Criticality	*Core Priority* – This allows the administrator to change images associated with an item to help the user understand more about the nature of an item to be viewed.
Technical Issues	Reordering images may require deletion and subsequent addition in the new order. This would increase the time required to perform such an operation. A new thumbnail has to be created to be in sync with the first image for the item
Schedule	Increment II of Construction phase
Risks	The system would become very slow to respond if the item had many images and a

	reorder required for moving many of them around.
Dependencies	RQ-10, RQ-11

Function	**Obtain Manager approval for the archive updates made by the Operator.**
Criticality	*Core Essential* – This allows the Manager to catch mistakes made by the Operator before the erroneous item is available to the public.
Technical Issues	Manager may not want to check each item individually, so the system should just be able to change the status of a group of items. Unapproved items should not be available to the public.
Schedule	Increment I of Construction phase
Risks	For want of time, the Manager may not be able to verify records of many items, thus the items would remain inaccessible for a long time.
Dependencies	RQ-12

Function	**Scanning of images can be done directly from the administrative subsystem.**
Criticality	*Optional* – This simplifies the scanning process and enables the administrator to readily do all administrative work from Mayordomo directly.
Technical Issues	TWAIN compliant scanners and scanner software required. No software supported by ISD for this purpose
Schedule	Increment II of Construction phase
Risks	The software may have integration issues with the rest of the system. It would also require dramatic changes to the User Interface for managing the image list.
Dependencies	RQ-13

Function	**The administrative subsystem allows for watermarking of images.**
Criticality	*Optional* – This allows the administrator to protect the images for copyright purposes.
Technical Issues	Would require handling bitmap images in Visual Basic. Updating the item should be delayed so that the final set of changes are written instead of writing each time the administrator makes a change or addition.
Schedule	Increment II of Construction phase
Risks	API for performing watermarking is currently not well understood by any developers or ISD consultants. Performance might be degraded due to the invocation of other DLLs. Installation problems could arise
Dependencies	RQ-14

Function	**The administrative subsystem can be used to view item and print details.**
Criticality	*Core Priority* – This allows the administrator to view item information to verify correctness and print out for maintain as a record.
Technical Issues	Printing directly to the printer requires printer interfacing and precise layout of item information. No prototype has been developed for this. Viewing the items requires a search based on a large number of fields or their combination.
Schedule	Increment II of Construction phase
Risks	Printing requires interface to external logic. Search could become complicated and there could be a large number of matching results.
Dependencies	RQ-15

Function	**The administrative subsystem has a user access manager that allows for changing details for an administrator**
Criticality	*Core Priority* – This allows the Manager to add new users to the administrator subsystem and adjust the privileges of the users, and change the details for an administrator.
Technical Issues	Initials cannot be removed, because they are they identify the creator/modifier of an item.
Schedule	Increment II of Construction phase

Risks	Initials once assigned cannot be changed. The administrator privileges have to be set to none to disable that administrator
Dependencies	RQ-16, RQ-17

Function	**Context-related help on Mayordomo**
Criticality	*Core Priority* – This allows the administrator to get help on the functionality of the administrator subsystem.
Technical Issues	Help creation facilities not available. Time too short and personnel not available to provide grammatically sound and simple explanation of features online
Schedule	Increment II of Construction phase
Risks	Help may not be too helpful if the administrator has to wade through a lot of text to find the required information
Dependencies	RQ-18

Function	**Mayordomo would include a splash screen**
Criticality	*Core Optional* – This allows the administrator to view the identity of the software.
Technical Issues	The use of a splash screen makes the system difficult to reuse, unless a name is used which allows the actual picture to be replaced.
Schedule	Increment II of Construction phase
Risks	The splash screen may be irritating to the administrator.
Dependencies	RQ-19

2.2.1.2 User Subsystem

[Archive User Subsystem (SSRD 2.3.1.2)]

Function	**The user subsystem allows the user to find the archive items by browsing or by keyword.**
Criticality	*Core Essential* – Search: This is the primary way of finding an item in the archive. Optional – Browse. Many users do not resort to using a browse.
Technical Issues	Many different ways of browsing. Browsing can be done within collections, item types, country, or any other fields.
Schedule	Increment I of Construction phase. Browsing would be completed in Increment II.
Risks	Browsing methods may not be logical for every user since they are looking for different items. The interface may be too complicated to use because the design has not been properly tested and analyzed.
Dependencies	RQ-20, RQ-21

