2004

An online repository directory for the Society of California Archivists

LeeAnn Rippetue Erwine
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Erwine, LeeAnn Rippetue, "An online repository directory for the Society of California Archivists" (2004). *Master's Theses*. 2643.

DOI: <https://doi.org/10.31979/etd.k92n-ryd2>
https://scholarworks.sjsu.edu/etd_theses/2643

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

**AN ONLINE REPOSITORY DIRECTORY
FOR THE
SOCIETY OF CALIFORNIA ARCHIVISTS**

**A Thesis Presented to
The Faculty of
The School of Library and Information Science
San José State University**

**In Partial Fulfillment
of the Requirements for the Degree
Master of Library and Information Science**

**by
LeeAnn Rippetue Erwine**

August 2004

UMI Number: 1424514

Copyright 2004 by
Erwine, LeeAnn Rippetue

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 1424514

Copyright 2005 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2004

LeeAnn Rippetue Erwine

ALL RIGHTS RESERVED

APPROVED FOR THE SCHOOL OF
LIBRARY AND INFORMATION SCIENCE

Dr. Debra Hansen, Chair

Dr. Judith Weedman

Dr. Dan Fuller

APPROVED FOR THE UNIVERSITY

ABSTRACT

AN ONLINE REPOSITORY DIRECTORY FOR THE SOCIETY OF CALIFORNIA ARCHIVISTS

By LeeAnn Rippetue Erwine

The purpose of this thesis project was to develop a searchable online repository directory for the Society of California Archivists (SCA). An in-depth review of other archival organizations' Websites in both the United States and abroad justified the need for such a directory. This was followed by a study of possible methodologies for the directory. XML and XSLT were chosen as the most practical method for implementing a data file-driven online Website. Several steps were involved in the creation of the Website, including data processing, stylesheet and website design, and usability testing. The result was the creation of an initial repository directory containing 80 records now residing on the SCA's Website. On-going improvements and maintenance of the online repository directory will continue beyond the scope of this project as a service to the archival profession.

DEDICATIONS AND ACKNOWLEDGEMENTS

DEDICATIONS

In loving memory of my parents Lee and Lillian Rippetue
To Tom, Brittany, and Tori.

ACKNOWLEDGEMENTS

Among the many people who helped during the writing of this Master's Thesis I would like to thank Dr. Judith Weedman , Dr. Dan Fuller, and Dr. Geoffrey Liu for their valuable advise and constructive criticism. I would like to offer special thanks to Dr. Debra Hansen for the countless hours of editing and never-ending support without which this huge undertaking would never have been accomplished. Finally I would like to thank my family. Many times the work on this project took me away from them, yet they never failed to be patient and understanding.

TABLE OF CONTENTS

LIST OF FIGURES	viii
CHAPTER	
1. INTRODUCTION	1
Background of the Project	1
Scope of the Project	3
Description of the Current Directory	4
Critique of the Print Format	11
Benefits of an Online Directory	12
Conclusion	15
2. REVIEW OF ARCHIVAL SOCIETY WEBSITES	16
United States Websites	17
International Websites	23
Transitioning From Print to an Online Format	31
Recommendations for the SCA	32
Benefits to the Archival Profession	34
3. METHODOLOGY	36
Markup Languages	36
DBMS	45
Transitioning from a Print to an Online Directory	47
4. PROJECT IMPLEMENTATION	51
Data Processing	51

	Stylesheet Encoding	58
	Stylesheet Design	60
	Website Design	64
	Usability Testing	74
5.	CONCLUSION	80
	Discussion	81
	Recommendations	86
	Future Plans	89
	Significance of the Project	91
	REFERENCES	110
	APPENDIXES	
A.	Society of California Archivists Directory of Archival and Manuscript Repositories in California Data Sheet	92
B.	Example of a Filemaker Record	98
C.	Example of an XML Record	104
D.	XML DTD and Commentary	105
E.	Stylesheet Encoding	107

LIST OF FIGURES

<i>Figure 1.</i> Sample main entry for archival repository	5
<i>Figure 2.</i> Sample page from Repository Name Index.	6
<i>Figure 3.</i> Sample page from County Index	7
<i>Figure 4.</i> Sample page from Type of Institution Index	8
<i>Figure 5.</i> Sample page from General Subject Index	9
<i>Figure 6.</i> Sample page from Notable Holdings Index	10
<i>Figure 7.</i> Archival society Websites in the United States	16
<i>Figure 8.</i> Sample PDF search screen	18
<i>Figure 9.</i> Search PDF screen showing search results	19
<i>Figure 10.</i> LA as Subject keyword search screen	20
<i>Figure 11.</i> LA as Subject search options	20
<i>Figure 12.</i> Indiana Archival and Historical Repositories search screen	22
<i>Figure 13.</i> International archival society Websites	24
<i>Figure 14.</i> ARCHON A-Z repositories screen	25
<i>Figure 15.</i> Search ARCHON screen with a JavaScript search interface	26
<i>Figure 16.</i> ARCHON county pull-down menu	26
<i>Figure 17.</i> Full-Text display of an ARCHON record	27
<i>Figure 18.</i> ARCHIVENET search screen	28
<i>Figure 19.</i> Repository directory format of archival Websites	29
<i>Figure 20.</i> Example of HTML and XML data and encoding	37
<i>Figure 21.</i> Example of an XML data file	38

<i>Figure 22.</i> XSLT tasks and results	40
<i>Figure 23.</i> Browser display of an XML file	42
<i>Figure 24.</i> Example of XSLT encoding	42
<i>Figure 25.</i> Example of HTML encoding	43
<i>Figure 26.</i> Example of a query	46
<i>Figure 27.</i> Data processing flow chart	52
<i>Figure 28.</i> XSLT encoding excerpt from repositories.xslt	59
<i>Figure 29.</i> First stylesheet (repsoritories.xslt) in browser view	61
<i>Figure 30.</i> Second stylesheet (results.xslt) in browser view	63
<i>Figure 31.</i> Main page (wireframemain)	65
<i>Figure 32.</i> Find by name page (wireframename)	66
<i>Figure 33.</i> Find by keywords page (wireframekeyword)	67
<i>Figure 34.</i> Find by location page (wireframelocation)	68
<i>Figure 35.</i> Find by repository type page (wireframetype)	69
<i>Figure 36.</i> Online Repository Directory Main Menu (demomain)	70
<i>Figure 37.</i> Repository/Parent name search page (demoname)	78
<i>Figure 38.</i> Location search page (demolocation)	79
<i>Figure 39.</i> Repository type search page (demotype)	79
<i>Figure 40.</i> Format search page (demoformat)	80
<i>Figure 41.</i> Format list page (formatlist)	80
<i>Figure 42.</i> Holdings search page (demoholdings)	81

CHAPTER 1

INTRODUCTION

Trying to find up-to-date information about archival repositories in California is not an easy task for the average person, nor is it easy for the archival professional. This is because the last time a directory was published about California's repositories was in 1996, nearly eight years ago by the Society of California Archivists (SCA). It is easier, however, to find current information about repositories in other states and in other countries. This is because most archival organizations have discovered that using the World Wide Web is an effective method to manage and broadcast their repositories' information to a larger audience.

In order to keep up with the trend of using the World Wide Web for disseminating information about archival repositories, the SCA needed to develop an online repository directory. By doing this many problems of the previous print editions, such as outdated information due to time lags in publication, would be resolved. It was the purpose of this thesis to create a searchable online repository directory for the SCA. The thesis involved studying past SCA directories, current archival Website practices, and current online data managing methodologies. The thesis resulted in an online repository directory with features that were not only unavailable in any of the previous print editions, but were also not available in any other single archival Website.

Background of the Project

The 1st edition of the *Directory of Archival and Manuscript Repositories in California* in 1975 was a modest volume of 108 pages (Hansen, 2001). The 2nd edition,

along with the 1st, was a joint effort of the Society of California Archivists and the California State Archives. The 3rd edition was issued in 1991 and the 4th edition followed five years later in 1996. Since its humble beginnings in 1975, each volume has grown considerably expanding in scope, content, and access points. The most recent edition is nearly 500 pages and reports on the location and holdings of 1,030 repositories (SCA, 1996).

The SCA realizing that its repository directory was out of date conducted a preliminary investigation in 2002 into other formats used by other archival organizations. This was done in anticipation of revising and producing a 5th edition of the directory. The investigation revealed that in addition to the information in the repository directory being out of date, the format was out of date as well. A survey of archival directories nationwide found that only the Michigan Archival Association had a directory available solely in print. The rest of the organizations' directories were online. Exploring a little deeper, the SCA wanted to know how many printed archival directories were still for sale. The Books In Print Website showed 13 repository directories listed, but only 6 were still in print. Amazon.com listed about 38 archival directories; however, all but 4 were listed as out of print or had limited availability. Although it was is not a systematic nor totally unbiased survey, a trend clearly emerged that showed archival organizations are moving away from print in favor of the Internet to disseminate information. Based on these findings that the Internet is the preferred method of other archival associations to post information, the SCA Publications Committee came to the conclusion that print

directories were being replaced by online versions and recommended that the SCA seriously explore the option of having an online directory (Hansen, 2001).

This directory has been a useful tool for individuals and organizations interested in the preservation and usage of archival materials in California. Each entry contains information about each repository's physical location, contact information, and its holdings. Each successive volume strives to improve upon its predecessor in order to meet the needs of information seekers. The 4th edition, which contained five indexes, improved upon the 3rd edition, which contained three indexes. The five indexes in the 4th edition are: repository name, repositories by county, general subject, type of institution, and notable holdings. In addition to the added indexes, the 4th edition also contained more information in each entry such as FAX number, e-mail address, personal contact information, and URL address (SCA, 1996).

Scope of the Project

The scope of the thesis project was to create searchable Web pages by using XML for the data file and XSLT for the stylesheets. (The term "data file" will be used from now on to mean a flat, XML tagged text file of records, as opposed to "database" in the sense of a relational database.) Included in the process of creating the online directory was the construction of a wire-frame structure of the Web pages as well as a usability test to ensure that the online directory meets the needs of the both the researcher and the SCA.

Description of the Current Directory

In order to understand how a data file-driven online directory would benefit the SCA, it is first necessary to examine the current print directory. Information in the current 4th edition is organized into one main directory with five additional indexes for cross-referencing. These indexes are: Repository Name; County; Type of Institution; General Subject; and Notable Holdings. The main directory contains one full set of records for each repository and is organized into an alphabetical list of cities in California. Within each city, repositories are alphabetized by their name. Figure 1 below shows some of the repositories in San Raphael. The LucasFilm, Ltd. Research Library entry is a complete record and has the following fields: address, county name, and telephone and fax numbers. This is followed by a contact person, hours of operation, type of institution, general subjects, holdings, formats, copying permitted and, lastly, other fees.

<p>275</p> <p>San Rafael</p> <hr/> <p>LUCASFILM, LTD. RESEARCH LIBRARY*</p> <p>Mailing Address P.O. Box 2009, San Rafael 94912</p> <p>County: Marin</p> <p>Tel. (415) 662-1911 or (415) 662-1512</p> <p>Fax: (415) 662-1628</p> <p>CONTACT: Deborah Fine, Director, Jo Donaldson, Researcher</p> <p>OPEN: Library open by appointment only to film production personnel on a fee basis</p> <p>TYPE OF INSTITUTION: Corporate</p> <p>GENERAL SUBJECTS: Performing arts and entertainment (photography)</p> <p>HOLDINGS: Paramount Studios Research Library, photographs and films of LucasFilm productions</p> <p>FORMATS: Books, clippings, over 400 vertical file drawers of photographs, and thousands of films/videotapes.</p>	<p>COPYING PERMITTED? Yes, facilities available 10 cents per page.</p> <p>OTHER FEES: Non-members pay a \$5 per day use fee.</p> <p>MARIN COUNTY LIBRARY. ANNE KENT CALIFORNIA ROOM</p> <p>Civic Center, San Rafael 94903-4177</p> <p>County: Marin</p> <p>Tel: (415) 499-7419 or (415) 499-6051</p> <p>Fax: (415) 499-3726</p> <p>CONTACT: Jocelyn A. Moss, Senior Library Assistant</p> <p>OPEN: Mon, Wed and Fri, 10-6; Tues and Thurs, 10-9; Sat, 10-5</p> <p>TYPE OF INSTITUTION: County Library</p> <p>VOLUME: 15 in. ft.</p> <p>INCLUSIVE DATES: 1850 to present</p> <p>GENERAL SUBJECTS: California and the West, and local history (including architecture.</p>
--	---

Figure 1. Sample main entry for archival repository.

Information about the LucasFilm, Ltd. Research Library can only be found in the main directory under the city of San Raphael. If the information seeker does not know the name of the city the repository is in, he or she must consult one of the indexes to find it. If the name of the repository is known, the researcher can go to the Repository Name Index to find their repository. Page 362 in the Repository Name Index refers the information seeker back to LucasFilm, Ltd. on page 276.

362

Repository Name Index

Lodi Public Library	102	Los Angeles Times	
Logicon, Inc.	274	Editorial Library	121
Loma Linda University, Del E. Webb		History Center (Inactive)	122
Memorial Library	102	Los Angeles Zoo	114
Lompoc City Clerk's Department	103	Los Cerritos Historic Sals. Rancho	107
Lompoc Public Library	104	Los Fundadores, the Founders and Friends of	
Lompoc Valley Historical Society	104	Santa Clara County	286
London (Jack) Research Center, Library	78	Los Gatos Public Library	142
London (Jack) State Historic Park, Research		Los Osos Branch Library (San Luis Obispo	
Center	78	County Library)	142
Long Beach Historical Society	105	Los Padres National Forest	80
Long Beach Press Telegram	105	Louis B. Mayer Library, American	
Long Beach Public Library	105	Film Institute	106
Longshoremen's and Warehousemen's Union,		Louis E. Stockmeir Regional History	
International	246	Library, De Anza College	52
Los Alamitos Historic Site and Gardens, Rancho	106	Louise M. Darling Biomedical Library,	
Los Altos Public Library	107	University of California, Los Angeles	130
Los Angeles City Archives	112	Love (Malcolm A.) Library, San Diego	
Los Angeles City Clerk's Office,		State University	235
Los Angeles City Archives	112	Lowie (Robert H.) Museum of Anthropology,	
Los Angeles City Department of Recreation		University of California, Berkeley	24
and Parks, Hollywood Museum Collection	117	Loyola Law School, William B. Rains	
Los Angeles County Arboretum	5	Library, Special Collections	122
Los Angeles County Metropolitan		Loyola Marymount University,	
Transportation Authority	118	Charles Von der Ahe Library	122
Los Angeles County Museum of Art,		Lucasfilm, Ltd.	276
Costumes and Textile Archives	118	Lucretia Little History Room, Mill Valley	
Los Angeles County Natural History		Public Library	150
Museum, Seaver Center for Western		Lutheran Church in America, Evangelical,	
History Research	125	Region 2, Archives Center	17
Los Angeles County Public Library,		Lutheran Church-Missouri Synod,	
American Indian Resource Center	87	California-Nevada-Hawaii District	247
Asian Pacific Resource Center	154		

Figure 2. Sample page from Repository Name Index.

LucasFilm, Ltd. can also be found on page 385 in the County Index for Marin County and on page 398 in the Type of Institution Index. It is listed on page 448 in the General Subject Index under "Performing Arts and Entertainment" as well. Examples of the latter three indexes are shown in Figures 3-5.

County Index		385
MADERA		
Cities represented: Madera, North Fork, and Oakhurst.		
Madera County Historical Society and Museum	143	
Madera County Library		
North Fork Branch	153	
Oakhurst Branch	157	
Sierra Historic Sites Association		
Fresno Flats Historical Park	157	
MARIN		
Cities represented: Mill Valley, Novato, Point Reyes, San Quentin, San Rafael, Sausalito, Tiburon, and Tomales.		
Belvedere Tiburon Landmarks Society	315	
California State Department of Parks and Recreation, Marin District	275	
Fireman's Fund Archives	166	
Golden Gate Baptist Theological Seminary	150	
LucasFilm, Ltd.	276	
Marin County Historical Society	278	
Marin County Library, Anne Kent California Room	278	
Marin Independent Journal	156	
Mill Valley Public Library		
Lucretia Little History Room	150	
Novato History Museum	166	
Point Reyes National Seashore	187	
San Quentin Museum Association	275	
Sausalito Public Library	298	
Sisters of Saint Dominic, Congregation of the Most Holy Name	277	
Tomales History Center	316	
U.S. Army Corps of Engineers, Bay Model Visitor Center, Marinship 1942-1945	299	
Mendocino College	320	
Mendocino County Historical Society and Held-Posge Memorial Home	320	
Mendocino County Library	321	
Mendocino County Museum	332	
Mendocino Historical Research, Kelley House Museum	147	
Temple of Kwan Tai, Inc.	148	
MERCED		
Cities represented: Atwater, Los Banos, and Merced.		
Castle Air Museum Foundation	7	
Merced County Courthouse Museum	149	
Merced County Library	150	
Milken Museum	141	
MODOC		
City represented: Alturas.		
Modoc County Library	2	
Modoc National Forest	2	
MONO		
Cities represented: Bodie and Bridgeport.		
Bodie State Historic Park	32	
Mono County Free Library	34	
MONTEREY		
Cities represented: Aromas, Carmel, King City, Monterey, Pacific Grove, and Salinas.		

Figure 3. Sample page from County Index.