Function	**The user subsystem allows the user to view and print information including images on archive items.**
Criticality	*Core Essential* – These item pages hold all the content of a particular item.
Technical Issues	Selecting and deselecting items track the items of interest from a set of items. Requires a memory of previous user actions.
Schedule	Increment I of Construction phase
Risks	Selecting an item may be a problem if that item is suddenly deleted from the DBMS.
Dependencies	RQ-22, RQ-23, RQ-27

Function	**The user subsystem allows the user to make paging requests for archive materials.**
Criticality	*Core Priority* – This retrieval request system will cut down time for users since the item will be available by the time they arrives at the Boeckmann Center.
Technical Issues	There is only one-way communication from User to Administrator so the Administrator cannot reply back.
Schedule	Increment II of Construction phase.

Risks	Misuse of facility either inadvertently or on purpose by Users for paging archive materials from storage. E-mail information could change, thus breaking the link between users and the administrators
Dependencies	RQ-24

Function	**The user subsystem provides an online help manual for user assistance.**
Criticality	*Core Priority* - Help manual is useful for inexperienced users.
Technical Issues	User Manuals should be easily comprehensible and useful for performing any defined system responsibility.
Schedule	Increment I of Construction phase
Risks	Help manuals will cause more confusion if designed poorly.
Dependencies	RQ-25

Function	**The user subsystem allows for emailing of archive items.**
Criticality	*Optional* – The patron need items to be emailed.
Technical Issues	The user provided email may not be a valid one. The emails would bounce if that happens resulting in accumulated garbage in the archive.
Schedule	Increment II of Construction phase
Risks	Unnecessary use of disk space to accommodate the bounced emails
Dependencies	RQ-26

Function	**The user subsystem allows for bookmarking of items.**
Criticality	*Core Priority* – Bookmarking permits the user to get keep track of various pages of interest to make them easier to view again.
Technical Issues	Bookmarking requires all the information to be provided in the URL which makes generation of URLs and recovery of information from the URL difficult. Item descriptions should be cached for at least 15 days to reduce load on the server
Schedule	Increment II of Construction phase
Risks	The bookmarked items may be moved to review and not be available again
Dependencies	RQ-28

Function	**The user subsystem includes Acknowledgements and Overview information**
Criticality	*Core Priority* – Acknowledgements and Overview information is important to the user who wishes to find out more about the Hispanic Digital Archive.
Technical Issues	It would be difficult to the customer to change the contents of these pages as they are not directly maintainable from Mayordomo
Schedule	Increment II of Construction phase
Risks	The information may not always be accurate or up-to-date.
Dependencies	RQ-29

Function	**The user subsystem includes logos, e-mail and contact information**
Criticality	*Core Priority* – Logos are used to preserve the uniqueness of the user interface. E-mail and contact information is helpful for the patron who wishes to know how to contact individuals associated with the Hispanic Digital Archive for assistance.
Technical Issues	Legal requirements for using specific images at specific locations in the web page would complicate matters.
Schedule	Increment II of Construction phase
Risks	E-mail and contact information could be out of date.
Dependencies	RQ-30

2.2.2 Interface Requirements

The Hispanic Digital Archive will be housed in the Information Services Division (ISD) at USC. ISD will use the interfaces provided by the IBM Digital Library to maintain the HDA. There is also a specific requirement to enter data into the HDA via IBM Digital Library routines. Our front-end software will need to communicate to the data in the IBM Digital Library using the Digital Library Dynamic Page Builder. Another option considered was CGI, but it was not adopted as its use is strongly discouraged by ISD.

Operators, who enter in data, will need to interface with the administration subsystem for the Hispanic Digital Archive. Operators will need to learn the correct format of entering data and procedures such as scanning images, entering sound or movies, and typing in textual input. Operators are required to use this specific interface because the design of the administrator access page is fixed and is not able to be customized for different operators.

Users will interact with the user subsystem pages using a set interface. This interface requires that users know how to use search engine techniques. The search engine interface is designed to be self-explanatory. Information on each of the searchable categories is provided in case users are not familiar with standard searching techniques. When designing the graphical user interface for the user subsystem, we assume that the users know very little about computers and searching techniques. This is because the general public, a heterogeneous group of people who may not possess extensive computer knowledge, will use the Hispanic Digital Archive.

Scanners will interface to software for instant image conversion and scanner control through programs using the TWAIN scanner interface is supported by AccuSoft's ImageGear software. This will enable program applications such as the Hispanic Digital Archive's Administrative subsystem to directly control scanner operations and functionality. [SSRD 4, SSRD 5, SSAD 3.1.2, SSAD 3.1.3]

2.2.3 Quality Requirements

2.2.3.1 Dependability/Fault-Tolerance

Availability

This system will always be online except for the 2-hour downtime each week for any software upgrades or modifications to be made to either the User subsystem or the Administrator subsystem. Since updates to the DBMS can be made while the system is still online, these updates will not affect the availability of the HDA so they can be performed anytime. Two hours will be enough time for all software maintenance to be performed. This is described below in the Serviceability portion of the Quality Requirements.

In addition to the 2-hour downtime, network failures and congestion will affect availability. These failures and congestion will result in delays and interruptions between Users and the HDA when Users make requests and queries. Other than the weekly 2-hour software maintenance downtime, the availability of the Hispanic Digital Archive will depend on the integrity of the network, which connects Users to the HDA. To alleviate a potential congestion problem on the digital archive's end, concurrent usage monitoring is used to detect a shift in the number of concurrent Users. An increase in the number of concurrent Users will cause congestion since the HDA will need to handle more requests resulting in slower replies. More threads can be assigned to the HDA so that more concurrent users can be accommodated. This solution does not affect congestion occurring in the network. Network congestion is a separate issue that cannot be mitigated by the HDA. [SSRD 3.1.1]

2.2.3.2 Security

Privacy/Confidentiality

The Hispanic Digital Archive is secure from the User subsystem end because the Net.Data macros being used for accessing archive items do not allow any modification of information. This also avoids direct access to the IBM Digital Library DBMS.

Administrators use two letter initials to enter the Administrator subsystem. The system will recognize the type of administrators by determining their privilege level. Operators and Managers will use different passcodes so the system will be able authorize the amount of privileges based on the type of the administrator. The IBM Digital Library contains a password authorization system, which we will use for the Administrators to gain entry into the Administrator subsystem. Separating the User subsystem from the Administrator subsystem prevents any security breaches by Users. Users only have read and querying capabilities from the User subsystem end. Additions, modifications, and deletions to the Hispanic Digital Archive can only be performed through the Administrator subsystem. [SSRD 3.2.1]

Integrity

From the Administrator subsystem end, Operators are able to make modifications, but these changes are pending and can be disapproved by the Managers. Managers have complete access to the digital archive and can add, modify, and delete information.

Once the item information is updated and the change has been approved, it will be immediately available to all users. This happens because the DBMS in the IBM Digital Library will return the same information to any client that requests it. This ensures integrity of data across all Users so that they will be viewing the same data in the same format as other Users. [SSRD 3.2.2]

2.2.3.3 Usability

Ease of learning

Learning to use the Hispanic Digital Archive can be facilitated through several means. An HDA User's manual will be provided to familiarize the administrators with functionality and the capability of the system. This manual will be written in a simple manner using few technical terms so that Users of all computer skill levels will be able to understand the manual after 2 hours of studying. It would also be used for training operators for data entry. Another online user's manual will be available on the Internet for the Patron subsystem.

There are also help sections directly embedded in the User subsystem. These include general "Help" buttons on most pages that need explanation.

The design of the User subsystem and the Administrator subsystem will lend itself to ease of learning. Many of the pages are self-explanatory even for Users with little computer knowledge. To do this, we have designed prototypes, which will be used for testing and obtain feedback in order to improve the ease of learning of the system. [SSRD 3.3.1]

Ease of Use

Usability is the foundation of the design of our system. We made the assumption that the typical user of the HDA is not very computer savvy. This is because the system is made available to the general public, which consists of many people who are not too familiar with computers and digital archives. Thus, by making our system user-friendly and straightforward, users of all types of computer skill levels can be accommodated.

Our prototype tests and user reviews demonstrate the usability of the HDA. This is our way of ensuring usability of the HDA. Users and Customers will examine and test the prototype while looking for any usability issues. Feedback from these tests is valuable since it provides developers with a User's perspective on the system. [SSRD 3.3.2]

Help Requirements

Help will be provided for the Administrator subsystem and the User subsystem through the use of both on-line and paper based help manuals. The online help for Administrator subsystem would be in the form of context related help, where buttons would be provided on each screen for help with using the screen. The manuals will provide the necessary explanations for Administrators to successfully add, modify, and delete items, item types, and other fields. These manuals will also teach Users to perform successful searches and retrieval requests of items. [SSRD 3.3.3]

2.2.3.4 Performance

Response Time

Search results will return a list of matched items in less than 10 seconds after the search query is submitted by the User. The network speed could be a bottleneck in this case; the IBM Digital Library DBMS contains many indices for fast searching and there will be no table joins involved. The largest table will be the Items table, which will contain the 50,000 items in the Hispanic Digital Archive.