ACADEMIC (Continued)		CORPORATE	
University of California, Santa Cruz (Continued)		Amgen, Inc.	314
University Library, Special Collections	293	Automobile Club of Southern California	109
University of Judaism	134	Bank of America	238
University of La Verne	96	Bay Meadows Race Track	272
University of Redlands	196	Bockman Instruments, Inc.	73
University of San Diego	236	CTB / McGraw-Hill Library	154
University of San Francisco	260	Chevron Services	277
University of Southern California		Communications Institute	
Boeckmann Center for Iberian		Communications Library	243
and Latin American Studies	135	Contra Costa Times	327
Cinema-Television Library and		Disney (Walt) Archives	35
Archives of Performing Arts	135	Disney (Walt) Imagineering Information	
Hancock (Allan) Library of Biology		Research Center	78
and Oceanography	135	Educational Communications	
Hosoe Library of Philosophy	136	Environmental Resource Library	113
Johnson (Arlene) Social Work Library		Fireman's Fund Archives	166
California Social Welfare Archives	136	First American Title Insurance Company	278
Korean Heritage Library	136	Fresno Bee	70
Library, Department of Special Collections	137	Hayward Daily Review	83
Library, Department of Special Collections		Hewlett-Packard Company	178
American Literature Collection	137	Hughes Aircraft Company	116
Library, Department of Special Collections		Levi Strauss and Co.	247
Feuchtwanger (Lion) Memorial Library	138	Logicon, Inc.	274
Library, Department of Special Collections		Long Beach Press Telegram	105
Regional History Collection	138	Los Angeles Daily News	336
Library, University Archives	139	Los Angeles Times	
Music Library, Ingolf Dahl Archive	139	Editorial Library	121
Schoenberg (Arnold) Institute	139	History Center	122
School of Architecture, Gamble House		LucasFilm, Ltd.	276

Figure 4. Sample page from Type of Institution Index.

448	General Subject Index
PERFORMING ARTS AND ENTERTAINMENT (Continued)	
Del Norte County Historical Society and Museum	55
Disney Archives	35
Disney Imagineering Information Research Center	78
El Monte Historical Society	87
Fantasy Foundation	114
Fashion Institute of Design and Merchandising	114
Fulerton Public Library	75
Getty Center for the History of Art and the Humanities	295
Harrison Memorial Library	38
Hart (William S.) Museum	162
Hearst San Simeon State Historical Monument	277
Historical Society of Long Beach	105
Hollywood Bowl Museum	116
Hollywood Film Archive	85
Holy Names College	169
Humboldt State University Library Humboldt County Collection and University Archives	6
Inner City Cultural Center	85
International Longshoremen's and Warehousemen's Union	246
Kern County Library, Beale Memorial Library	12
Kern River Valley Historical Society and Museum	52
Library Association of La Jolla	94
Long Beach Public Library	105
Los Angeles City Department of Recreation and Parks, Hollywood Museum Collection	117
Los Angeles County Public Library Black Resource Center	118
Los Angeles Philharmonic Archives	118
Los Angeles Public Library Goldwyn Hollywood Regional Library Photograph Collections	120
Los Angeles Times Editorial Library	121
LucasFilm Ltd.	276
Reagan Presidential Library	
Riverside Municipal Museum	
Rogers (Roy) and Dale Evans Museum	
Roseville Public Library	
Sacramento Archives and Museum	
Sacramento Valley Museum	
San Bernardino Public Library, California Room	
San Diego Public Library, California Room	
San Diego State University Malcolm A. Love Library	
San Francisco Academy of Comic Art	
San Francisco Conservatory of Music	
San Francisco Performing Arts Library and Museum	
San Francisco Public Library	
San Francisco History Room and Archives	
San Francisco State University Department of Anthropology Adan E. Treganza Anthropology Museum	
Labor Archives and Research Center	
Leonard Library, Frank V. de Bellis Collection of the California State University	
Leonard Library, Special Collections/Archives	
San Jose State University Library Special Collections	
Steinbeck Research Center	
Santa Barbara Mission	
Santa Maria Valley Historical Museum	
Seaver Center for Western History Research, Natural History Museum of Los Angeles County	
Sierra Valley Historical Society and Museum	
Smithsonian Institution Archives of American Art	
Society of California Pioneers	
Stanford University Archive of Recorded Sound	
Libraries, Department of Special Collections	
Stockton-San Joaquin County Public Library	
Templeton Historical Museum Society	
Thousand Oaks Library, Special Collections American Library of Radio and Television	

Figure 5. Sample page from General Subject Index.

The General Subject Index shown above in Figure 5 is a controlled vocabulary comprised of the following terms:

Agriculture	Journalism	Printing
Archaeology and	Labor	Psychology
Anthropology	Law	Race/Ethnicity
Architecture	Literature	Religion
Art	Maritime	Science
Biology	Military	Science Fiction
Business and Economics	Mining	Sexuality
California and the West	Natural Resources and the	Social History
Computer Technology	Environment	Sports
Education	Oral History	Telecommunications
Engineering	Performing Arts and	Transportation

Foreign Nations
Genealogy
Geography and Mapping

Entertainment
Philosophy
Photographs and
Photography

United States History
Women

The last index, Notable Holdings, is also a controlled vocabulary. It does not list repositories by name but rather by terms that describes their holdings. One of LucasFilm's notable holdings is the Paramount Studios Research Library, which is listed on page 481 of the 4th edition simply as "Paramount Studios."

Notable Holdings Index			481
Orange County Transportation Authority, 174	Panama, 21, 283	Pickford, Mary, 30	
Orange County Transportation Commission, 174	Panama Canal, 253	Pico, Andres, 152	
Orange Empire Railway Museum, 183	Panama Pacific World's Fair, 70	Pico (Andres) Adobe, 152	
Oregon, 14, 40-41, 56, 159	Paraguay, 205	Pico, Pio, 125, 332	
	Paramount Studios, 30, 128, 276	Piedmont, 184	
	Pardee, Enoch H., 171	Pinnacles National Monument, 177	
	Pardee Family, 171	Pitcairn Islands, 3	

Figure 6. Sample page from Notable Holdings Index.

If the information seeker wants to know which repositories have notable holdings related to Paramount Studios, he or she must browse through all of the repositories described on the pages indicated in the index to find the term "Paramount Studios." This is time consuming for the user. For example, there are two repositories listed on page 30. The term "Paramount" not "Paramount Studios" is found buried in a long list of the holdings belonging to the Academy of Motion Picture Arts and Sciences Library. Similarly, page 128 has three repository descriptions through which the user must browse. The term "Paramount Pictures," not "Paramount Studios," is found in the University of California, Los Angeles, Arts Library Special Collections. The LucasFilm, Ltd. Research Library described on page 276 has the exact term "Paramount Studios" in

the list of holdings, but there are also two other repositories listed on that page through which the information seeker must browse. This is very problematic for the researcher. Because the index only refers the researcher to a page rather than to a specific repository, time is wasted browsing the entire page.

Critique of the Print Format

The SCA Repository Directory has many other limitations. Because it is a print document, the researcher is limited to browsing the main directory and using the five indexes for searching for repository information. Timeliness is another problem the researcher has with the print directory. By the time a print directory is issued, some of the information has already become outdated. Furthermore, obtaining the latest edition can be a problem for an information seeker. A person who has ordered a print copy of the directory edition from the SCA Website will typically have to wait anywhere between 4 and 6 weeks to receive it in the mail.

Timeliness is also a matter with which the SCA must contend. A great deal of time and effort is needed to create a print directory. The LucasFilm repository appears (or is referred to) six separate times in the directory. For the person compiling and maintaining the directory this is a tremendous effort, not only in time but also in editorial effort. With each new index created, redundant information is retyped thereby increasing the chance of an error being made by the typist. This misinformation is ultimately passed on to the researcher.

Benefits of an Online Directory

Fortunately for the researcher, a data file-driven online directory is the solution for the limitations caused by the inherent nature of the print directory. An online searchable directory, designed with good navigation, is much faster and more flexible than the static print directory. More search options can be offered since the directory does not have a size limitation like that of a hard copy. While there are numerous possibilities of searches that could be done in the online directory, the following is a partial list of some of the new search options that might be added into the Website:

- List the parent institutions of the repositories in the database
- List the names of the contact personnel (alphabetized) and their job titles along with their phone numbers and URLs
- List all of the educational repositories in a given county or city

A data file-driven online directory, using a search interface such as JavaScript, would allow the researcher to enter a term in a dialog box and perform either a full-text keyword search on all records or perform a search from selected fields. These search options are far more flexible than the indexes of the print edition. For example, if the researcher performed a search on the term “Paramount Studios” in the notable holdings field in the data file, only the LucasFilm, Ltd. record would be retrieved. Similarly, if the researcher was not sure of the exact terminology, but knew the term was related to Paramount Pictures or Paramount Studios, he or she could perform a search on the word “Paramount” and retrieve all records with “Paramount” in the notable holdings field. This is still an improvement over the print index because it eliminates those records in the

print directory which have absolutely nothing to do with “Paramount” but happen to be on the same page as those repositories which do have the term “Paramount” in their notable holdings.

In addition to having greater searching capabilities in the online directory, the researcher will also have better access to the repositories themselves. This will be made possible by including hyperlinks to URLs and e-mail addresses in the repository description, which puts the researcher in direct communication with the repository by merely clicking on a hyperlink. Another possible feature that could be used in the online directory might be hyperlinks to MapQuest that would provide driving directions and a map so the researcher could find the physical location of the repository. Needless to say, this type of communication between the researcher and the repository was not possible before with the print directory. This is truly an innovative feature that is likely to make a researcher a repeat user of the online directory.

Just as the online directory would be faster and easier for the researcher to use than a print directory, it would also be faster and easier for the SCA to update and maintain. The issue of timeliness has always been a problem for the SCA in maintaining the print directory. It currently takes several years between gathering the data and publishing a new edition of the print directory. Because of this time lag, not only is much of the information in the 4th edition no longer current, it became outdated before the directory was even published. A data file-driven online directory would be an ideal solution to this problem because it could be updated at predetermined intervals or whenever deemed necessary. The task of updating an online data file throughout the year would be more

efficient and not be as overwhelming as the previous method of collecting and printing the information.

Redundancy is another problem that an online data file would solve. As stated earlier, the print directory is very time consuming to maintain because redundant information has to be retyped into each of the five indexes. The retyping of this information also allows for a higher chance of errors to be inadvertently made by the typist. This, however, would not occur in an online data file. All of the information about a repository would be contained in a single record making it much easier to update and maintain.

Another benefit the SCA would receive from having an online repository directory would be higher exposure within the archival community. Between November 4, 2002, and January 4, 2003, the SCA Website received 1,746 hits with 175 of those hits going to the publications page selling the repository directory. Interestingly, the SCA Website was not accessed solely from the United States and Canada, but also from Australia, the United Kingdom, Germany, Netherlands, Japan, Singapore, Spain, and Brazil (Carolyn Garner, personal communication, January 4, 1999). Based on this fact, it is of no doubt that an online directory would cause the SCA to receive much more exposure than is currently provided by the print directory. The greater exposure generated by the online directory, in turn, causes higher exposure of the smaller and lesser-known repositories listed in the directory.

Conclusion

A new edition of the SCA repository directory is long overdue. Publishing a new version in print is possible, though not practical. There are many disadvantages in maintaining the repository directory in a print format. Not only is it very difficult to keep information current in print, but also it is also difficult to disseminate it to a large audience. On the other hand, there are many advantages in creating an online repository directory. Not only would it be easier for keeping information current, it would also be accessible to a larger audience. Other archival societies in the United States have already realized this fact and have turned to the World Wide for posting information about repositories in their jurisdictions. Chapter 2 will explore how other archival societies are using the World Wide Web to get a better understanding of how an online directory might be created.

CHAPTER 2

REVIEW OF ARCHIVAL SOCIETY WEBSITES

In order for the online repository directory to have the most efficacy, it is helpful to examine how other archival organizations within the United States as well as internationally use their Websites to provide repository information. In the United States, there are a number of archival organizations on the World Wide Web whose jurisdictions range from the national level all the way down to the city level. A list of archival websites in the United States appears in Figure 7. While no two Websites are alike, many share similarities in their method of displaying repository information.

National Level:

ArchivesUSA

<http://archives.chadwyck.com/repository/search>

Regional Level

Society of Rocky Mountain Archivists

<http://ahc.uwyo.edu/srma/publications.htm>

Society of Southwest Archivists

<http://info.lib.uh.edu/ssa/ssa.htm>

Society of Northwest Archivists

<http://osulibrary.oregonstate.edu/archives/misc/nwa.html>

State Level

Society of Alabama Archivists

<http://www.auburn.edu/sala/>

Society of Florida Archivists

<http://www.florida-archivists.org/>

Society of Georgia Archivists

<http://www.soga.org/resources/gar.html>

Society of Indiana Archivists
<http://archives1.archives.nd.edu/sia/guide/reposito.htm>

Louisiana Archives and Manuscripts Association
<http://nutrias.org/lama/guide/guide.htm>

Michigan Archival Association
<http://maa.hypermart.net/index.htm>

Society of North Carolina Archivists
<http://www.rtpnet.org/~snca/>

City Level

Archivists Round Table of Metropolitan New York
<http://www.nycarchivists.org/guide/rguide.html>

Greater New Orleans Archivists
<http://nutrias.org/gnoa/norepos.htm>

ARC - LA [Los Angeles] as Subject
<http://www.usc.edu/isd/archives/arc/lasubject/index.html>

Figure 7. Archival society Websites in the United States.

United States Websites

Of the 16 United States Websites listed in Figure 7, all but 4 (the Society of Alabama Archivists, the Society of Tennessee Archivists, the Society of Northwest Archivists, and the Society of Southwest Archivists) offer information about archival repositories either online, in print, or a combination thereof (SAA, 2003; SNA, 2003; SSA, 2003; and STA, 2003). Two Websites in Figure 7 have restricted access to their information. ArchivesUSA, whose target audience is larger institutions, requires an annual subscription (ArchivesUSA, n.d.). The Society of Rocky Mountain Archivists limits access to their print directory of regional archives to its members (SRMA, 2001). The Society of Indiana Archivists, the Michigan Archival Association, and the Society of

North Carolina Archivists also have print directories for sale to the public, but provide online information about their repositories as well (MAA, 2003; SIA, n.d.; SNCA, n.d.).

The majority of Websites with repository information used HTML to display the data. However, two organizations in Figure 7 chose other methods. The Society of Florida Archivists used PDF, a file format created by Adobe, initially to provide a standard form for storing and editing printed publishable documents. The PDF directory is arranged similar to the SCA's printed repository directory, with a primary sort by city and a secondary sort by repository name. The PDF document is WYSIWYG (what you see is what you get) by nature. It is static, appearing the way in which it was typed. This format is very limiting for a researcher. While PDF is capable of searching for a word or phrase, it is not the most efficient method for finding repository information. If a researcher wants to search the Society of Florida Archivists' PDF directory, he or she would enter a word or phrase in the dialog box illustrated below.

Search PDF Hide

What word or phrase would you like to search for?

☐ Whole words only

☐ Case-Sensitive

☐ Search in Bookmarks

☐ Search in Comments

Search

Figure 8. Sample PDF search screen.

The results from the search are displayed in a list. The researcher could click on one of the results in the list and be taken directly to the place in the document where the word or

phrase resides. For example, a search on the word “miami” found 40 instances of its use in the PDF document.

Figure 9. Search PDF screen showing search results.

By looking at the results for the word “miami,” it is obvious this is not very helpful for a researcher because the results appear out of context. This might be acceptable if there were not a large number of results returned, but in this instance the search is cumbersome and time consuming. A PDF directory would not be a viable option for the SCA; especially if one considers that the Florida Society of Archivists’ directory is only 31 pages long (SFA, 2003). The SCA repository directory is 10 times that size.

A second method used by archival organizations is a JavaScript search interface. For example, the Archival Research Center (ARC) in Los Angeles uses this interface in a Website entitled LA as Subject. This site that allows information seekers to perform a

keyword search or select from a list of types of institutions as predetermined by the encoder.

Keyword Search

OR

Institutions

[Additional Search Options](#)

Figure 10. LA as Subject keyword search screen.

A second search interface allows the user to select type of institution:

- Archive (95)
- Archive (95)
- Artist (7)
- Artist (7)
- Charitable Organization (2)
- Church/Mosque/Synagogue/Temple (9)
- College/University (42)
- Commercial Business (6)
- Community Arts Organization (26)
- Community Foundation (1)
- Community Organization (2)
- Community Service Organization (34)

Figure 11. LA as Subject search options.

The additional search options allow the researcher to search on holdings types, languages, and time periods. While this method is not being used as much by archival organizations in the United States, it is being used by some organizations internationally. This type of search interface is much more effective than the PDF keyword search and would be suitable for the needs of the SCA (LA as Subject, 2003).

For those Websites using simple HTML, the data was displayed in one of two ways. One popular approach was to provide a list of hyperlinks to the URLs of the repositories, sorted alphabetically by repository name. While this is a simple and fast approach to

display information, it is only good for displaying those repositories with URLs. It is then the responsibility of the information seeker to click on the hyperlink and search each repository's Website to obtain the desired information. This approach does not work for those repositories not yet on the World Wide Web. The organizations using this method are: the Society of Georgia Archivists (SGA), the Michigan Archival Association (MAA), the Society of North Carolina Archivists (SNC), and the Society of Ohio Archivists (SOA) (MAA, 2003; SGA, n.d.; SNC, n.d.; SOA, n.d.).

While hyperlinks to those repositories with URLs will be an important feature of the SCA's online repository directory, they cannot be the sole method for providing repository information. The major drawback with relying on URL hyperlinks is that each Website displays information differently; therefore, descriptions are not standardized. While creating URL hyperlinks is quick and easy for the HTML encoder, it is problematic for the researcher who will have to navigate through each Website to find information. The SCA's online directory, on the other hand, will offer both standardized information to the researcher along with URL hyperlinks. This will benefit those repositories that do not have Websites and would be otherwise excluded from the directory. It also gives the researcher the option of using the hyperlink if he or she is unsatisfied with the information already contained in the online repository directory description.