Another factor affecting response time is the number of concurrent Users. This will require the HDA to handle more total requests and, in turn, spend more time in returning replies. More threads can be assigned to the HDA to handle this problem, as described above in the Availability section of 2.2.3.1 Dependability/Fault-Tolerance. [SSRD 3.4.1]

Network Bandwidth Usage

Many of the general public Users use 33.6Kbps modems since this is a mainstream modem speed. Users are not expected to wait more than 15 seconds for a specific item page containing images to be downloaded. With a 33.6Kbps connection, 63 Kilobytes can be downloaded in 15 seconds. The amount of bytes downloaded will actually be slightly lower if processing and search time for that item is accounted for. In most cases, items will contain approximately 1 KB of text and five images of 10 KB in size. However, the images are loaded when the user clicks on a thumbnail, so only the thumbnails and text will be initially loaded for each item. With thumbnails being approximately 1 KB, the initial load of five thumbnails and text for an item would require no more than 2-3 seconds for downloading. [SSRD 3.4.2]

2.2.3.5 Portability

This Hispanic Digital Archive will be available to students, researchers, and the general public as long as they have access to a computer with a modern Web browser such as Netscape 4.x or Internet Explorer 4.x. Only the User subsystem, not the entire system, needs to be portable. This is because the Administrative subsystem will only need to be used on one specific platform housed by the Boeckmann Center. Availability of the User subsystem will increase as testing of the front-end continues on different types of Web browsers and computer platforms. However, our developers and quality assurance personnel would need to conduct testing for the Patron subsystem on various platforms.

Administrators will always have access to the User subsystem and the Administrator subsystem because they will be well tested on the administrator's browsers and computer platforms. [SSRD 3.5]

2.2.3.6 Reusability

The software that we have developed for the Hispanic Digital Archive can be reused for other digital archives requiring similar goals and functionality since the basis for the archive is the proposed Dublin Core standard. The design of HDA also promotes reusability. Item types and Item field options can be modified, added or deleted. In this way, other archives could be represented as digital archives by using the components of the Hispanic Digital Archive. The components such as the User subsystem and the Administrator subsystem are specifically designed to run off the IBM Digital Library. The IBM Digital Library at USC's ISD uses IBM's DB2 as a database management system, but also can operate using the Oracle DBMS. IBM Digital Library runs on IBM AIX and Microsoft Windows NT operating systems.

Therefore, our HDA software components can be reused for other systems that try to achieve similar goals if they are running the IBM Digital Library. This is not such an unachievable requirement if the system

reusing our components is located within the domain of USC. USC has licensed IBM Digital Library so that any facility located within USC can use it. [SSRD 3.6]

2.2.3.7 Compatibility

Legality

Some of the material in the Hispanic Digital Archive is copyrighted. The copyrights of these materials have to be protected and so they cannot be completely digitized and published on the Internet. There are some rare articles in the public domain possessed by the Boeckmann Center. A method has been devised to protect copyrighted material by imprinting a watermark on all sensitive material. This watermark would appear as a faint image in the background of the image so to not detract from the content of the image, but also to make evident that this particular image is copyrighted. The image can still be copied by downloading it to a remote machine, but the watermark will remain and indicate that it is copyrighted. The IBM Digital Library provides the means to use watermarking on images. The IBM DL has tools to facilitate the use of these watermarking techniques.

In terms of protecting copyrighted text, all sensitive text appearing in the HDA will be displayed as images. Because text from materials is displayed as images, watermarking techniques may apply to text scanned as images. [SSRD 3.7.1]

2.2.4 Evolution Requirements

The scalability of the HDA is important in the case that many more items beyond the initial 50,000 need to be entered into the system. The DBMS will be able to handle these increases provided that disk space is available. DBMSs have a configurable number of threads so increasing the number of concurrent Users can be accommodated by increasing the number of threads. The performance of the system will decrease because of the increase in Users, but the hardware is also scalable to accept the addition of servers. IBM Digital Library is able to handle these scalability factors because, like any other DBMS, IBM Digital Library can be scaled to allow for more concurrent User access.

The HDA will evolve to use more efficient scanning methods and procedures. Currently, the system is designed for Administrators to manually scan materials to be entered through the Administrator subsystem. Since this process is tedious and labor-intensive, a more automated method must be devised. Specialized outside companies offer services to scan in physical material and store their images into microfilm media. This microfilm media serves as a preserver of the material and also allows for easy conversion from microfilm to digitized images, which the HDA can incorporate into its DBMS. Although this method of microfilming will save Administrator time, costs of these services are a trade-off that must be considered.

As another future requirement, the HDA will allow administrators to add new Operators and Managers through the Administrator subsystem. The passwords could also be altered for individual Operators and Managers. Administrators could also be upgraded from an Operator to a Manager.

The design of the system will allow for more features to be added in future versions. The Archive could house other media types such as audio and video. A 3-tier architecture will allow for the addition of new capabilities because adding features to the server will be more transparent to the client [SSRD 6]. A student team in 1999-2000 can build all the incomplete features and other requirements considered but not implemented in 1998-99.

2.3 Operational Concept Satisfaction

The Hispanic Digital Archive can be used in two different scenarios: User and Administrative.

The User scenario entails users such as students, researchers, and the general public who need to retrieve data from the HDA.

User access:

Users start at the User access page and enter in keyword(s) to narrow down the particular item they are searching for. The User will also be able to make advanced reservations of a particular item by using the Retrieval Request system. A message will be sent to an administrator at the Boeckmann Center who will obtain that item, typically from the East Library. This gives time for the administrator to retrieve an item so that it is available by the time the User visits the Boeckmann Center. The HDA will allow Users to browse the archive if they do not have a particular item to search. This browse capability will be developed if time permits, otherwise it will be incorporated in future versions. Items can also be selected or deselected which places and then collectively bookmark this list. An exception-handling scenario is when a user searches for a keyword which returns no items. This will cause a page to appear indicating no items were found.

The Administrative scenario can be broken down into two categories: Manager scenario and Operator scenario.

Manager maintenance of data:
The Manager can add, modify, and delete items from the archive. The Manager also approves of changes made by the Operator. Full access and privileges are given to the Manager. The Manager can also add new types and fields to the database schema through interfaces in the Administrator subsystem. An exception-handling scenario is when an administrator enters the wrong password, in which a message will appear indicating an unacceptable password was entered giving the administrator the option to enter a new password.

Operator entry of data:
The Operator has the ability to add and modify data only. This includes scanning images and entering textual information. The Operator's changes are considered to be pending changes since the Manager needs to approve them. The Operator might enter incorrect data such as searching for an Item # which does not exist. In this case a message will be returned indicating the failure to find that item. [OCD 5]

2.4 Stakeholder Concurrence

All of the agreements, were voted on by stakeholders and were passed. The participating stakeholders include Developers, Users, Customers, and the Project Manager. Barbara Robinson represented the Customer and User needs and requirements. She participated in voting on all the agreements of the WinWin negotiations. In the conclusion of our WinWin negotiation process, all of the system requirements, or "Win conditions", were covered by the agreements either directly or indirectly. Here, indirectly means an agreement was reached through an issue and an option of a Win condition. These agreements each have a priority level (Very High, High, Medium, Low, Very Low) which determines it's importance and, hence, the order in which to implement the objective of each agreement. [OCD 4.1.1, LCP 3.2]

3 Process Rationale

This section describes our development methodology by listing system requirements by their priority level. Those requirements with the highest priority are developed first. Requirements determined to have low priorities will be developed in future upgrades by a development team other than the CSCI 577b student group.

3.1 System Priorities

The main priority of the system, in addition to successfully designing and building the archive, is to design a system which is able to be implemented in a 12-week span. The system would be implemented in CSCI 577b during January-April 1999. Based on the priorities assigned to each Win condition, each requirement is categorized as an Essential Capability, a Desirable Capability, or an Optional Capability [SSRD 2.3.1, LCP 3.1]. These are further mapped to phase priorities and divided into *Core Essential, Core Priority* and *Optional requirements* for the production phase [LCP 2.2.2].

Essential Capability – *Core Essential*:
Provide global access to the archive, prominent recognition to the sponsor, provide multi-platform support, a visually attractive main page, a prototype that demonstrates functionality, create minimal amount of input

text or metadata, create virtual collection groups, different ways of representing different data, flexibility to add/delete data types, use DBMS, simple GUI, understandable search results, keyword based search, ease of use for administrative functions.