Some organizations provided more in-depth information about repositories (such as location, contact information, and hours of operation) directly from their Websites. By encoding all of their data within HTML tags, all repositories can be included on the

Website, not just those with URLs. This method is more labor intensive for the maintenance of the Website because it requires more encoding, but it is easier for the information seeker because all data is displayed in a standard format. The organizations using this method are: the Society of Indiana Archivists (SIA), the Louisiana Archives and Manuscripts Association (LAMA), the Archivists Round Table of Metropolitan New York (ART), and the Greater New Orleans Archivists (GNOA) (SIA, n.d.; LAMA, 1998; ART, 2003; GNOA, 2003). This is a far better method than simple URL hyperlinks because it does not exclude those repositories without URL addresses. In a similar fashion, the SCA Website will also include those repositories without URL addresses. One interesting way some archival Websites used HTML was to make their directory appear to be a database. For example, the following is from the Society of Indiana Archivists' Website. It contains list of counties in Indiana.

Figure 12. Indiana Archival and Historical Repositories search screen.

The researcher can select a county from the list, which takes him or her to a page with an alphabetized list of cities for that county. Then under each city name is an alphabetized list of repository names. This, however, is not a database, but rather 85 separate Web pages created by the encoder with hyperlinks to represent each county. This is a

monumental task to create and maintain. The alphabetizing that occurs on this Website is not an automated feature; it is manually done by the encoder. Each time a new repository is added the encoder must manually insert it in the correct alphabetical location within the HTML code (SIA, n.d.).

Because this Website requires such high maintenance, it does not include any other search options. There is no way for a researcher to locate a repository by its type or by its name. In order to find a repository in a particular city, the researcher must first know the county in which that city is located. Clearly, this is very limiting and not very effective for the researcher. Although this Web page has the look and feel of a database, it is not one. An actual database or data file would be far easier to maintain. Alphabetizing would be automated and not manually done by the encoder. Furthermore, data could be presented in a variety of ways because it would not have the limitations of HTML encoding. For this reason, HTML alone is not the solution for the SCA's repository directory.

International Websites

For comparison purposes, ten international archival Websites were also examined to see how they differed from those in the United States. A list of these Websites appears in Figure 13. Like archival societies in the United States, most international archival societies provide online information about their repositories. Likewise, they also use HTML in the same manner as the United States. However, more international societies use JavaScript as a search interface, thereby making their Websites more interactive for the information seeker. The organizations using a JavaScript for a keyword search

interface are: ARCHON in the United Kingdom and Archivenet in the Netherlands/Flanders (ARCHON, 2002; Archivenet, 2002).

Australia

Australian Society of Archivists

<http://www.archivists.org.au/>

Brazil

AAB (Associacao dos Archivistas Brasileiros)

<http://www.aab.org.br/>

CONARQ

<http://www.arquivonacional.gov.br/conarq/qsomos/Listarq.htm>

Canada

BCAIN (British Columbia Archival Information Network)

<http://aabc.bc.ca/aabc/bcain.html>

Europe

European Archival Network

<http://www.european-archival.net/>

Israel

IAA (Israel Archives Association)

<http://www.tau.ac.il/humanities/iaa/index.html>

Netherlands / Flanders

Archivenet

<http://www.archiefnet.nl/index.asp?taal=en>

Norway

National Archives of Norway

<http://www.riksarkivet.no/english/about/contact.html>

Spain

ANABAD

<http://www.anabad.org/>

United Kingdom
ARCHON
<http://www.hmc.gov.uk/archon/archondirectory.htm>

Figure 13. International archival society Websites.

The ARCHON Website allows the researcher to do two types of searches. One option is to select a country within the United Kingdom. After the researcher selects a country, an alphabetized list of repository names for the entire country is shown.

Figure 14. ARCHON A-Z repositories screen.

However, the user also has the option of selecting “Search ARCHON.” This is an interactive search using JavaScript for the interface. The researcher has the option of searching on ARCHON reference number, repository name, town, or county.

Search ARCHON

Enter you search criteria and select the Search button.
Please enter whole words as acronyms will not work. N.B:
If an ARCHON reference is entered only this field will be
searched.

ARCHON reference:

Repository name:

Town:

County:

Figure 15. Search ARCHON screen with a JavaScript search interface.

Quite helpful is the county dialog box that has a pull-down menu of county names (ARCHON, A-Z Repositories, 2002).

Aberdeenshire
Anglesey
Angus
County Antrim
Argyllshire
County Armagh
Ayrshire
Bedfordshire
Berkshire
Berwickshire

Figure 16. ARCHON county pull-down menu.

One noteworthy feature of ARCHON is that each record has a unique number (primary key). This is important in a database; by making each record unique it is instantly identifiable for the programmer when making revisions to the database. The following figure is a complete ARCHON record, showing all of its fields including the ARCHON reference code (ARCHON, Search ARCHON, 2002).

ARCHON DIRECTORY

Archon Code : 0336
Lancaster University Library

Contact Details

Lancaster University
 Bailrigg
 Lancaster LA1 4YH
 England

Tel 01524 592535
 Fax 01524 63806

Email library@lancs.ac.uk

URL <http://libweb.lancs.ac.uk/>

Librarian : Ms JM Whiteside

Access Information

- Open : M,Tu,Th,F 8.45-21.45; W 9.30-21.45; Sat 10-16.45
- Readers ticket required
- Book in advance

Figure 17. Full-text display of an ARCHON record.

This record also has hyperlinks to e-mail and URL addresses. This is a good example of how a JavaScript interface, combined with a controlled vocabulary (city names, county names, repository names), allows the user some flexibility when searching for a record while, at the same time, presents the data in a standardized format. An interface similar to this will be included in the SCA's online repository directory (ARCHON, ARCHON Code 0336, 2002).

Another Website that also allows the researcher a few options from its main page is Archivenet in Netherlands/Flanders. The researcher can select a region, then either a town or province within that region, and then, finally, the letter of the alphabet with which the name of the town or province begins. The researcher can also perform a keyword search from a dialog box using a JavaScript search interface. This appears to be

a keyword search on repository names, although it is not clear since the site is not entirely in English. However, there are no instructions for the researcher on the Website, which give guidelines for performing a keyword search.

Figure 18. ARCHIVENET search screen.

This is an effective Website and more flexible than most of the archival Websites in the United States. It also has its limitations. There is no alphabetized list of repositories in the entire region. One must go through each letter of the alphabet to compile a full list of repositories. Repositories also are not listed by their institutional type (Archivenet, n.d.).

To get an overall view of the numerous ways repository information is presented a categorized list of common formats used by archival organizations appears in Figure 19. While this is by no means a list of every archival organization, it does help to point out similarities and differences among organizations. In summary, most archival organizations in the United States and abroad use their Websites to provide online information about repositories in their jurisdiction. While most organizations in the

United States use HTML without any other search interface, some European archival organizations are turning to JavaScript for an interactive search interface.

No Repository Directory (Print or Online)

Society of Alabama Archivists: <http://www.auburn.edu/sala/>

Society of Tennessee Archivists: <http://www.geocities.com/tennarchivists/>

Society of Northwest Archivists:
<http://osulibrary.oregonstate.edu/archives/misc/nwa.html>

Society of Southwest Archivists: <http://info.lib.uh.edu/ssa/ssa.htm>

Asociacion Espanola de Archiveros – ANABAD: <http://www.anabad.org/>

Print Directory

Society of Rocky Mountain Archivists:
<http://ahc.uwyo.edu/srma/publications.htm>

Only available to members

Society of California Archivists:
<http://www.calarchivists.org/>

For sale to public

PDF Directory

Society of Florida Archivists:
<http://www.florida-archivists.org/>

Sorted by city/repository name

HTML Directories

Unknown Encoding

ArchivesUSA: Available by subscription only, encoding unknown
<http://archives.chadwyck.com/repository/search>

Hyperlinks to repositories

Society of Georgia Archivists:
<http://www.soga.org/resources/gar.html>

Sorted by repository name

Michigan Archival Association:
<http://maa.hypermart.net/index.htm>

Sorted by repository name
Google search (does not work),
JavaScript search (does not work)
*Also has print directory

Society of North Carolina Archivists:
<http://www.rtpnet.org/~snca/>

Sorted by repository name

Society of Ohio Archivists:
<http://www.ohiojunction.net/soa/links.html>

Sorted by repository name

IAA - Israel Archives Association:
<http://www.tau.ac.il/humanities/iaa/index.html>

Sorted by repository name

AAB – Associacao dos Archivistas
Brazileiros :
<http://www.aab.org.br/>

No apparent sort order

<i>Repository information provided within Website</i>	
ARC - LA as Subject : http://www.usc.edu/isd/archives/arc/lasubject/index.html	JavaScript 1. Keyword search 2. Repository type search
Archivenet: http://www.archiefnet.nl/index.asp?taal=en	Location sort (HTML) JavaScript Search Interface Keyword search
Archivists Round Table of Metropolitan New York: http://www.nycarchivists.org/guide/rguide.html	1. Repository name sort 2. Subject index
ARCHON: http://www.hmc.gov.uk/archon/archondirectory.htm	JavaScript/ASP Active Server Pages Map Search Country and city search Repository name search
Australian Society of Archivists: http://www.archivists.org.au/	1. Repository name sort 2. Location Sort Google Search Interface
BCAIN (British Columbia Archival Information Network): http://aabc.bc.ca/aabc/bcain.html	1. Repository name sort 2. Repository type sort 3. Location sort 4. Map search Other Search Interface
CONARQ – Conselho Nacional do Arquivos: http://www - arquivonacional.gov.br/conarq/qsomos/Listarq.htm	Sorted by location
European Archival Network: http://www.european-archival.net/	JavaScript 1. Keyword search 2. Map search 3. Repository name sort
Greater New Orleans Archivists: http://nutrias.org/gnoa/norepos.htm	Sorted by repository name
Louisiana Archives and Manuscripts Association: http://nutrias.org/lama/guide/guide.htm	City/repository name sort
National Archives of Norway: http://www.riksarkivet.no/english/about/contact.html	No apparent sort order
Society of Indiana Archivists: http://archives1.archives.nd.edu/sia/index.html	1. County/city/repository name sort 2. Repository name sort (does not work) *Also has print directory

GFigure 19. Repository directory format of archival Websites.

Transitioning From Print to an Online Format

Interestingly, some archival organizations use their Websites to discuss their transition to an online format. The Society of Indiana Archivists explained the process they underwent to create an online repository directory. According to the SIA, over 800 survey forms were mailed to every known archive, library, museum, historical society, and genealogical organization in the state. Over 300 repositories are listed in this 1st edition of their online repository directory. Most institutions and organizations that responded to the survey are included in the directory. A few institutions, whose responses indicated only artifact collections, were intentionally omitted (SIA, n.d.).

Just as some archival organizations in the United States used their Websites to discuss how they made the transition from print to the Web, Australia has done the same. In 1992 the Australian Society of Archivists (ASA) published the *Australian Directory of Archives*, which existed in word-processed format. In 1995 the ASA was approached by the Australian Science Archives with a proposal to create a World Wide Web version of the directory. The ASA accepted the proposal knowing that an online repository directory would be easier to maintain and would also gain international recognition. The first step in transitioning to an online format was to export the word-processed files into an Access database. Meanwhile the design of HTML Web pages occurred and revisions were made to the data as individual repositories provided updates for the Web version. Once the revisions were completed, MS Access software generated HTML pages. The online repository directory also included a new search interface and a section that

provided hyperlinks to World Wide Web resources. The *Australian Directory of Archives* made its public debut on May 31, 1996 (ASA, Introduction, 2003).

Other archival organizations use their Websites to share the origins of their online directories. Two city-level archival organizations, the Archivists Round Table of Metropolitan New York (ART) and the Greater New Orleans Archivists (GNOA), first used brochures to publish information about their cities' repositories. A list of repositories in the City of New Orleans went online in 1996, with New York following in 1998 (ART, 2003; GNOA, 2003). A statewide organization, the Louisiana Archives and Manuscript Association (LAMA), went online as a response to the publication of the Society of Southwest Archivists (SSA) repository guide in 1993. The Society of Southwest Archivists overlooked several of Louisiana's smaller repositories. By going online, LAMA also felt that they could improve upon the SSA guide by including newer communication contact points such as e-mail addresses or Internet URLs (LAMA, 1998).

Recommendations for the SCA

After examining the archival Websites in this Review of Archival Society Websites, it is possible to make some recommendations for the SCA's online repository directory. Some of the methods used by other Websites are not good options for the SCA. For instance, PDF is a poor choice as the keyword searches are not very helpful and difficult to navigate and the large size of the SCA's online repository directory would make PDF searching extremely difficult for a researcher.

Another option that will not work is using HTML alone to display repository information. It is not powerful enough to maintain efficiently the number of records

required by the SCA's online repository directory. HTML also cannot perform the functions that can be done in a database or data file, such as multilevel sorting. Without this capability it would be difficult for a researcher to navigate SCA's directory. The large online repository directory would also be too labor intensive to be properly maintained in HTML. Therefore HTML as the sole method for the SCA's online repository directory is not practical.

However, there are some features of HTML-based archival Websites that should be incorporated into the SCA's online repository directory. Hyperlinks to URLs should be used in the SCA's online repository directory, though only as an enhancement. By including URLs as part of a standardized format, researchers can rely on the SCA's online repository directory to provide consistent information from the data file, while still being able to explore a repository's Website for information not necessarily provided in the SCA's online repository directory. A search interface such as JavaScript is another feature that should be included in the SCA's online repository directory that allows a full-text keyword search. This would allow the researcher access to the repositories based on the subjects of their collections. This searching feature is lacking in all of the other archival Websites.

Since there are over 1,000 repositories listed in the SCA repository directory, the most efficient method to use will be one that makes maintaining the large number of records easier. This can best be accomplished by storing the repository information in a data file such as XML. If the records are stored in XML and displayed using XSLT, the hyperlinks to URLs, e-mail addresses, and different search interfaces can be incorporated.

A more in-depth look at how XML and XSLT can work together for the benefit of the SCA's online repository directory is provided in the next chapter, "Methodology".

Benefits to the Archival Profession

If the SCA follows the above recommendations, the archival profession in the United States will benefit as a whole. Never before in the United States has an archival Website, at the state level, employed a database or data file to manage repository information online. Managing the repository data online will be easier and more efficient for the SCA to maintain and keep current. The researcher will benefit not only from having access to current data, but also from the reliability of standardized information made possible by a data file. Moreover, the researcher will not be limited to the standardized data format, but will also have the flexibility to click on hyperlinks to visit a repository's Website, visit the Online Archive of California, or to contact a repository via e-mail. Furthermore, the SCA Website will offer a wide number of pre-programmed search options in addition to the flexibility of keyword searches. By implementing the suggested recommendations, not only will the SCA's online repository directory benefit the researcher and the SCA but ultimately the individual repositories, since all of the new features promote greater communication between the user and the archives.

While many other archival Websites, domestically and abroad, have offered some of these features, no single Website has all of the features proposed for the new SCA online repository directory. The SCA's online repository directory offers a new and improved way for the SCA to maintain repository information and for the researcher to

search for repository information. There is no doubt that its of great significance to archivists and the archival profession.

CHAPTER 3

METHODOLOGY

Although HTML has been used by most archival organizations in the United States to construct their web pages, it is not the most efficient method for handling the large number of records in the SCA's online repository directory. There are two other methods that are better suited for the purposes of the SCA's online repository directory. The first method is to store the records in a database management system (DBMS). The other approach is to store the records in a flat data file such as XML. To understand why an XML data file or a DBMS are more efficient than HTML in managing the online repository directory, it is helpful to discuss the nature of HTML, XML, and a DBMS.

Markup Languages

The W3C (World Wide Web Consortium) describes HTML (Hypertext Markup Language) as the *lingua franca* for publishing hypertext on the World Wide Web. HTML uses a predefined set of tags to display text. For example, <h1> and </h1> are tags that structure text into headings. There are also tags that structure text into other formats, including paragraphs, lists, tables, and hypertext links, among others (W3 HTML Home Page, 2002). An HTML document contains all the data, along with all the tags that format it. Although HTML is the sole approach used by other organizations to list information about repositories, it is not the best choice for the SCA. Many of the other organizations are using HTML to manage information about a small group of repositories. The SCA, however, will have over 1,000 repositories in its directory, and

using HTML is simply not practical nor is it sophisticated enough to perform the tasks desired by the association.

There is another markup language that can be used more efficiently than HTML alone. This other markup language is XML (Extensible Markup Language), a meta-language that was accepted by the WC3 in 1998 (Flynn, n.d.). It is considered a cross-platform software and hardware-independent tool for transmitting information. XML and HTML have different goals. HTML is used for displaying data, while XML is used for containing and describing the data. It also provides access to information. Data in an XML document exists in a data file and is structured into pre-defined elements and attributes by the encoder. Because the author creates the structure of the data, the elements can be tailored for a particular audience and written in a terminology relevant to them (STEP Stürtz Electronic Publishing GmbH, n.d.). To illustrate this, Figure 20 gives examples of HTML and XML codes that appear identical to the browser but, in fact, do different functions.

<p>HTML:</p> <pre><p>P266 Laptop
Friendly Computer Shop
\$1438</pre>	<p>XML:</p> <pre><product> <model>P266 Laptop</model> <dealer>Friendly Computer Shop</dealer> <price>\$1438</price> </product></pre>
<p>The HTML data is structured with tags into a paragraph with two line breaks. The browser has no way to interpret the meaning of this data.</p>	<p><product>, <model>, <dealer>, and <price> are elements defined by the author. With XML, the browser can associate the data with the elements.</p>

Figure 20. Example of HTML and XML data and encoding (Goldfarb, 1999).

In Figure 20, XML is shown to be very powerful, yet XML is not designed to replace HTML. Through the help of XSLT (Extensible Stylesheet Language Transformations), XML and HTML can be brought together. XSLT is an XML transducer, which produces an external presentation of the data (Salminen, 2003). XSLT will allow for the data in XML to be manipulated while still allowing HTML to format the data. In this way XML and HTML can be used together to compliment one another (W3 Schools, XML, n.d.). By using XML, HTML, and XSLT together, the SCA's online repository directory will be more flexible than many of the other online repository directories. XSLT will allow the elements and attributes of the XML file to be selected or excluded and sorted.

To demonstrate how XML and XSLT work together, the following is a simplified version of what a repository data file might look like. This is a fictional database and not the actual SCA repository data file.

```
<institutions>
  <record>1
    <parent>University of California, Irvine</parent>
    <repository>James Irvine Research Library</repository>
    <type>Educational</type>
    <county>Orange</country>
    <city>Irvine</city>
  </record>

  <record>2
    <parent>Orange County Citrus Growers</parent>
    <repository>Sunkist Archives</repository>
    <type>Corporate</type>
    <county>Orange</country>
    <city>Orange</city>
  </record>

  <record>3
```

```

    <parent>Brooks College</parent>
    <repository>Brooks Fashion Center</repository>
    <type>Educational</type>
    <county>Los Angeles</county>
    <city>Long Beach</city>
  </record>

  <record>4
    <parent>Crystal Cathedral</parent>
    <repository>Robert Schuler Reading Room</repository>
    <type>Religious</type>
    <county>Orange</county>
    <city>Garden Grove</city>
  </record>

  <record>5
    <parent>The County of Orange</parent>
    <repository>Orange Historical Society</repository>
    <type>Government</type>
    <county>Orange</county>
    <city>Orange</city>
  </record>
</institutions>

```

Figure 21. Example of an XML data file.

This data file is arranged hierarchically into three levels. This first level is the data file is called `<institutions>`; the entire data for the data file is contained between the `<institutions>` and `</institutions>` elements. Beneath the `<institutions>` element is a sub element called `<record>`. Each record has its own unique number as well as the sub elements `<parent>`, `<repository>`, `<type>`, `<county>`, and `<city>`.

With the data now contained within the elements, XSLT can perform many functions. For example, here are some of the things XSLT could do with our sample data file:

XSLT Task	XSLT Result
<p>1. Find all of the records in Orange County, sorted alphabetically on repository name, and display the repository name.</p>	<p>Records 1, 2, 4, and 5 would be found and would be displayed in this order:</p> <p>James Irvine Research Library Orange Historical Society Robert Schuler Reading Room Sunkist Archives</p>
<p>2. Find all of the records in Orange County, with a primary alphabetical sort on city name, then a secondary alphabetical sort on repository.</p> <p>Display the city name and the repository name.</p>	<p>Records 1, 2, 4, and 5 would be found and would be displayed in this order:</p> <p>Garden Grove Robert Schuler Reading Room Irvine James Irvine Research Library Orange Orange Historical Society Sunkist Archives</p>
<p>3. Find all of the educational repositories in the entire data file, with a primary alphabetical sort on county.</p> <p>Display the county name, record number, parent institution, and the repository.</p>	<p>Records 1 and 3 would be found and displayed in this order:</p> <p>Los Angeles Record: 3 Brooks College Brooks Fashion Center</p> <p>Orange Record: 1 University of California, Irvine James Irvine Research Library</p>
<p>4. Find all of the repositories in the data file with a primary alphabetical sort on type of institution.</p> <p>Display the type, record number, and repository.</p> <p>(Note that since a secondary alphabetical sort on repository was not performed, the James Irvine</p>	<p>Records 1 – 5 would be found and displayed in this order:</p> <p>Corporate Record: 2 Sunkist Archives</p> <p>Educational Record: 1 James Irvine Research Library</p>

Research Library appears before the Brooks Fashion Center.)	Record: 3 Brooks Fashion Center Government Record: 5 Orange Historical Society Religious Record: 4 Robert Schuler Reading Room
--	---

Figure 22. *XSLT tasks and results.*

The above examples in Figure 22 show how XSLT sorts data. In addition, XSLT can also display the results of a search by using HTML. This is done by applying HTML tags to the XML elements requiring only a few lines of code for each element. By using HTML in this way, far less encoding is actually needed than with a Web page encoded entirely in HTML. Below is a list of the XML elements in our data file. Next to each element is a description of the attributes to display the data. These attributes are applied to the font.

<type> Bold, Tahoma, Color Black, Size 12
<record> Normal, Tahoma, Color Black, Size 12
<repository> Italics, Arial, Color Black, Size 14

If the XSLT stylesheet used HTML tags to control the appearance of the XML elements, the result would look like this in the browser:

Figure 23. Browser display of an XML file.

The above example only requires a few lines of encoding in XSLT for the three elements:

<type>, <record>, and <repository>. This is what the XSLT encoding would look like:

```
<xsl:template match="type">
  <p style="color:black; font-family:Tahoma; font-size 12pt;fontstyle:bold">
    <xsl:value-of select="."/>
  </p>
</xsl:template>

      <xsl:template match="record">

        <p style="color:red; font-family:Tahoma; font-size 12pt">
          <xsl:value-of select="."/>
        </p>
        <xsl:template match="repository">
          <p style="color:black; font-family:Arial; font-size 14 pt">
            <xsl:value-of select="."/>
          </p>
        </xsl:template>
      </xsl:template>
```

Figure 24. Example of XSLT encoding.

However, if the same table was done solely in HTML, the following encoding would be needed for both the data and the formatting:

```
<body>
<p class="MsoNormal" style="text-align:justify"><b>
<span style="font-size:14.0pt">Corporate</span></b></p>
<p class="MsoNormal" style="text-align:justify"></p>
<p class="MsoNormal" style="text-align:justify"><span style="color:red">
Record: 2</span></p>
<p class="MsoNormal" style="text-align:justify"><i>
<span style="font-family:Arial">Sunkist Archives</span></i></p>
<p class="MsoNormal" style="text-align:justify"></p>
<p class="MsoNormal" style="text-align:justify"><b>
<span style="font-size:14.0pt">Educational</span></b></p>
<p class="MsoNormal" style="text-align:justify"></p>
<p class="MsoNormal" style="text-align:justify">
<span style="color:red">
Record: 1</span></p>
<p class="MsoNormal" style="text-align:justify"><i>
<span style="font-family:Arial">James Irvine Research Library</span></i></p>
<p class="MsoNormal" style="text-align:justify"></p>
<p class="MsoNormal" style="text-align:justify">
<span style="color:red">
Record: 3</span></p>
<p class="MsoNormal" style="text-align:justify"><i>
<span style="font-family:Arial">Brooks Fashion Center</span></i></p>
<p class="MsoNormal" style="text-align:justify"></p>
<p class="MsoNormal" style="text-align:justify"><b>
<span style="font-size:14.0pt">Government</span></b></p>
<p class="MsoNormal" style="text-align:justify"></p>
<p class="MsoNormal" style="text-align:justify">
<span style="color:red">Record:&nbsp;5</span></p>
<p class="MsoNormal" style="text-align:justify"><i>
<span style="font-family:Arial">Orange Historical Society </span></i></p>
<p class="MsoNormal" style="text-align:justify"></p>
<p class="MsoNormal" style="text-align:justify"><b>
<span style="font-size:14.0pt">Religious</span></b></p>
<p class="MsoNormal" style="text-align:justify"><span
style="color:red">&nbsp;</span></p>
<p class="MsoNormal" style="text-align:justify"><span style="color:red">
Record: 4</span></p>
</body>
```

Figure 25. Example of HTML encoding.

This illustrates the efficiency of using XSLT and HTML together. The HTML example has no automated sorting capability. The entire layout of the HTML document must first be decided by the encoder who would then type the data word for word on the Web page. Referring back to the Society of Indiana's Website, 85 separate Web pages were designed to handle one search: primary sort on county, secondary sort on city, and tertiary sort on repository name. This would have been much more efficient if done in XSLT. The XSLT encoding is easier to read and would have taken far less time to do. Furthermore, search options could have been programmed into the Website for the researcher.

Another function that HTML cannot do is calculation. In contrast, XSLT will handle simple mathematical calculations. For example, the number of academic archives in San Diego County could be counted, thereby allowing statistics to be generated from the XML file. One of the most exciting advantages of using XML and XSLT together is that one source file (XML) can be used to generate several stylesheets, each displaying different aspects of the information. In this way one set of stylesheets could be used by information seekers, another by archival institutions, and, perhaps, another by the SCA board members.

There are additional advantages regarding the creation, revision, and cost of maintaining an XML file. XML files can be easily created using a PC database program such as MS Access or Filemaker. After the file has been created, it can simply be exported into an XML editor program and then checked for validation and/or form. Maintaining a file in XML is much easier than updating a file in HTML. This is because

a new repository can simply be added at the end of the XML file. The order of the repositories is not important in an XML file because XSLT performs all of the sorting based on the criteria given. This is not possible in HTML. A final advantage of using XML, XSLT, and HTML together is that they are non-proprietary products, thereby costing the SCA nothing for their use (STEP Stürtz Electronic Publishing GmbH, n.d.). This is important since proprietary database products such as Oracle and SQL Server, discussed below, can be quite costly.

DBMS

A Database Management System (DBMS) is the second possibility for the online repository directory. A DBMS is a software system that facilitates the creation and maintenance of an electronic database. There are several database management products on the market including Oracle, SQL, MySQL, and Microsoft Access. Most DBMSs are based upon relational databases. A relational database is a way to manage data inside a set of two-dimensional tables with relationships between the tables. Each table consists of rows and columns. The relationships between the tables are defined by one table having a column with the same meaning as a column in another table (SmartSoft Ltd., n.d).

Using the online repository directory as an example, there might be two tables: repositories and parent institutions. The repositories table might have these columns: repository name, address, city, and parent name. The parent institutions table might have these columns: parent name, address, and city. The online repository directory would, in fact, have additional tables that would contain information about a repository's type,

various holdings, and formats. One of the advantages of storing information in tables is that it eliminates redundant data entry. For example, if the University of California, Irvine - a parent institution - acquired a second repository, only information about the new repository would have to be entered into a table. Since data about the University of California, Irvine, already exists in the Parent Institutions Table, it would not have to be entered again.

Another advantage of storing data in a relational database is that it can be retrieved and viewed several different ways. Data is viewed by performing queries on the tables. A query selects data from a table or tables that matches a specific condition. For example, information about a repository and its associated parent institution, ordered by the repository name, could be viewed with the following query:

```
SELECT Repositories.Repository_Name, Repositories.Address, Repositories.City,
       Parent_Institution.Parent_Name, Parent_Institution.Address,
       Parent_Institution.City

FROM Repositories, Parent_Institution

WHERE Repositories.Parent_Name = Parent_Institution.Parent_Name

ORDER BY Repositories.Repository_Name
```

Figure 26. Example of a query.

Although a DBMS is very powerful and would theoretically work for the online repository directory, it has many disadvantages and is not suited for the SCA. One of the main reasons for not selecting a DBMS for the SCA directory is that many DBMSs can be quite costly. Oracle's Standard Edition One is \$4,995 with a \$745 annual licensing fee (Oracle Corporation, 2004). Furthermore, DBMSs are usually maintained by a database administrator (DBA). The role of a DBA is to make sure the database is

running, has enough space, and is performing at peak efficiency. The DBA also backs up the database and troubleshoots any problems. In California a full-time DBA's salary can range from \$50,000-\$90,000 a year, while a contractor DBA can earn between \$30 and \$50 per hour. This is far more than the SCA could afford, as it is a voluntary association for archivists not a business.

In addition to the cost, a high level of expertise is also needed for the creation and maintenance of the database. Creating the database with the appropriate rights and privileges for a Web application requires a high degree of knowledge. Once the database is created, further programming expertise is required to transfer the data from the flat files into the appropriate database tables. After the database was on the World Wide Web, it would be maintained by a volunteer member of the SCA. Based on the fact that a DBMS would be too expensive for the SCA, and that high-level expertise is required for both its development and maintenance, a DBMS currently does not meet the needs for the SCA's online repository. It is more demanding than appropriate for a voluntary association. However, an XML data file does meet the requirements for the project in terms of cost, ease of maintenance, and functionality. Therefore, it is recommended that XML be selected for the online repository directory.

Transitioning from a Print to an Online Directory

After the SCA Publications Committee recommended that the next version of the SCA repository directory be made into a Web-based document, the publications committee started doing some preliminary work to prepare for the transition. Since the directory's 4th edition was published seven years ago, the SCA Publications Committee

knew that much of the information needed updating. They created a survey form asking for updated information that was mailed to the repositories listed in the 1996 edition. The repositories had the option of filling out the survey and mailing it back to the SCA Publications Committee or submitting their information in electronic format. The SCA Publications Committee then input the data from the returned surveys into a Filemaker PC database. The Publications Committee is currently editing the data in Filemaker in order to give it an overall standardized appearance.

The goal of this project is to create a data file-driven online repository directory based upon the information obtained from these surveys that will be incorporated into the SCA's current Website. The data file will be created by exporting the data, currently in Filemaker, into an XML file. The online version will have more fields for displaying information than the print directory making it more useful for the researcher. The following fields will be found in the online version:

- Repository Name
- Parent/Sponsoring Institution
- Contact Information
- Location (address, city, zip code, county, telephone, fax, e-mail, URL)
- Access information
- Type of repository
- Description of holdings
- Descriptors (subject, geographical, individuals, families, organizations)
- Extent of holdings

- Dates (bulk dates, oldest record dates)
- Formats
- Finding Aids
- Union Lists
- Controlled Subject Headings

There will be two different types of searches made available. One type will be advertised on the SCA Website for general public use. The other type will provide searches of importance to the SCA officers, but not to the general public. The SCA administrative searches would either not be advertised on the Website or would be accessible by password.

After working with SCA officers to determine search criteria, a wireframe, or prototype of the Web pages' structure, will be created. This will be done in HTML mainly to help plan the navigation of the Web pages. It will also help determine layout and basic content for each Web page (Stanford, 2003). Once the wireframe structure is completed, it will be shown to the SCA liaison for comments before further design steps are taken. After a basic structure has been agreed upon, XML stylesheets for the online repository directory will be created.

When there is a significant subset of the data ready in XML format (approximately 100 records), a usability test will be conducted to determine ease of use in navigating the Website. The test will be informal in nature, using approximately ten participants. The participants will come from the target audience for the Website, namely archivists, librarians, MLIS students, and information professionals. The test will ask participants to

perform specific tasks and rate the ease of use and navigation. Because the test is informal, the participants will do the testing from the convenience of their own computers and will e-mail the results back to me. The recommendations from the usability test will be incorporated into Website and will then be submitted to the SCA Publications committee before the Website is released to the public (gotomedia, n.d.).

CHAPTER 4

PROJECT IMPLEMENTATION

After reviewing other archival organizations' Websites and selecting a methodology best suited for the SCA's online repository directory, the process of creating an online repository directory finally began. There were many steps involved in transforming the outdated print directory into a current searchable Website. The main steps involved in accomplishing this task were data processing, stylesheet design, Website design, and usability testing. The end result of all these tasks was a test version of the online repository directory, available on the SCA Website at <http://www.calarchivists.org/erwine/demo/demomain.html>.

Data Processing

Data processing, the first step in creating the online repository directory, involved gathering data, transferring it into a Filemaker database, and then exporting it from Filemaker into XML. This process is depicted in a flowchart appearing in Figure 27. The effort of gathering updated information for the electronic version of the repository directory began in the fall of 2002 when the SCA Publications Committee developed a survey to be mailed out to those repositories listed in the directory's last edition. The survey entitled the "Directory of Archival and Manuscript Repositories in California Data Sheet," can be found in Appendix A. In the summer of 2003 surveys were mailed to over 1,000 repositories throughout California. The repositories had the option of filling out the surveys and returning them by mail or to fill out the surveys

Figure 27. Data processing flow chart.

electronically via a Website on the School of Information Science's server.

Approximately 100 repositories submitted the surveys electronically while approximately 300 repositories returned their surveys by mail (Hansen, 2003).

The electronically submitted surveys were converted into an Excel spreadsheet using a Java program. The data for each survey was linearly contained in one row with 161 columns that represented the different categories of information requested on the survey. At this point a data structure was created in Filemaker containing field names that matched the columns in the spreadsheet, along with the appropriate values to match the data fields. The Filemaker database was then slightly modified with the addition of a comments field for document conversion and data entry information and a record number field so that each record could be assigned a unique identification number. The data from Excel was then exported into Filemaker. The information from the surveys returned by mail was manually keyed into an Excel spreadsheet and a similar conversion to the Filemaker database was made. A sample Filemaker record for the Levi Strauss & Co. Archives appears in Appendix B.

With the electronically submitted records and the manually keyed records residing in the same database, it was possible to convert the Filemaker records into XML.

Although many of the Filemaker field names remained the same in XML, it was decided to streamline the new XML structure to reduce the excessive number of field names, thereby making the records more manageable for maintenance and data entry. Each record in the new XML structure contains 43 elements, a great improvement over the 163

field names used in the Filemaker structure. The streamlining mainly occurred in the areas of the repository types, formats, and holdings.