Desirable Capability – *Core Priority*:
Ensure scalability of the system, dual levels of administrative access, low cost of development, use IBM Digital Library, provide training manual, minimize steps to reach information, do not provide too much data, high speed system, concise objective on front page, representative sample of material, and category based browsing capabilities, allow users to page archive material in advance, integrated scanning support, itemized reports.

Optional Capability - *Optional*:
System should interface with SIRSI search engine, user profile management, provide hit counter on main page, provide hyperlinks to other Web resources, jukebox with audio streams, voice recognition, familiar formats for input, access to site should be straightforward, minimize false hits, maximize correct hits, easy navigation, feedback during processing, user control of audio/video, select/deselect for user list of search results and statistical archive reports.

3.2 Process Match to System Priorities

We use the Spiral model to design the Hispanic Digital Archive. Once a cycle has been completed and stakeholders' Win conditions have been identified, the next cycle is used to elaborate on the conditions and possibly modify them. The Spiral model allows us to consider all possible requirements first before working out details in subsequent cycles.

We are not performing incremental steps in our system analysis and design. Using incremental steps requires that we complete one section of the requirements in full detail before beginning the next step. This means that we would specify all Win conditions before establishing next level objectives, constraints, and alternatives. Building the prototype incrementally would allow us to gain feedback and tune functionality accordingly.

Incremental steps will, however, be used during the construction stage of the Hispanic Digital Archive. This is because certain components depend on the functioning of other components. In this way, the core capabilities can be completed first while additional features can be completed later. In case our schedule is not met and the implementation runs longer than expected, the core requirements and capabilities will be satisfied and the system will still be able to function.

The choice of process model we have decided on is the Evolutionary Development Model. The growth envelope of our objectives and constraints is on the range of limited to medium since we have determined most of our objectives and constraints. During the inception of the Hispanic Digital Archive, the understanding of the requirements was fairly low because we did not fully grasp all of the concepts and capabilities of the archive. Robustness is not required to be too high so low-medium is fine since the we are not dealing with critical life and death scenarios. There are many aspects to our architecture and the ability of the HDA to function depends on a logical working architecture. So we must have a high understanding of it. [LCP 2.1]

3.3 Consistency of Priorities, Process and Resources

Resources of the Hispanic Digital Archive will be efficiently used and allocated. Every member of our group has defined roles and responsibilities to ensure that all aspects of our project are covered. For example, the System Architect handles the structure of how the HDA will work. The Architect deals with issues concerning implementation alternatives and interaction between different system components. The skills and strengths of group members are also fully utilized. The group member with the most knowledge and experience in Web design and IDEs implements our prototype design. Moreover, other group members gain invaluable insight and first-hand knowledge from the skilled group member. [LCP 3.1]

4 Risk Assessment

This section identifies the major risks involved in this project. These risks are managed and controlled as described in [LCP 4.1].

4.1 Personnel Shortfalls and Group Member Interaction

Description: Staffing the project with incompatible or inappropriate personnel.
Risk Exposure: Quality and schedule of completion of the project could be sacrificed.
Actions to Mitigate Risk: The team for the fall semester that is going to design the system architecture has already been formed on the basis of compatibility of team members and a balanced mix of domain requirements. These include user interface skills, analysis skills, programming ability, design knowledge, as well as web development experience and management ability. For the spring semester, 3 team members will continue on the project. This results in less time for training new team members and a better understanding of the project by continuing members.

Description: The availability of personnel is not guaranteed.
Risk Exposure: Critical personnel can become sick or otherwise become unavailable temporarily and, as a result, critical tasks get delayed.
Actions to Mitigate Risk: Assigning a supporting member to every responsibility will mitigate the shortfalls. Periodic walkthroughs and in-process reviews would be conducted to assess the progress of work. This would assist the primary responsible member to complete the task.

Description: Communication breakdown between the developers and the customer.
Risk Exposure: Functionality does not meet specifications and requirements are not satisfied.
Actions to Mitigate Risk: The Project Manager conveys the customer's understanding and requirements of the desired product to the development group. In turn, the Project Manager informs the customer of the developers' technical limitations and system limitations.

Description: Unbalanced work allocation.
Risk Exposure: If there are too many people working on a particular area of the system, then other areas may be neglected.
Actions to Mitigate Risk: Group members' strengths need to be utilized and, at the same time, the entire scope of the project needs to be covered so that maximum production can occur. Ensure all group members are working in parallel on different tasks.