In the case of the repository types, there are 33 possible repository types, from which 2 can be selected to describe a repository. In the Filemaker database, each of the 33 repository types has its own field. The Filemaker record in Appendix B shows that the Levi Strauss & Co. Archives is a business repository, indicated by the word “checkbox” appearing next to the business field. The 32 other repository types listed in the data structure not describing the Levi Strauss & Co. Archives are identified with the word “null”. Admittedly this is not the most efficient database design, especially for data entry purposes. However, the reason for the numerous fields in the Filemaker data structure is that it was generated from the electronically submitted records. The electronic format was designed to make data submission easier for the repositories, not for database efficiency. The XML data structure is less complicated than the Filemaker structure. Repository types are contained within one element called “<Reptypes>”. Within <Reptypes> repository types may be identified with the optional attributes of “Type1” and “Type2.” Appendix C shows how the Levi Strauss & Co. Archives record appears in the XML data file. It only takes the following line of code to identify it as being a business repository:

```
<Reptypes Type1="Business"></Reptypes>
```

Attributes are commonly used in XML and HTML to more precisely describe the elements to which they are assigned. For example, the element in the stylesheet example in Appendix E has the attributes of color, face, and size. Each of these attributes

describes different characteristics of the element. In many cases attributes can be replaced with child elements. The above Levi Strauss record could also have been written in the following way:

```
<Reptypes>
  <Type1>Business</Type1>
  <Type2/>
</Reptypes>
```

The main difference is that the data appears on four lines instead of one when encoded using elements instead of attributes. Because attributes and elements can perform the same function, it is ultimately the decision of the encoder to decide the definition of the structure.

There was one other instance in which the electronically submitted surveys generated an excessive number of fields in the Filemaker database. This was in the case of identifying a repository's formats. Appendix B shows that there were 35 fields used for containing information about a repository's formats. Again, this was streamlined in the XML structure. A repository's formats are now described in one element called "<Formats>". However, the format types were not identified with attributes as was done with the repository types. Instead, they were included together as one string of data separated by commas as shown below:

```
<Formats>Architectural, Paintings, Photographs, Posters, Artifacts, Books,
Ephemera, Films, Institutional, Legal, Newspapers, Oral Histories, Pamphlets,
Periodicals, Sound Recordings, Subject Files</Formats>
```

This method was selected because it is easier not only for data entry purposes but also to read. This differs from the attributes method used earlier when describing repository types. Attributes were chosen to describe repository types because the repositories are limited to describing themselves with no more than two repository types. In the case of formats, however, it is plausible that a repository could have as many as 35 different formats. Therefore, listing each repository's formats as one string of data was the more practical method.

Finally, there was one more area in the Filemaker data structure that was streamlined in XML. This occurred in describing a repository's holdings. The Filemaker database included five subcategories in which a repository's holdings could be described: subjects, geographical areas, individuals, families, and organizations. These are shown in Appendix B. Each of these subcategories used 10 fields for containing data. For example, the 10 fields used to describe subjects are: "Subjects," "_Subjects," "F3," "F4," "F5," "F6," "F7," "F8," "F9," and "F10". The purpose of this structure was to limit a repository to using no more than 10 descriptors when describing the subjects of its holdings. The subjects subcategory was streamlined in XML in the same way the formats field was streamlined; one string of data contains all of a repository's subjects:

<Subjects>Costume and textiles, Advertising/Marketing, Dry goods, Popular culture, Gold Rush, Manufacturing, Dude ranches </Subjects>

The other subcategories were streamlined in the same way and can be seen in Appendix C.

To ensure that the XML data file does not become corrupted, its structure is defined by a DTD (Document Type Definition). A DTD consists of a list of elements and attributes used in the data structure. During data entry, if an element or attribute not listed in the DTD was typed, an error message would appear indicating that the record was invalid and not well formed. In addition to defining elements and attributes, a DTD also sets rules that control the data. For instance, in the online directory data file, there are two rules associated with the RECORDID attribute of the <ROW> element.

`<!ATTLIST ROW RECORDID CDATA #REQUIRED>`

According to the encoding, a record number must be entered for each RECORDID and cannot be left blank. This is purpose of term “#REQUIRED”. Furthermore, the data entered for the RECORDID must be alphanumeric. If non-alphanumeric characters were mistakenly entered as data for the RECORDID, an error message indicating invalid data would appear.

The DTD protects the integrity of the XML data file by ensuring that all data within it is well formed and valid. Although it is not necessary to have a DTD for an XML file, it is advantageous because it prevents encoding errors from occurring during data entry. A DTD allows several people to enter data into the system in a standardized format. The DTD is created as an external document that is accessed from inside the repository directory source file. The DTD used for the online repository directory, along with a brief commentary, appears in Appendix D.

Stylesheet Encoding

The encoding of the XSL stylesheets occurred in two phases. The first phase was the construction of a stylesheet based solely on the code generated from the Filemaker database. After designing the first stylesheet, the XML structure was streamlined and required the creation of a second stylesheet. Since the first stylesheet is not the one being used in the SCA Website, its encoding will not be displayed in its entirety in this report. However, there were some important lessons learned while designing the first stylesheet that led to an improved second stylesheet. Therefore, significant parts of the first stylesheet's encoding will be discussed.

One of the goals of the first stylesheet, "repositories.xslt," was to see how much effort was needed to create it based on the XML code imported from Filemaker. The initial encoding of the first stylesheet was challenging because each XML record contained over 160 lines of encoding, and it took several lines of encoding in the XSL stylesheet to reflect the large number of XML elements. The first XSL stylesheet, in fact, had 1,062 lines of code.

One area of the XSL stylesheet that required several lines of encoding was the part used for selecting and displaying repository types. This is because in the original XML structure each of the 33 repository types existed as independent elements. If a repository's type was identified as "museum," it was indicated by the word "checkbox" next to the <museum> element. In order to select and display the repository types in XSL, each element had to be tested to see if it contained the word "checkbox". This was done by using the <xsl:choose> statement. If an element contained the word "checkbox,"

the repository type was displayed. If an element did not contain the word “checkbox,” it was ignored. Figure 28 below shows an excerpt of this encoding.

```
<xsl:choose>
  <xsl:when test="Military[.='checkbox']">
    <span class="mid">Military / </span>
  </xsl:when>
</xsl:choose>
<xsl:choose>
  <xsl:when test="Museum[.='checkbox']">
    <span class="mid">Museum / </span>
  </xsl:when>
</xsl:choose>
<xsl:choose>
  <xsl:when test="Newspaper[.='checkbox']">
    <span class="mid">Newspaper / </span>
  </xsl:when>
</xsl:choose>
```

Figure 28. XSLT encoding excerpt from repositories.xslt.

The `<xsl:choose>` statement was also used in another part of the stylesheet for selecting and displaying the repositories’ formats. The results were generated the same way. If an XML element identified a format with the term “checkbox,” it was displayed on the stylesheet.

Although Figure 28 only shows a small part of the encoding used to select and display the repository type, it demonstrates that five lines of encoding are required to display each repository type. This occurs in each `<xsl:choose>` statement. It should be mentioned that it is possible to write each `<xsl:choose>` statement in one line as follows:

```
<xsl:choose><xsl:when test="Film[.='checkbox']"><span class="mid">Film
/ </span></xsl:when> </xsl:choose>
```

However, XSL encoding is typically written as it appears in the above figure because it is easier to read. No matter which style is used for typing the `<xsl:choose>` statement, several lines of encoding are still required to display the repository types.

The large number of elements required to describe repository types and formats demanded the XML structure and subsequent stylesheet be streamlined. A new stylesheet, entitled “results.xslt,” was created to serve as the prototype for all the stylesheets used in the SCA’s online repository directory Website. The encoding for the second stylesheet is shown in Appendix E. The areas of the encoding for selecting and displaying the repository types and formats have been highlighted to show the efficiency of the second stylesheet as compared to its predecessor. The second style sheet is greatly improved over the first stylesheet because it uses fewer lines of code. Consisting now of only 176 rather than 1,000 lines of code, the XML database is easier to maintain and read.

Stylesheet Design

While one of the goals of the first stylesheet was to gauge the difficulty of working with the imported Filemaker data, another object was to view all of the data in the browser. This was needed to get a rough idea, visually, of how much data each record held. After this was known, a second spreadsheet could be developed. Therefore, the first stylesheet simply displayed the data from the XML structure. This was quite beneficial because it was easier to visualize different layouts for future stylesheets. It was also helpful in an unexpected way; it made it easier to see problems with the data, such as typographical errors and inconsistencies in entry formats from record to record.

Although it would have been possible to discover these problems by examining the data within the XML structure, it was easier to view the data in the browser using the first stylesheet. Figure 29 shows what the record for the Historical Society of Long Beach looks like using the first stylesheet.

Historical Society of Long Beach	
Parent Institution:	
Repository Type(s): Historical Society / Museum /	
Contact Name: Julie Bartolotto - Executive Director	
Location	Contact
210 E. Ocean Boulevard - Arcade Level	Phone: (562) 495-1210
Long Beach, California 90802	FAX: (562) 495-1281
County: Los Angeles	Email: hsfb@historicalsocietylb.org
Mailing Address	URL: http://www.historicalocietylb.org
	Contact Options: Telephone, Mail, Email
Access	
Open to Public: Yes	
Hours: Tuesday, Wednesday, Friday 1pm - 5pm Thursday 1pm - 7pm	
Special Requirements:	
Services offered: Photocopying, Photograph reproduction, Scanning, Fee-based research services	
Holdings Narrative: Holdings include over 30,000 photographic images of Long Beach and its vicinity, Long Beach Cultural Heritage Surveys documenting the architecture in historic neighborhoods throughout Long Beach, pre-1940 maps and atlases, and a variety of ephemera and memorabilia.	
Subject terms (f3-f10): Art and Architecture Maritime and naval history Business, Education, Genealogy, Government, Local history, Entertainment, Transportation, Oil industry	

Geo terms (f11-f18): Long Beach Signal Hill	
Individuals terms (f19-f26):	
Families terms (f27-f34): Bixby	
Organizations: United States Navy	
_Orgs:	
Organizations terms (f35-f42):	
Linear feet:	
Bulk Dates: 1890-1980	Oldest Date: 1880
Format: Architectural / Artifacts / Books / Digital / Electronic / Ephemera / Films / Institutional / Legal / Maps / Music / Sheet Music / Newspapers / Oral Histories / Pamphlets / Periodicals / Personal Papers / _Photographs / Inventories / Databases / Card Catalog /	
OPAC:	
Citation:	
Catalogs?:	

Figure 29. First stylesheet (repositories.xslt) in browser view.

After viewing the records using the first stylesheet, it was obvious that the initial design was not practical. The above record shows the many of its inherent problems. One noticeable problem is that the data overflows onto two screens. Also, the data's organization is illogical. For instance, the contact name "Julie Bartolotto" and her title appear on the left-hand side of the screen above the repository's location, while other contact information, such as phone numbers, e-mail, and URL, appear in a separate column on the right-hand side of the screen. The page also lacks a sense of style because the font sizes, styles, and colors are inconsistent. There is also a problem with the bottom part of the stylesheet that indicates there was confusion as to displaying catalog information. Lastly, the record number is missing.

Taking all of the flaws of the first stylesheet into consideration, it was possible to construct an improved second stylesheet. The problems of data overflow, poor layout, and inconsistent style, among others, were corrected in the second stylesheet, “results.xslt.” Figure 30 displays the new stylesheet’s layout and format.

<p>Record: 0052</p> <p>Historical Society of Long Beach</p> <p>Parent Institution:</p> <p>Type - Historical Society, Museum</p> <p>Address: 219 E. Ocean Boulevard, Arcade Level Long Beach, CA 90802</p> <p>County: Los Angeles</p> <p>Contact: Julie Bartolotto - Executive Director</p> <p>Phone: (562) 495-1210 Fax: (562) 495-1281</p> <p>Email: hsib@historicalsocietylb.org</p> <p>URL: http://www.historicalocietylb.org</p>	<p>Holdings Narrative: Holdings include over 30,000 photographic images of Long Beach and its vicinity, Long Beach Cultural Heritage Survey documenting the architecture in historic neighborhoods throughout Long Beach, pre-1940 maps and atlases, and a variety of ephemera and memorabilia.</p> <p>Subjects: Art and Architecture, Maritime and naval history, Business Education, Genealogy, Government, Local history, Entertainment, Transportation, Oil industry</p> <p>Geographical Areas: Long Beach, Signal Hill</p> <p>Individuals:</p> <p>Families: Bixby</p> <p>Organizations: United States Navy</p> <p>Linear Feet: Bulk Dates: 1890-1980 Oldest: 1880</p> <p>Formats: Architectural, Artifacts, Books, Digital, Electronic, Ephemera, Film, Institutional, Legal, Maps, Music, Sheet Music, Newspapers, Oral History, Pamphlets, Periodicals, Personal Papers, Photographs</p> <p>Published Guides:</p> <p>Finding Aids: Inventories, Databases, Card Catalog</p> <p>Union Lists or Networks:</p>
--	---

Figure 30. Second stylesheet (results.xslt) in browser view.

The improvements in the second stylesheet are immediately noticeable. The data now fits on one screen and has a more logical layout. This was done by dividing the screen into two columns. Information about the repository itself is in the left column while information about the repository’s collections, finding aids, guides, and union lists are listed in the right column. The left column was divided into three vertical sections to make the information more presentable. The gray top section contains information about the repository’s identity and physical location. If a repository has a separate mailing address, it also appears in this section. If the mailing address is the same as the street

address, only the street address is displayed. The yellow middle section has information on how to contact the repository. The last section, in blue, shows if a repository is public or private, its hours, special user requirements, and services provided.

Another improvement in the second stylesheet is that it is more visually pleasing. This was done by selecting a color palette and uniform fonts. Because this stylesheet is going to be incorporated into the existing SCA Website (www.calarchivists.org), it was designed using their color palette and default font. The font sizes were varied as needed. A larger font size was used to make certain information more prominent, such as repository name and type. Smaller sized fonts were used to get as much data as possible onto one screen and also to keep text wrapping to a minimum. As a result of the changes made, the second stylesheet not only has a greatly improved layout, it also has the look of an SCA Web page.

Website Design

Once the first stylesheet was complete, the project entered the design phase. The layout for the first Website began as a wireframe structure, hand-drawn on paper, that included notes on navigation ideas as well as possible searching strategies. The ideas on paper were then refined and an HTML wireframe website created. This preliminary website allowed users to search on a variety of fields using a JavaScript search interface that displayed the results with the first stylesheet, `repositories.xslt`. Users were asked for feedback after using the wireframe website, and their comments were incorporated into a final Website presented to the SCA Board of Directors. A discussion of the users' feedback appears in Chapter 5.

The wireframe Website includes the following five web pages:

- wireframemain – the Web site's home page
- wireframename – searches on a repository or parent institution name
- wireframekeyword – performs broad or narrow keyword searches
- wireframelocation – allows users to search by city or county
- wireframetype – searches on repository type

The above Web pages in the wireframe website were designed to look like an SCA Web page by using the SCA color pallet as well as the SCA logo. Further standardization was achieved by using a uniform navigation bar across the top of all Web pages. The navigation bar allows users to go to the other pages on the wireframe Website. Individual screen images of the wireframe Website are shown below in Figures 31-35. Below each figure is a brief commentary on its functionality.

Find by Name Find by Keyword Find by Location Find by Repository Type Go To Directory Main Page

The Society of
California Archivists

Testing Network
Find an Archival Repository
Main Page

Find an Archival Repository

Search by:

Name: Repository Name or Parent Name

Type: Displays a List of Repository Types

Location: County or City

Keywords: Performs broad or narrow keyword searches

Figure 31. Main page (wireframemain).

The body of the main page contains hyperlinks to the web pages listed in the navigation bar. Each hyperlink provides a brief explanation about the type of searches that can be performed on the other web pages. For example, by looking at the main page, a user would know that he or she would be able to search on a repository's name or parent name when doing a name search. The main page is the only Web page on the wireframe site that does not use a JavaScript search interface.

Find by Name	Find by Keywords	Find by Location	Find by Repository Type	Go To Directory Main Page
------------------------------	----------------------------------	----------------------------------	---	---

The Society of California Archivists

Testing Version

Online Directory

Name Page

Find a repository by name or by parent institution name

Enter your search terms

Then click on a button below

Figure 32. Find by name page (wireframename).

The find by name page, wireframename, allows users to perform an interactive search on repository or parent institution name. The user enters the search terms in the dialog box and then selects “Search By Repository Name” or “Search By Parent Institution Name.” The user may select the reset button to refresh the page to perform another search.

Find by Name	Find by Keywords	Find by Location	Find by Repository Type	Go To Directory Main Page
------------------------------	----------------------------------	----------------------------------	---	---

The Society of California Archivists

Testing Version
 Online Directory - Keyword Page

Search on keyword(s)

For a broader search click below

For a narrower search click on a button below

Search Holdings	Search Subject	Geographic Terms	Individuals	Organizations	Families
---------------------------------	--------------------------------	----------------------------------	-----------------------------	-------------------------------	--------------------------

Figure 33. Find by keywords page (wireframekeyword).

The keyword search page allows the user to perform two different types of searches. A user can either perform a broad keyword search, which searches all fields in the data file to find the term. The user also can perform a narrower search on holdings (holdings narrative) to locate specific subjects, geographic terms, individuals, organizations, or families.

Find by Name	Find by Keywords	Find by Location	Find by Repository Type	Go To Directory Main Page
------------------------------	----------------------------------	----------------------------------	---	---

The Society of California Archivists

Testing Version

Online Directory
Location Page

Find a repository by county or city

Enter your search terms

Then click on a button below

Search By City Name	Search By County Name
Reset	

Figure 34. Find by location page (wireframelocation).

From the location page the user enters the name of a city or county in the dialog box and selects the appropriate button to perform the search.

Find by Name	Find by Keywords	Find by Location	Find by Repository Type	Go To Directory Main Page
------------------------------	----------------------------------	----------------------------------	---	---

The Society of California Archivists

Testing Version
 Online Directory
 Type Page

Find a repository by type

Enter one of the following repository types listed below:

<ul style="list-style-type: none"> • Academic • Business • Dance • Government: (State, Federal) 	<ul style="list-style-type: none"> • Entertainment: (Film, Dance, Opera, Orchestra, Radio, Theatre, TV) 	<ul style="list-style-type: none"> • Historical Site • Historical Society • Law • Medical • Military 	<ul style="list-style-type: none"> • Museum • Newspaper • Professional • Private • Public 	<ul style="list-style-type: none"> • Religious • Social • Transportation • Other
---	--	---	--	--

Figure 35. Find by repository type page (wireframetype).

This page allows for the search of a repository by type of institution or library (i.e., academic, government, public, etc.). A list of repository types appears on the screen. The user performs the search by entering the repository type into the dialog box.

After the wireframe Website was tested, the creation of a second Website began. The second Website, called the “demo Website,” was intended to demonstrate the online repository directory to the SCA Board of Directors. The demo Website has many of the same features as the wireframe Website, such as use of the SCA logo and color palette, but it also has some differences. The main difference between the two Websites is that the demo Website has an additional Web page for searching by format. The Web pages in the demo Website are also shorter in length when viewed in the browser. These enhancements were made based on the feedback from the original wireframe Website’s testers.

The demo Website, depicted in Figures 36 – 42, can be accessed from <http://www.calarchivists.org/erwine/demo/demomain.html>. It includes the following seven Web pages:

Figure 36. Online Repository Directory Main Menu (demomain).

This is the main menu of the demo Website. The layout was based on an actual SCA Web page, which uses the far left-hand column for navigation to other SCA Web pages. The main Web page is entitled “Online Repository Directory.” It consists of a two-column table with hyperlinks to other Web pages in the Online Repository Directory along with a brief explanation of the types of searches possible from these pages.

Return to Main Search	Repository/Parent Name Search	Location Search	Repository Type Search	Format Search	Holdings Search
---------------------------------------	---	---------------------------------	--	-------------------------------	---------------------------------

Enter the name of a repository or parent institution

[Search By Repository Name](#)
[Search By Parent Institution Name](#)
[Reset](#)

Figure 37. Repository/Parent name search page (demoname).

Demoname allows users to perform an interactive search on a repository or parent institution name. The user enters the search terms in the dialog box and then selects “Search By Repository Name” or “Search By Parent Institution Name.”

Return to Main Search	Repository/Parent Name Search	Location Search	Repository Type Search	Format Search	Holdings Search
---------------------------------------	---	---------------------------------	--	-------------------------------	---------------------------------

Location Search

Enter the name of a city or county

[Search City](#)
[Search County](#)
[Reset](#)

Figure 38. Location search page (demolocation).

This page allows a user to search for a repository by city or county. The user enters the search terms in the dialog box and then selects “Search City” or “Search County.”

Return to Main Search	Repository/Parent Name Search	Location Search	Repository Type Search	Format Search	Holdings Search
---------------------------------------	---	---------------------------------	--	-------------------------------	---------------------------------

Search for a repository by type

Academic	Historical Site	Radio
Business	Historical Society	Religious
City	Museum	Research
County	Other	State
Entertainment	Private	Transportation
Government	Public	

View Search Results:

Sorted by Name	Sorted by City	Sorted by County	Reset
--------------------------------	--------------------------------	----------------------------------	-----------------------

Figure 39. Repository type search page (demotype).

This page allows for the search of a repository by its type. The list has bullets and parentheses removed and fewer items making it easier to read than the original version. The user enters a repository type (selecting from a list on the screen) into the dialog box and selects “Sorted by Name,” “Sorted by City,” or “Sorted by County” to view the search results.

Return to Main Search	Repository/Parent Name Search	Location Search	Repository Type Search	Format Search	Holdings Search
---------------------------------------	---	---------------------------------	--	-------------------------------	---------------------------------

Enter a format type. Example: Books, Maps, Newspapers

To view a complete list of format types click [here](#)

View Search Results:

Sorted by Name	Sorted by County	Sorted by City	Reset
--------------------------------	----------------------------------	--------------------------------	-----------------------

Figure 40. Format search page (demoformat).

This Web page allows a user to search for archival formats. The user can have the data retrieved and then sort the results alphabetically by repository name, county, or city.

Because the list of archival formats is long, a hyperlink to a sub Web page listing the formats was created to help the user.

Formats	
Architectural	Maps
Art	Microform
Artifacts	Music
Audiotapes	Newspapers
Books	Oral Histories
CDs	Paintings
DAT	Pamphlets
Digital	Papers
Drawings	Periodicals
DVD	Personal Papers
Electronic	Photographs
Ephemera	Posters
Films	Records
Genealogy	Sheet music
Institutional	Sound Recordings
Legal	Subject Files

[Return To Format Search](#)

Figure 41. Format list page (formatlist).

This listing of archival formats is accessible only from demoformat.html.

Navigation is restricted to demoformat.html via a hyperlink at the bottom of the screen that says “Return To Format Search.”

Return to Main Search	Repository/Parent Name Search	Location Search	Repository Type Search	Format Search	Holdings Search
Search Holdings Fields					
<input type="text"/>					
<input type="button" value="Holdings Narrative"/> <input type="button" value="Individuals"/> <input type="button" value="Subjects"/> <input type="button" value="Geographical Areas"/> <input type="button" value="Families"/> <input type="button" value="Organizations"/> <input type="button" value="Reset"/>					

Figure 42. Holdings search page (demoholdings).

At first glance this looks like a new Web page but, in fact, it is a redesign of its predecessor wireframekeyword.html. It allows a search of the fields associated with the

repositories' holdings. The user enters a search term into the dialog box and then selects one of the following fields to search: holdings narrative, individuals, subjects, geographical areas, families, or organizations. Since this screen is dedicated to searching holdings, it no longer has the ability to perform a broad search of every field in the data file. The capability to search every field will be restored before the Website is released to the public. This is discussed in Chapter 5.

Usability Testing

User testing for the online repository directory was conducted in two phases. The first phase involved informal testing of the wireframe Website along with the first stylesheet, `repositories.xslt`, by a small group of testers. The testers had many good ideas for improving the Website and stylesheet. After the testing was done, these ideas were incorporated into a second Website, the demo Website. The second phase of testing involved the evaluation of the demo Website. The demo Website, however, was evaluated in a different manner than the original wireframe Website. Instead of being evaluated by a small group of testers, it was presented to the SCA Board of Directors at their annual meeting. After the presentation, the Board members were asked for their feedback on the demo Website. Although the testing styles differed, both yielded many good ideas that will greatly benefit the online repository directory.

The testing of the wireframe Website and stylesheet involved three testers. Two of the testers were on the SCA Publications Committee and one was an information technology specialist. The testers were emailed the URL of the wireframe Website and asked to evaluate it at their convenience. There were two main objectives in the

wireframe Website test. The first objective was to test the functionality of the Website. In order to do this, the testers were asked to perform a few simple tasks. First, they were asked to navigate from screen to screen to check for any broken hyperlinks. Next, they were asked to perform searches from each of the interactive Web pages. This was done to check the integrity of the JavaScript search interface. The second objective of the testing was to get feedback on the overall design of both the Website and the stylesheet. For this part of the evaluation, the testers were simply asked for their opinion of the design features such as layout, colors, and fonts.

The feedback from the testers regarding both objectives of the wireframe Website test was very insightful. The testers reported that accomplishing the first task, Website navigation, was easy. They found no broken hyperlinks and indicated that navigation went smoothly. The next task of performing searches from each of the interactive Web pages went well, but the users did require a small amount of assistance. Because there were only ten records in the wireframe Website data file, testers were provided three repository names, three cities, and three counties so they could perform a few initial searches. They were not, however, given any terms describing the holdings or the repository types. It was hoped that after performing a few searches, they could get the holdings terms and types on their own by looking at the search results displayed in the stylesheet. Testers were asked to search on city location, county location, repository type, repository name, and by keywords. This task went fairly well, but did reveal a problem with the Website design. The stylesheet did not display on the screen after the testers entered their terms into the dialog box because the Web page was too big. The

testers needed to scroll down to the lower part of the screen to view the stylesheet. After they were told to do this, all of the testers were able to perform successful searches from all of the Web pages. The discovery of this problem led to an improved design on the subsequent demo Website. All interactive Web pages were designed have a smaller screen size. Thus, when searches are performed, the stylesheet appears on the current screen in its entirety so users can view their search results without having to scroll down. Another problem that was brought to light after the testers performed searches was the fact that there was no way to search for archival formats. Two users thought that this capability should be included on the next stylesheet.

The second objective, feedback on the overall design, was also quite helpful. One of the first things that testers commented on was the look of the Website. Because two of the users were on the SCA Publications Committee, they immediately noticed the SCA logo and color pallet and were very pleased with the look. They also noticed that the stylesheet did not use the SCA color pallet. They suggested that the next stylesheet be designed using the SCA's colors and fonts so that they matched each other. Another comment made about the stylesheet concerned the layout. All three testers said that while it was nice to see all of the fields displayed, the data took up a lot of screen space and did not have a logical or efficient layout. They felt that the layout could be greatly improved if the fields were rearranged into logical groupings and less screen space was used. The testers also noticed that there was an inconsistency of the font styles used in the stylesheet. The repository types and repository formats used one type of font while the rest of the stylesheet used another.

Another thing that the testers observed was that there were inconsistencies with the data. This was due to the fact each repository did their own data entry in the electronic surveys. One of the problems noticed was that some repositories left important fields blank, such as address and repository name. By viewing the records in the stylesheet, the testers could also see that there was a problem of non-standardization with some of the data. For example, the word “California” was typed three different ways: “California,” “Calif.,” and “CA.” There were also standardization problems with the way names were entered into the holdings fields. Some repositories typed the first name followed by the last name while others typed the last name followed by a comma then the first name. While the goal of testing the stylesheet was not to solve the data problems, it still served a useful purpose. It alerted the testers from the SCA Publications Committee that there is a data problem and that it will have to be solved before the repository directory goes public.

The testers’ constructive criticism of the wireframe Website and stylesheet was used to create the subsequent demo Website. The Web pages were redesigned to improve the presentation of search results. Additional Web pages were also created for searches on archival formats. The demo stylesheet also used the SCA color pallet so that the online directory blended in nicely with other pages of the SCA Website. The layout was greatly improved due to the regrouping of like fields and rearranging the data to fit on one screen. The consistent use of the SCA’s default font also led to a dramatically improved stylesheet. The user testing of the wireframe Website and stylesheet were of

great significance to the project. The testers' critiques are directly responsible for the very professional-looking demo Website and stylesheet.

As the testers' recommendations were being incorporated into the demo Website and second stylesheet, records were also being added to the data file. Finally on April 22, 2004, with 80 records now in the data file, the demo Website was presented to the SCA Board of Directors at their annual meeting. Although the board was aware of the Online Repository Directory project, it was their first opportunity to see how much progress had actually been made. During the presentation each screen of the demo Website was shown. Then various types of searches were performed to demonstrate the search interface as well as the layout of the second stylesheet. Interestingly, the records of two board members' institutions happened to be in the data file. Both of these records were displayed for the board members. One record was retrieved by searching on the city, and the other record was retrieved after searching by repository name.

The presentation of the online repository directory received an enthusiastic response from the SCA Board of Directors. After the presentation, the group was asked for their comments and suggestions. Although the Online Repository Directory was originally going to be incorporated into SCA's Website, the President thought it would be better if it were given its own URL. He thought this would make it easier for people to locate it on the Internet and keep it from getting buried in the SCA Website. Consequently, the board is now looking into obtaining another SCA URL for the repository directory. The board was also pleased with the layout of the stylesheet. One suggestion was to reduce navigation by redesigning some of the Web pages to allow more search options from

fewer pages. Another suggestion was to include a statement on the home page, explaining the nature of the repository directory to the users so that they understood that it was not intended to be used as a finding aid to specific archival collections. The statement should also let the user know that they could contact the repository or Online Archive of California for more in-depth information about specific holdings. At the end of the presentation the board members were given the URL of the demo Website so they could explore it at their convenience and respond later with any additional feedback. To date only one additional comment has been received from the SCA board since the demo Website was presented. Clearly the officers are satisfied with the Website, stylesheet, and data file. They would like to see the Website go public as soon as possible, adding new records as they become available. The board's comments were very encouraging and significant. Just as the comments from the user testing of the wireframe Website led to improvements in the demo Website, the feedback from the SCA board members will ultimately make the online repository directory a better Website.

CHAPTER 5

CONCLUSION

The success of the online repository directory can be attributed to the amount of groundwork that went into its development. Preliminary studies by the SCA Publications Committee in 2002 revealed that archival organizations in the United States were starting to post information about repositories in their jurisdiction on the World Wide Web. This led to more in-depth research of domestic and international archival Websites and their methodologies, which, in turn, led to an exploration of possible methodologies suited for putting the SCA's print directory online. After it was proven that an online repository directory was not only possible to create but could be better than other online directories, the actual design process began.

The online repository directory addressed many of the problems inherent in the directory's previous print editions. Because an electronic data file is easier to update, the online directory now offers more accurate and timely information. It is also superior to other archival societies' Websites by offering more powerful searching capabilities in addition to providing more information about repositories and their holdings. While the large amount of research involved in studying other organizations and exploring different methodologies gave an idea of the type of work required to create the online repository directory, it did not predict the tremendous amount of work it actually entailed.

Although the creation of the online repository was a larger undertaking than originally anticipated, it still was an invaluable experience. Overcoming many of the challenges encountered while working on the directory led to greater insight as to how it

could be better designed and maintained. Many of the lessons learned while working on the repository directory are discussed below. The discussion is followed by recommendations and future plans for the online directory's on-going development.

Discussion

As can be expected with any large project, many issues arose during the creation of the online repository directory. Challenges were present during all phases, from the initial data processing through the last step of usability testing. At times it was necessary to not follow the planned methodology. The unresolved issues did not affect the functionality of the Website and will be dealt with at an appropriate time in the future. Nevertheless, in the end the creation of the online repository directory succeeded. One of the most challenging parts of this project was the initial phase of data processing. The mailing of surveys to more than 1,000 California's archival repositories alone was significantly labor intensive. In collecting the data, the SCA allowed repositories to complete the surveys by mail or electronically via e-mail. This option, determined before the Website methodology was established, created problems for the web designer using the data.

Incorrect data was a problem shared by both the mailed and electronic surveys. Often important fields, such as the repository name or address, were left blank. Another common problem was that the data was incorrect. This was mostly noticeable with URLs. Many URLs did not function because they were either non-existent or had typographical errors. Approximately 20-25% of the URLs in the current data file were found to be invalid and had to be corrected. Another reason for incorrect data was that

certain survey questions were not understood by the individuals answering them. For example, one repository stated that the linear feet of their collection was a 9' x 12' room. When data errors such as these occur in the future, follow-up work will be needed to ensure that the online directory has accurate and standardized information.

Other data problems arose when the typing of entries was inconsistent. This happened because the repositories did not have standardized guidelines to follow when entering the data. The three most common problems were with the state name, entry of telephone numbers, and people's names. As noted earlier, there were many variations in the way the word "California" was typed. For example, some chose to spell the word out, some abbreviated it to "Calif.", while others used "CA." Similarly, there were variations in the way telephone numbers were typed. Some used spaces to separate number blocks; for example, in a number like "714 555 1212," some used hyphens, while others chose periods to separate segments. There were also variations in the way telephone prefixes were typed. Some prefixes appear within parentheses while others do not. Lastly, people's names were typed in different styles. This was mostly noticeable in the holdings fields of family names and individual names. Some names were typed with the first name followed by last name, while other names were typed with the last name followed by a comma and then the first name. While none of these standardization problems prevent the online directory from functioning, the inconsistent appearance causes it to have an unprofessional look.

Although the repositories cannot be faulted for data style inconsistencies, some of the repositories were guilty of not following directions. The most common error

occurred when describing the repository type. Each repository was told to select no more than two repository types from a list provided. Some repositories entered three repository types. Similarly, when repositories were asked to enter one e-mail address for a contact person, some repositories submitted two or more. In both of these cases, the extra information caused problems for the data structure. Since several repositories entered three repository types, a third repository type attribute has been added to the DTD. The extra repository type will remain in the record temporarily until the repository can be contacted to clarify its repository type. In the case of the extra e-mail address, the second address was deleted from the record. The decision to delete the second e-mail address was based on the fact that it did not make sense for a repository to have more than one contact person. Also, since most repositories only provided one contact e-mail address, changing the entire XML structure to accommodate only a few records was not warranted.

The use of checkboxes on the survey made data entry easier for the repositories, but negatively affected data processing. Many problems occurred when the checkboxes were converted from Filemaker into XML. They generated an excessive amount of encoding which was difficult to read through. The unnecessary encoding also slowed down the data file. Another problem of the checkboxes was that the corresponding stylesheet required an excessive amount of encoding. Finally, the checkboxes did not work with the JavaScript search interface. Because the checkboxes caused so many problems, they were not used in the XML structure. Instead the XML structure was modified to include the data represented by the checkboxes.