Description: The Feasibility Rationale Description does not cover all issues.
Risk Exposure: Scrap is generated requiring significant rework.
Actions to Mitigate Risk: Immediately document any potential risks so that they are not forgotten.

4.2 Schedules and budgets

Description: Project is not completed in 12 weeks.
Risk Exposure: Fail to deliver on time or deliver untested functionality.
Actions to Mitigate Risk: Descope project to deliver guaranteed set of functionality in given time.

Description: Operators not be hired in time to begin the data entry.
Risk Exposure: This would push back the entire schedule and the live date.
Actions to Mitigate Risk: Plan for staffing qualified Operators before data entry stage.

4.3 COTS and external components

Description: Using right COTS for the wrong reason.
Risk Exposure: COTS product would introduce the loss of controllability of the system or may not provide critical features required for the system.
Actions to Mitigate Risk: In order to prevent such events, early initiatives would be taken to explore alternatives and perform an analysis into the suitability of those products. Those COTS tools that are likely to affect performance would be benchmarked against other available options.

Description: Using the wrong COTS package is another potential source of risk.
Risk Exposure: Increased development time, maintenance complexity and costs.
Actions to Mitigate Risk: Do not try to force COTS to work for wrong objective.

4.4 Requirements mismatch, Quality mismatch

Description: Mismatch in the understanding of requirements by the various stakeholders.
Risk Exposure: Loss of customer and user satisfaction and schedule slippage.
Actions to Mitigate Risk: Use prototyping to specify the project requirements. Quality mismatches would be avoided by providing the users with early prototypes.

Description: Developed system is not scalable.
Risk Exposure: All archive materials are not archived.
Actions to Mitigate Risk: Architect and design for scalability from the start.

Description: Project goals and responsibilities are not met.
Damage: An inadequate and unwanted system is produced.
Actions to Mitigate Risk: The main objective of the system needs to be placed into perspective at all times. The number one goal is to solve a particular problem and a particular need.

4.5 User interface mismatch

Description: Inadequate user interface.
Risk Exposure: Users find interface difficult to use.
Actions to Mitigate Risk: Various types of users would have to be identified and profiled while designing the interface. These would then be used for verification of system capabilities. A prototype would be developed and presented to the potential users with feedback obtained from them before progressing into development of complete functionality. Designs would adhere to platform specific interface standards by using GUI frameworks and foundation classes.

Description: Prototype does not demonstrate ease of use.
Risk Exposure: Designed system requires high learning time.
Actions to Mitigate Risk: Because the prototype is a representation of the actual system, the prototype must be geared for the same type of users, i.e. possessing little computer experience.

Description: An easy-to-use user interface leads to an insecure system.
Risk Exposure: Archive items and information damaged.
Actions to Mitigate Risk: A separation of the Administrative and User subsystem is an option to minimize security breaches. Plan for regular backup of archived information.

4.6 Architecture inadequacy

Description: Rigid architecture makes adjustments difficult.
Damage: An architecture that cannot meet all the requirements is certain to lead to breakage and rebuild.
Actions to Mitigate Risk: A flexible architecture allows ease of change and extensibility. The project would use a layered architecture, as it is suitable for this purpose. Domain experts and architecture experts would be consulted to assess the architectural suitability early enough.

Description: Architecture is not designed to ensure security of the system.
Risk Exposure: The integrity of the archive is damaged.
Actions to Mitigate Risk: Only administrators will be able to perform data modification functions.

4.7 Requirements changes

Description: Customer constantly gets new ideas and requests them in the product.
Risk Exposure: Time overrun and a lot of rework can also impact the performance of the system.

Actions to Mitigate Risk: Build a flexible architecture to more easily adapt to changing requirements. Addition of changes and requirements to the system would be controlled. Changes requiring large amount of work would be postponed to further releases.

Description: Changing requirements
Risk Exposure: A lot of wasted effort and rework.
Actions to Mitigate Risk: Each requirement must be analyzed careful to ensure accuracy and correctness. This includes aspects of the system architecture.

4.8 Straining computer science skills

Description: Algorithms and technologies require advanced or unproven techniques.
Risk Exposure: Schedule slippage and an untestable system.
Actions to Mitigate Risk: Those requirements that need such straining skills and knowledge would be culled apart from other requirements. A cost-estimate analysis would be conducted to evaluate the feasibility of providing such features. If possible, available software would be reused after due licensing.