The checkboxes were not the only feature changed when the data file was converted from Filemaker to XML. One field, “Other Services,” was deleted altogether. After viewing the records of over 100 surveys, it became apparent that repositories were not using this field. Most likely the repositories considered the basic “Services” field sufficient to contain all information about a repository’s services and that the “Other Services” field was redundant. However, should data be found in the “Other Services” field, it would be manually typed into the “Services” field. Another Filemaker field, “_Contact,” was also deleted because it contained redundant information. This field was used to list all the ways a repository could be contacted such as mail, e-mail, and fax. However, since some of this information resides in other fields, this field was not really necessary.

Another change that occurred was the renaming of the field “URL” in Filemaker to “Web” in XML. This was done because the stylesheet would not display the data in a field entitled “URL.” However, after the name was changed to “Web” the stylesheet functioned properly.

In addition to these modifications, one final change was made to the XML structure. While designing the stylesheet it was noticed that there was no field name or element to indicate the mailing address’s state. The state name was initially omitted because it was assumed that the state was “California.” However, without the state included in the mailing address, mailing labels could not be generated. Therefore the element <_State> was added to hold this information in the directory.

Although the changes made to the XML structure were modified from the original Filemaker structure due to fields being added, removed, or renamed, it is interesting to note how closely the XML structure reflects the proposed fields identified in Chapter 3. There is one difference, however, that should be discussed. The project methodology mentioned including a field for controlled subject headings; this did not occur. Instead, the subject terms that were provided by repositories are the ones currently in use for searching. It was decided that it would be better to develop a controlled vocabulary in a thesaurus at a later date. The original thought was that a controlled vocabulary would be developed after looking at the subject terms provided by the repositories. Unfortunately, not enough surveys have been returned from the repositories to accomplish this task. Once more data is available, a controlled vocabulary will be established.

Another difference between the planned methodology and the actual implementation occurred with the number of testers forecast in the usability testing. Although the plan was to have ten users in the usability test, due to time constraints and scheduling problems, only three testers were available for the wireframe test. In hindsight though, ten might have been too many for this initial phase. Trying to incorporate the comments of ten people into a single stylesheet or Website might have been overwhelming. Fortunately, the feedback from the small group of testers was sufficient. Feedback from the SCA Board of Directors was also not planned as part of the usability testing. However, the demo Website was completed at the same time as the SCA annual meeting and the opportunity presented itself. As more people in the archival

community become aware of the online repository directory, it will get easier to do periodic usability testing when new features and design changes are implemented.

Recommendations

The following are recommendations for preparing the online directory for its upcoming public release as well as for long-term maintenance. These recommendations are based on the experience of getting the online repository directory to its current state and are intended to serve as a guideline for future work on the project. Issues addressed include software, procedures, data standardization, staff duties, and maintenance.

Software

1. XML Spy – Two different XML editors were used on the repository database, XML Spy and Xmetal. There was a notable difference in performance of two products. When data was converted from Filemaker to XML, Xmetal concatenated it into one string making it impossible to edit without great difficulty. XML Spy, however, handled the conversion without difficulty. Each Filemaker field became an XML element and was transferred to its own line. This made the conversion process quite easy. The menus in XML Spy are also somewhat easier to navigate. It is recommended that XML Spy be used as the XML editor for the SCA's online repository directory.
2. Filemaker – Filemaker is no longer needed for storing the repositories' records now that the XML structure has been defined with a DTD. Having the records reside in two data files (Filemaker and XML) is not practical and can lead to confusion when

updating records. Therefore, all existing Filemaker records that have not been converted to XML need to be converted and validated against the XML DTD.

Procedures

1. Electronic surveys – A new electronic survey form needs to be designed. Although the electronic surveys were instrumental in receiving data from the repositories, the fields they generated did not meet the needs of the repository directory. A new survey form should be developed based on the DTD and posted on a Web page. Included on a SCA Web page would be an example of a survey form filled out with sample data for reference. Required fields would be indicated with an asterisk next to them. This would lessen the chance of important fields such as repository name or repository address being left blank. The survey's Web page should also post the name of an SCA representative who could be contacted by phone and/or e-mail to answer questions about the survey.
2. Survey submission – Any repository desiring to submit their data electronically could copy the blank survey form into an MS Word document and fill it out. Once the survey was filled out, the repository would return it to the SCA as an attachment to e-mail. The SCA would then modify the form by assigning a unique repository number to the record and checking that all required fields were filled in. In the event of missing data, the repository would be contacted. The record would then be validated against the DTD and appended to the rest of the data file. By gathering data using a new survey form, future data collection will be faster, easier, and more accurate.

3. Mailed surveys – In addition to designing a new electronic form, a new hard copy survey needs to be designed. This new form should be modeled after the DTD to make it easier for data entry into XML. The hard copy survey should also include a name and phone number and/or e-mail of an SCA representative who could help answer any questions a repository might have.

Data Standardization

1. Data inconsistencies – One problem with the records is the inconsistent way in which the data appears. Before the online repository directory is released to the public, these inconsistencies must be resolved as noted above. Standardization needs to be applied to the way addresses and telephone numbers are entered. Also, it must be decided whether or not the terms used in the holdings fields will be typed entirely in lower case or with initial capitalization. Finally, all personal names must be typed in a consistent format. Once the records are in a standardized format, the online directory will have more credibility as a professional document.

Personnel

1. Tasks – Several duties are required in maintaining the online repository directory. Because these tasks would be performed by volunteers on a part-time basis, they could be divided among three or four people. Needed in the maintenance of the directory are the following: 1) a person the repositories can contact if they have questions about the Website or filling out the survey; 2) a proofreader to verify that all data, including e-mail and URL hyperlinks, are correct; 3) at least one

person trained in using the DTD within XML Spy; and 4) one or two people empowered to upload the files to the SCA fileserver.

Maintenance

1. Updating records – On a yearly basis the repositories in the directory should be contacted to see if there are any changes to their records, such as contact information, holdings and formats.
2. Backup – To prevent records from being lost or corrupted, a backup of the data file will be made every time it is updated. One backup copy will reside on the SCA file server and another will be kept on a CD.
3. Uploading – The data file will be updated on a regular basis as needed, and uploaded to the SCA file server on a monthly basis.

Future Plans

Work on the online repository directory will not end after it is released to the public later this year. It will be a constantly evolving Website, striving to meet the needs of the layperson and the archival professional. Although the directory does function in its current state, there are many improvements that can be made that will make it an even more powerful tool. It is hoped that within a year of the directory's release, the following will be accomplished:

1. Thesaurus – As more repositories submit their surveys, a controlled vocabulary of holdings descriptors will be developed to enable more precise and accurate searches.

2. Additional fields – Two additional fields containing driving directions as well as repository images (when available) will be added to the data file. This will help researchers physically find a repository without having to contact the repository directly for directions.
3. Additional stylesheets – The current repository directory only uses one stylesheet. However, additional stylesheets can be designed to display selected fields for the users. Customized stylesheets will also be designed for the SCA Publications Committee and support staff. These stylesheets will perform calculations to track statistical information and will display specific fields when performing maintenance tasks, such as hyperlink verification.
4. Printing options – Because the online repository directory will replace the Filemaker database, it needs to be able to perform all of the Filemaker printing functions. This includes the ability to print mailing labels from the data file as well as the option of printing the directory or sections on demand.
5. Website enhancements – The current Website will have added features including a hyperlink to the Online Archive of California from the online repository directory's main menu. Web pages will be consolidated and the search interface will be modified in order to provide the capability of performing searches with easier navigation.
6. Training – During the next year I will train at least one other person in all duties required to support the directory so that it will continue to be maintained after my departure from the project.

Significance of the Project

The successful creation of the SCA's online repository directory is a significant accomplishment. For the researcher it offers instantaneous information about California's many archival and manuscript repositories. For the SCA it is the long-awaited solution for managing and making available repository data in a timely and efficient manner. Most important, the online repository directory represents a major improvement on previous methods used by other archival societies. Not only does it contain more detailed information about each repository, it also offers additional and more precise searching capabilities. Once it is made public later this year, SCA's online repository directory can serve as a model for other archival societies to emulate.

Appendix A

Society of California Archivists
Directory of Archival and Manuscript Repositories in California
Data Sheet

Repository Name (e.g., Special Collections):

Parent (or Sponsoring) Institution (e.g., Oxnard Public Library):

Contact Information (Please specify individual and/or office which serves as first point of contact, e.g., Local History Curator):

Name _____ (and/or) Title _____

Street Address (line 1):

Street Address (line 2):

City:

Zip Code:

County:

Mailing Address, line 1 (if different from above):

Mailing Address (line 2):

City:

Zip Code:

County:

Telephone:

Fax:

Email:

URL:

Access to Facility and Holdings (Please check all that apply):

Is your repository open to researchers and/or the public? ☐ Yes ☐ No

Are there any user restrictions (please specify)?

What hours are the repository generally open to the public?

How can users contact your repository?

☐ Telephone

☐ Mail

☐ E-mail

What services does your repository provide?

☐ Photocopying

☐ Scanning

☐ Microfilming

☐ Fee-based research services

☐ Other services (please specify):

Please note any restrictions or conditions that users should know about when using your facility:

Institution or Type of Repository (Please check no more than 2):

☐ Academic (university or college)

☐ Financial

☐ Business/Corporate

☐ Entertainment

☐ Film

☐ Dance

☐ Opera

☐ Orchestra/Symphony

☐ Radio

☐ Television

☐ Theater

- ☐ Government
- ☐ City
- ☐ County
- ☐ Federal
- ☐ State
- ☐ Historical Site
- ☐ Historical Society
- ☐ Law
- ☐ Medical
- ☐ Military
- ☐ Museum or Cultural Institution
- ☐ Newspaper
- ☐ Professional or Trade Organization
- ☐ Private Library
- ☐ Public Library
- ☐ Religious
- ☐ Research Institution (such as botanical, not university or college)
- ☐ Social/Fraternal Organization
- ☐ Transportation
- ☐ Other (please specify):

Holdings (Please submit a short description of your major collecting areas and notable holdings):

Descriptors (Please list the major subjects, geographical areas, individual and family names, and organizations represented in your collections. This information will be used in compiling the directory's indexes):

Subjects:

Geographical areas (regions, counties, or towns):

Individuals:

Families:

Organizations:

Extent of Holdings in Linear Feet:

Bulk Dates (major time periods covered by holdings):

Dates of Oldest Records:

Formats (refers to the major *types* materials in your repository rather than the subject matter of your collections):

- ___ Architectural Drawing, Models, etc.
- ___ Art Works
 - ___ Paintings
 - ___ Drawings
 - ___ Photographs
- ___ Posters (including lithographs)
- ___ Artifacts
 - ___ Books
- ___ Digital works (created and/or stored digitally such as Websites, online periodicals, digitally created artwork, music)
- ___ Electronic records (e-mails, PDF files, spreadsheets, databases, etc.)
- ___ Ephemera (advertisements, program notes, tickets, flyers)
- ___ Films/videotapes
- ___ Genealogical information
- ___ Institutional/business records (not generated electronically)
- ___ Legal Documents
 - ___ Maps
 - ___ Music
 - ___ Sheet music or manuscripts
- ___ Phonograph records
- ___ Audiotapes (reel-to-reel or cassette)
- ___ CDs
- ___ Digital Audiotapes (DAT)
- ___ DVDs
 - ___ Microform
 - ___ Newspapers
 - ___ Oral Histories

- ☐ Pamphlets
- ☐ Periodicals
- ☐ Personal/family papers (including diaries, scrapbooks, letters, etc.)
- ☐ Photographs
- ☐ Specialty papers (vellum, papyrus, etc.)
- ☐ Subject Files

Finding Aids:

- ☐ Inventories and guides
- ☐ Databases
- ☐ Card catalog
- ☐ OPAC
- ☐ Repository Website
- ☐ Published guides to your collections (please provide full bibliographic citations, if possible):

Union Lists or Networks to which your institution contributes:

- ☐ Online Archive of California
- ☐ National Union Catalog of Manuscript Collections (1959-1993)
- ☐ OCLC
- ☐ RLIN
- ☐ Archives USA
- ☐ Other (please specify):

Appendix B

Example of a Filemaker Record

Repository	Levi Strauss & Co. Archives
Parent Inst	Levi Strauss & Co.
Contact	Lynn Downey
Title	Historian
Address 1	1155 Battery Street
Address 2	
Address City	San Francisco
Address State	CA
Address ZIP	94111
County	San Francisco
Mailing 1	
Mailing 2	
_City	
_ZIP	
_County	
Telephone	415-501-6577
FAX	415-501-6443
Email	ldowney@levi.com
Web	www.levistrauss.com
Open	no
Requirements	
Hours	
_Contact	phone, mail, email
Services	photocopying
Other Serv.	null
Academic	null
Financial	null
Business	checkbox
Entertainment	null

Film null
 Dance null
 Opera null
 Orchestra null
 Radio null
 TV null
 Theater null
 Other null
 Government null
 City null
 County null
 Federal null
 State null
 Hist Site null
 Hist Society null
 Law null
 Medical null
 Military null
 Museum null
 Newspaper null
 Prof null
 Private null
 Public null
 Religious null
 Research null
 Social null
 Trans null
 _Other null

Specify other

Holdings _Narrative_ Garments, documents, photographs, artifacts and marketing materials

Subjects Costume and textiles, Advertising/Marketing,
 _Subjects Advertising/Marketing, Dry goods, Popular
 f3 Dry goods
 f4 Popular culture
 f5 Gold Rush
 f6
 f7 Manufacturing
 f8
 f9 Dude ranches
 f10
 Geo California, San Francisco, Western U.S.
 _Geo San Francisco
 f11 Western U.S.
 f12
 f13
 f14
 f15
 f16
 f17
 f18
 Individuals Levi Strauss, Sigmund Stern, Elise Stern Haas,
 _Individuals Sigmund Stern,
 f19 Elise Stern Haas,
 f20 Haas, Walter
 f21 Haas, Walter Jr.
 f22 Haas, Peter
 f23 Haas, Robert
 f24 Daniel Koshland
 f25
 f26

Families	Stern, Haas
_Families	Haas
f27	
f28	
f29	
f30	
f31	
f32	
f33	
f34	
Orgs	Levi Strauss & Co., Levi Strauss Foundation
_Orgs	
f35	Levi Strauss Foundation
f36	
f37	
f38	
f39	
f40	
f41	
f42	
Linear feet	N/A
Bulk Dates	1854-present
Oldest	1854
Architectural	checkbox
Art	null
Paintings	checkbox
Drawings	null
Photographs	checkbox
Posters	checkbox
Artifacts	checkbox

Books	checkbox
Digital	null
Electronic	null
Ephemera	checkbox
Films	checkbox
Genealogy	null
Institutional	checkbox
Legal	checkbox
Maps	null
Music	null
Sheet music	null
Records	null
Audiotapes	null
CDs	null
DAT	null
DVD	null
Microform	null
Newspapers	checkbox
Oral Histories	checkbox
Pamphlets	checkbox
Periodicals	checkbox
Personal papers	null
_Photographs	checkbox
Sound recordings	checkbox
Papers	null
Subject Files	checkbox
f43	null
Specify	
Inventories	checkbox
Databases	null

Card Catalog	null
OPAC	null
Website	null
Pub guides	null
Citation	
OAC	null
NUC	null
OCLC	null
RLIN	null
ArchivesUSA	null
f44	null
_Specify	
Comments	Imported from Excel, 07/11/03, JG
Record ID	0051

Appendix C

Example of an XML Record

```

<ROW RECORDID="0051">
  <Names>
    <Repository>Levi Strauss & Co. Archives</Repository>
    <Parent_Inst>Levi Strauss & Co.</Parent_Inst>
  </Names>
  <Contact>Lynn Downey</Contact>
  <Title>Historian</Title>
  <Address_1>1155 Battery Street</Address_1>
  <Address_2/>
  <Address_City>San Francisco</Address_City>
  <Address_State>CA</Address_State>
  <Address_ZIP>94111</Address_ZIP>
  <County>San Francisco</County>
  <Mailing_1/>
  <Mailing_2/>
  <_City/>
  <_ZIP/>
  <_County/>
  <_State/>
  <Telephone>415-501-6577</Telephone>
  <FAX>415-501-6443</FAX>
  <Email>ldowney@levi.com</Email>
  <Web_url="http://www.levistrauss.com">www.levistrauss.com</Web>
  <Open>no</Open>
  <Requirements/>
  <Hours/>
  <_Contact>phone, mail, email</_Contact>
  <Services>photocopying</Services>
  <Reptypes Type1="Business"></Reptypes>
  <Holdings>
    <Holdings__Narrative_>Garments, documents, photographs, artifacts and marketing materials relating
to the history of Levi Strauss & Co., 1854-present</Holdings__Narrative_>
    <Subjects>Costume and textiles, Advertising/Marketing, Dry goods, Popular culture, Gold Rush,
Manufacturing, Dude ranches </Subjects>
    <Geo>California, San Francisco, Western U.S.</Geo>
    <Individuals>Levi Strauss, Sigmund Stern, Elise Stern Haas, Walter Haas, Walter Hass Jr., Peter
Haas, Robert Haas, Daniel Koshland </Individuals>
    <Families>Stern, Haas</Families>
    <Orgs>Levi Strauss & Co., Levi Strauss Foundation</Orgs>
  </Holdings>
  <Linear_feet>N/A</Linear_feet>
  <Bulk_Dates>1854-present</Bulk_Dates>
  <Oldest>1854</Oldest>
  <Formats>Architectural, Paintings, Photographs, Posters, Artifacts, Books, Ephemera, Films, Institutional,
Legal, Newspapers, Oral Histories, Pamphlets, Periodicals, Sound Recordings, Subject Files
  </Formats>
  <Citation/>
  <Finding_Aids>Inventories</Finding_Aids>
  <Networks/>
</ROW>

```

Appendix D

XML DTD and Commentary

```

<!ELEMENT SCA (ROW)*>
<!ELEMENT ROW (Names, Contact, Title, Address_1, Address_2, Address_City, Address_State, Address_ZIP,
County, Mailing_1, Mailing_2, _City, _ZIP, _County, Telephone, FAX, Email, Web, Open, Requirements, Hours,
_Contact, Services, Reptypes, Holdings, Linear_feet, Bulk_Dates, Oldest, Formats, Citation, Finding_Aids,
Networks)>
<!ATTLIST ROW
  RECORDID CDATA #REQUIRED>
<!ELEMENT Names (Repository, Parent_Inst)>
<!ELEMENT Repository (#PCDATA)>
<!ELEMENT Parent_Inst (#PCDATA)>
<!ELEMENT Contact (#PCDATA)>
<!ELEMENT Title (#PCDATA)>
<!ELEMENT Address_1 (#PCDATA)>
<!ELEMENT Address_2 (#PCDATA)>
<!ELEMENT Address_City (#PCDATA)>
<!ELEMENT Address_State (#PCDATA)>
<!ELEMENT Address_ZIP (#PCDATA)>
<!ELEMENT County (#PCDATA)>
<!ELEMENT Mailing_1 (#PCDATA)>
<!ELEMENT Mailing_2 (#PCDATA)>
<!ELEMENT _City (#PCDATA)>
<!ELEMENT _ZIP (#PCDATA)>
<!ELEMENT _County (#PCDATA)>
<!ELEMENT Telephone (#PCDATA)>
<!ELEMENT FAX (#PCDATA)>
<!ELEMENT Email (#PCDATA)>
<!ELEMENT Web (#PCDATA)>
<!ATTLIST Web url CDATA #IMPLIED>
<!ELEMENT Open (#PCDATA)>
<!ELEMENT Requirements (#PCDATA)>
<!ELEMENT Hours (#PCDATA)>
<!ELEMENT _Contact (#PCDATA)>
<!ELEMENT Services (#PCDATA)>
<!ELEMENT Reptypes (#PCDATA)>
<!ATTLIST Reptypes
  Type1 CDATA #IMPLIED
  Type2 CDATA #IMPLIED>
<!ELEMENT Type (#PCDATA)>
<!ELEMENT Holdings (Holdings__Narrative_, Subjects, Geo, Individuals, Families, Orgs)>
<!ELEMENT Holdings__Narrative_ (#PCDATA)>
<!ELEMENT Subjects (#PCDATA)>
<!ELEMENT Geo (#PCDATA)>
<!ELEMENT Individuals (#PCDATA)>
<!ELEMENT Families (#PCDATA)>
<!ELEMENT Orgs (#PCDATA)>
<!ELEMENT Linear_feet (#PCDATA)>
<!ELEMENT Bulk_Dates (#PCDATA)>
<!ELEMENT Oldest (#PCDATA)>
<!ELEMENT Formats (#PCDATA)>
<!ELEMENT Citation (#PCDATA)>
<!ELEMENT Finding_Aids (#PCDATA)>
<!ELEMENT Networks (#PCDATA)>

```

Elements are defined using the !ELEMENT declaration and attributes are defined using the !ATTLIST declaration. The root element of the XML data file is <SCA>. Each record within <SCA> is contained inside the <ROW> element. Within <ROW> is a RECORDID attribute where each record is assigned a unique record number. This is a required attribute. Each <ROW> must contain a record number. The <ROW> element contains the following children elements: Names, Contact, Title, Address_1, Address_2, Address_City, Address_State, Address_ZIP, County, Mailing_1, Mailing_2, _City, _ZIP, _County, Telephone, FAX, Email, Web, Open, Requirements, Hours, _Contact, Services, Reptypes, Holdings, Linear_feet, Bulk_Dates, Oldest, Formats, Citation, Finding_Aids, Networks. Some of the child elements of <ROW> have their own children. <Names> is a parent of <Repository> and <Parent_Inst>. Likewise <Holdings> is a parent of <Holdings__Narrative_>, <Subjects>, <Geo>, <Individuals>, <Families>, and <Orgs>.

Appendix E

Stylesheet Encoding

```
<?xml version="1.0"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:template match="/">
    <html>
      <head>
        <meta content=""></meta>

      </head>
      <body>
        <center>
          <h2>Search Results:</h2>
        </center>
        <table border="3" cellpadding="3" cellspacing="2" width="100%" bgcolor="#FFFFCC">
          <tr/>
          <td width="100%">
            <xsl:for-each select="//ROW">
              <xsl:sort select="Record_ID" order="descending"/>
              <xsl:sort select="."/>
              <table border="3" cellpadding="2" cellspacing="5" width="100%" bgcolor="#FFFFCC">
                <tr>
                  <xsl:apply-templates select="Record_ID"/>
                </tr>
              </table>
            </xsl:for-each>
          </td>
        </table>
      </body>
    </html>
  </xsl:template>
  <xsl:template match="Record_ID">
    <td width="45%" valign="top" bgcolor="#C0C0C0">
      <font color="#666666" face="Arial" size="1">Record:&#160;<xsl:value-of select="."/>
      </font>
      <br/>
      <font color="#000080" face="Arial" size="2">
        <b>
          <xsl:value-of select=" ../Names/Repository"/>
        </b>
      </font>
      <br/>
      <font color="#000080" face="Arial" size="1">
        <xsl:value-of select=" ../Address_1"/>&#160;
        <xsl:value-of select=" ../Address_2"/>&#160;<br/>
        <xsl:value-of select=" ../Address_City"/>&#160;CA&#160;&#160;
        <xsl:value-of select=" ../Adress_ZIP"/>&#160;<br/>
      <br/>
      <font color="#333333">
        <b>Type -<xsl:value-of select=" ../Reptypes"/>
        </b>
      </font>
      <br/>
      <br/>
      <b>Contact:&#160;</b>
      <xsl:value-of select=" ../Contact"/>&#160;-
      <xsl:value-of select=" ../Title"/>
    </td>
  </xsl:template>
</xsl:stylesheet>
```

```

<br/>
<br/>
<b>Phone: &#160;</b>
<xsl:value-of select=" ../Telephone"/>&#160;&#160;&#160;&#160;&#160;<b>Fax: &#160;</b>
<xsl:value-of select=" ../FAX"/>
<br/>
<b>Email: &#160;</b>
<a>
  <xsl:attribute name="href">mailto:<xsl:value-of select=" ../Email"/></xsl:attribute>
  <span class="email">
    <xsl:value-of select=" ../Email"/>
  </span>
</a>
<br/>
<b>URL: &#160;</b>
<a>
  <xsl:attribute name="href"><xsl:value-of select=" ../Web/@url "/></xsl:attribute>
  <xsl:value-of select=" ../Web"/>
</a>
<br/>
<br/>
</font>
<font color="#666666" face="Arial" size="1">
  <xsl:value-of select=" ../Mailing_1"/>&#160;
  <xsl:value-of select=" ../Mailing_2"/>&#160;<br/>
  <xsl:value-of select=" ../_City"/>&#160;&#160;&#160;
  <xsl:value-of select=" ../_ZIP"/>&#160;<br/>
</font>
</td>
<td width="58%" rowspan="2" valign="top" align="justify">
  <font color="#000080" face="Arial" size="1">
    <p align="justify">
      <b>Holdings Narrative: &#160;</b>
      <xsl:value-of select=" ../Holdings__Narrative_"/>
    </p>
    <p align="justify">
      <b>Subjects: &#160;</b>
      <xsl:value-of select=" ../Subjects"/>
    </p>
    <p align="justify">
      <b>Geographical Areas: &#160;</b>
      <xsl:value-of select=" ../Geo"/>
    </p>
    <p align="justify">
      <b>Individuals: &#160;</b>
      <xsl:value-of select=" ../Individuals"/>
    </p>
    <p align="justify">
      <b>Families: &#160;</b>
      <xsl:value-of select=" ../Families"/>
    </p>
    <p align="justify">
      <b>Organizations: &#160;</b>
      <xsl:value-of select=" ../Orgs"/>
    </p>
    <p>
      <b>Linear Feet: &#160;</b>
      <xsl:value-of select=" ../Linear_feet"/>&#160;&#160;&#160;<b>Bulk
Dates: &#160;</b>&#160;<xsl:value-of select=" ../Bulk_Dates"/>&#160;&#160;&#160;
      <b>Oldest: </b>&#160;<xsl:value-of select=" ../Oldest"/>
    </p>
  <p align="justify">

```

```

        <b>Formats:&#160;</b>
        <xsl:value-of select=" ../Formats"/>
    </p>
    <p>
        <b>Published Guides:&#160;</b>
        <xsl:value-of select=" ../Citation"/>
    </p>
    <p>
        <b>Finding Aids:&#160;</b>
        <xsl:value-of select=" ../Finding_Aids"/>
    </p>
    <p>
        <b>Union Lists or Networks:&#160;</b>
        <xsl:value-of select=" ../Networks"/>
    </p>
</font>
</td>
<tr>
    <td width="42%" valign="top" bgcolor="#000080">
        <font color="#FFFFCC" size="1">
            <p>
                <b>Open To Public:&#160;</b>
                <xsl:value-of select=" ../Open"/>
            </p>
            <p>
                <b>Hours:&#160;</b>
                <xsl:value-of select=" ../Hours"/>
            </p>
            <p>
                <b>Requirements:&#160;</b>
                <xsl:value-of select=" ../Requirements"/>
            </p>
            <p>
                <b>Services:&#160;</b>
                <xsl:value-of select=" ../Services"/>
            </p>
        </font>
    </td>
</tr>
</xsl:template>
<xsl:template match="Parent_Inst">
    <td style="color:darkblue; font-family:Tahoma; font-size:8pt;background:lightyellow;fontstyle:bold"
    colspan="2">
        <b>Parent Institution:
    <xsl:value-of select="."/>
        </b>
    </td>
</xsl:template>
<xsl:template match="Email">
    <td style="color:red; font-family:Tahoma; font-size:10pt;background:lightyellow">
        Email:&#160;&#160;
        <a>
            <xsl:attribute name="href">mailto:<xsl:value-of select="."/></xsl:attribute>
            <span class="email">
                <xsl:value-of select="."/>
            </span>
        </a>
    </td>
</xsl:template>
</xsl:stylesheet>

```


REFERENCES

- ANABAD – Asociacion de Archiveros Bibliotecarios, Museologos y Documentalistas. (2001). Retrieved December 15, 2002, from <http://www.anabad.org/>
- Archivenet. (n.d.). Retrieved December 15, 2002, from <http://www.archiefnet.nl/index.asp?taal=en>
- ArchivesUSA. (n.d.). Retrieved August 29, 2003, from <http://archives.chadwyck.com/repository/search>
- Archivists Round Table of Metropolitan New York. (2003, October 2). Retrieved October 10, 2003, from <http://www.nycarfchivists.org/guide/rguide.html>
- ARCHON directory. (2002). *ARCHON A-Z repositories*. Retrieved October 10, 2003, from <http://www.hmc.gov.uk/archon/archondirectory.htm>
- ARCHON directory. (2002). *ARCHON code: 0336*. Retrieved October 10, 2003, from <http://www.hmc.gov.uk/archon/searches/locresult.asp>
- ARCHON directory. (2002). *Search ARCHON*. Retrieved October 10, 2003, from <http://www.hmc.gov.uk/archon/archondirectory.htm#search>
- Associacao dos Archivistas Brasileiros. (n.d.). *Home page*. Retrieved December 15, 2002, from <http://www.aab.org.br/>
- Australian Society of Archivists. (2003, May). *Introduction to the WWW edition*. Retrieved August 29, 2003, from http://www.archivists.org.au/directory/asa_intw.htm
- BCAIN (British Columbia Archival Information Network). (2002, December). *Home page*. Retrieved December 15, 2002, from <http://aabc.bc.ca/aabc/bcain.html>
- Boynton, J. R. (1998, May). *The new publishing: An information choreographer's manifesto*. Paper presented at the W3 Shaping The Future of HTML Conference, San Francisco, CA. Retrieved December 15, 2002, from <http://www.tdl.com/%7Ejrb/icm.html>
- CONARQ (Conselho Nacional de Arquivos). (n.d.). *Public archives web page*. Retrieved December 15, 2002, from <http://www.arquivonacional.gov.br/conarq/qsomos/Listarq.htm>
- European Archival Network (EAN). (n.d.). *Home page*. Retrieved December 15, 2002, from <http://www.european-archival.net/>

- Flynn, P. (Ed.). (n.d.). *The XML FAQ*. Retrieved December 15, 2002, from <http://www.ucc.ie/xml>
- Goldfarb, C. F. (1999). *XML in an instant: A non-geeky introduction*. Retrieved December 15, 2002, from <http://www.xml.org/xml/goldfarb.shtml>
- Gotomedia. (n.d.). *Usability testing*. Retrieved August 10, 2003, from <http://www.gotomedia.com/iceland/usability/>
- Greater New Orleans Archivists (2003, May 19). *Archival research repositories in New Orleans*. Retrieved October 10, 2003, from <http://nutrias.org/gnoa/norepos.htm>
- Hansen, D. (2001, October). Society of California Archivists Publication Committee quarterly report. Fullerton, CA: Author.
- Hansen, D. (2003, September). Society of California Archivists Publication Committee quarterly report. Fullerton, CA: Author.
- IAA (Israel Archives Association). (1996, June). *Home page*. Retrieved December 15, 2002, from <http://www.tau.ac.il/humanities/iaa/index.html>
- LA as Subject. (2003, October 7). *Additional search options*. Retrieved October 10, 2003, from <http://www.usc.edu/isd/archives/arc/lasubject/addsearch.html>
- LA as Subject. (2003, October 7). *Index page*. Retrieved October 10, 2003, from <http://www.usc.edu/isd/archives/arc/lasubject/index.html>
- L'ANAI (Associazione Nazionale Archivistica Italiana). (n.d.). *Home page*. Retrieved December 15, 2002, from <http://www.anai.org/anai/>
- Louisiana Archives and Manuscripts Association. (1998, November). *LAMA directory*. Retrieved, October 10, 2003, <http://nutrias.org/lama/guide/guide.htm>
- Michigan Archival Association. (2003). *Publications page*. Retrieved October 10, 2003, from <http://maa.hypermart.net/publications.htm>

Michigan Archival Association. (2003). *Michigan archival institutions page*, 3(1). Retrieved October 10, 2003, from <http://maa.hypermart.net/archives.htm>

National Archives of Norway. (2002, May). *Home page*. Retrieved December 15, 2002, from <http://www.riksarkivet.no/english/about/contact.html>

Oracle Corporation. (2004). *Products page*. Retrieved February 3, 2004, from http://www.oracle.com/database/index.html?Std_One.html

Salminen, A. (2003, October 14). *XML family of languages: Overview and classification of W3C specifications*. Retrieved November 6, 2003, from <http://www.cs.jyu.fi/~airi/xmlfamily.html>

SmartSoft Computing Ltd. (n.d.). *A short Oracle tutorial for beginners*. Retrieved January 25, 2004, from <http://www.smart-soft.co.uk/Oracle/oracle-tutorial-part2.htm>

Society of Alabama Archivists. (2003, January 27). *Home page*. Retrieved, October 10, 2003, from <http://www.auburn.edu/sala/>

Society of California Archivists. (1996). *Archival & manuscript repositories in California* (4th ed.). Sacramento: Society of California Archivists.

Society of California Archivists. (2002, April). *Home page*. Retrieved December 15, 2002, from <http://www.calarchivists.org/publications.html>

Society of Florida Archivists. (2003, January 5). *Home page*. Retrieved October 10, 2003, from <http://www.florida-archivists.org/>

Society of Georgia Archivists. (n.d.). *Home page*. Retrieved, October 10, 2003, from <http://www.soga.org/resources/gar.html>

Society of Indiana Archivists. (n.d.). *Indiana archival and historical repositories page*. Retrieved, October 10, 2003, from <http://archives1.archives.nd.edu/sia/guide/reposito.htm>

Society of North Carolina Archivists. (n.d.). *Links to archives and special collections Web pages*. Retrieved October 10, 2003, from <http://www.rtpnet.org/~snca/links.htm>

Society of North Carolina Archivists. (n.d.). *Publications page*. Retrieved October 10, 2003, from <http://www.rtpnet.org/~snca/publica.htm>

Society of Northwest Archivists. (2003, September 2). *Home page*.

Retrieved, October 10, 2003, from

<http://osulibrary.oregonstate.edu/archives/misc/nwa.html>

Society of Ohio Archivists. (n.d.). *Links to websites of interest*.

Retrieved October 10, 2003, from <http://www.ohiojunction.net/soa/links.html>

Society of Rocky Mountain Archivists. (2001, August 10). *Publications page*.

Retrieved October 10, 2003, from <http://ahc.uwyo.edu/srma/publications.htm>

Society of Southwest Archivists. (2003, February 25). *Home page*.

Retrieved October 10, 2003, from <http://info.lib.uh.edu/ssa/ssa.htm>

Society of Tennessee Archivists. (2003, March 7). *Home page*.

Retrieved October 10, 2003, from <http://www.geocities.com/tennarchivists/>

Stanford, J. (2003, January). *HTML wireframes and prototypes: All pain and no gain*. Retrieved August 10, 2003, from

http://www.boxesandarrows.com/archives/html_wireframes_and_prototypes_all_gain_and_no_pain.php

W3 HTML. (2002). *Home page*. Retrieved December 15, 2002, from

from <http://www.w3.org/MarkUp/>

W3 Schools, XML. (n.d.). *Introduction to XML*. Retrieved December 15, 2002,

from http://www.w3schools.com/xml/xml_what_is.asp