5 Analysis Results

The design of the Hispanic Digital Archive could have resulted in many variations all having the same goal. Many of our decisions were based on resource, cost, and feasibility factors. We tried to limit ourselves to using only ISD provided software for it is more readily available, free for the customer, and fully supported. Also, many software products were chosen because either one or several group members were already familiar with that product. The decision to use certain software tools during the implementation stage to cut down on learning time was made based on the members enrolled in CSCI 577b. The decisions to use products based on familiarity will not compromise the quality of the HDA because the feasibility of the product's usefulness is always the top priority.

So, the design of our system did not occur haphazardly because resource, cost, and feasibility factors all contributed to the architecture we have structured and the methodology we are following.

5.1 Architectural Issues

We decided to go for a multi-tier client-server architecture after considering several other options. A 3-tier system consists of a front-end client, a back-end server, and a DBMS. Using a 3-tier architecture would allow the system to expand and later provide service to other types of clients. Clients would request and supply data to the server transparent of what DBMS is underneath the server. This would allow the clients to operate normally even with the introduction of new data types or other data schema changes in the DBMS. We considered using a 2-tier system with only a front-end client communicating with a back-end DBMS using Java Database Connectivity (JDBC). Using a 2-tier design would mean that there would be no server application to take requests and perform functions on the DBMS. All DBMS activity would be performed directly from the client. The client would, therefore, need to have specific details about the structure the DBMS. This would cause the client to shift away from the "thin-client" paradigm. "Thin clients" allow applets to load faster and execute faster. Incorporating a server application hides much of the complexity of the DBMS and gives the client the ability to perform simpler, cleaner requests.

Using the above rationale led us to the decision of using the IBM Digital Library. We discovered that many of our requirements and functionality could be addressed by using IBM DL. This provides a complete solution to the Hispanic Digital Archive because it includes a DB2 DBMS, provides dynamic Web page creation, provides scalability, incorporates security and material protection methods such as watermarking, and offers utilities for customization of data entry and retrieval. One disadvantage of using IBM DL is that developers must adhere to certain methods of operation and coding specific to the application such as using the CGI-similar Net.Data and using its DDO objects. Also, the server applications and DBMS are not very portable. IBM DL can only run on IBM AIX and Microsoft Windows NT. IBM DL's DBMS can either use Oracle DBMS or IBM DB2.

5.2 Prototype

Our prototype effectively demonstrates the functionality of the system from the Users' point of view. A main focus of the graphical user interface (GUI) as suggested by Barbara Robinson (Customer) is appearance and visual attractiveness. Not only does the prototype and the resulting system have to have a clean, clear-cut design, the look and feel of the prototype must be appealing to users who are assumed to have little computer experience. The functionality, usability, and actions performed on the prototype will be further tested.

We will be demonstrating the functionality and verifying the acceptability of the user end of our prototype on a Windows NT/95 platform as well as a Macintosh platform. This is most suited for presentation since the client is familiar with the Macintosh operating system and the classrooms have PCs running Windows. The back-end applications only need to be functional on Windows platforms since multi-platform support is not necessary in this case.

5.3 Commercial-Off-The-Shelf Solutions

The following is a list of commercial off-the-shelf products which we considered using for the Hispanic Digital Archive.

DBMS:
We have chosen to use IBM Digital Library over some of the other popular vendor software applications for reasons stated in the Architectural Issues. Basically, IBM Digital Library offers a more complete solution for our Hispanic Digital Archive and, also, because it is provided to us free of charge by ISD. Some other products which were considered include:
Informix ORDBMS
Oracle 8

Middleware:
The use of middleware was a major decision we had to determine. Because we are using IBM Digital Library, we did not see the need to use CGI or CORBA middleware. Instead we will be using Net.Data, which is provided by IBM Digital Library. Some of the middleware products we were considering are as follows:
Inprise VisiBroker for Java (CORBA)
Iona OrbixWeb for Java (CORBA)
ISD Java/CGI scripts

Miscellaneous:
We have decided to use AccuSoft ImageGear 98 ActiveX-LT 16/32 for image conversion and scanner software. This supports the TWAIN interface which scanners use and also supports the conversion of many image types such as TIFF, JPEG, GIF, and BMP. ISD does not support this software so it will have to be purchased.

Future software:
For future functionality of the Hispanic Digital Library, some addition COTS will be needed. These COTS have yet to be determined but will be needed for the following extended capabilities:
Voice recognition software
Jukebox controller software
User profile management