

## CS577a Projects for 1996/97 and 1997/98

### **1. Architecture & Fine Arts Databases (Howard Smith)**

The Architecture and Fine Arts Library and Special Image Collections is developing several databases for material unique to that library. These include: artists' books, CD-ROM titles, video holdings, artists represented in the slide collection, and a list of architects and buildings in Los Angeles, among others. The purpose of this project is to develop a system of databases with an interface that is not only user-friendly and graphically well-designed, but a powerful search system.

### **2. Bella Lewitsky Archives (Melinda Hayes)**

The Bella Lewitsky collection includes costumes, photographs, sound recordings, videos, programs, posters, stage plots, and paper files relating to the Bella Lewitsky Dance Company, which was headquartered in Los Angeles. The company was one of the few successful U.S. modern dance companies located outside of New York. The purpose of this project is to make this collection accessible to the world via the Web; one special challenge is to make the costumes digitally accessible.

### **3. Business School Working Papers (Ken Hickson, Julie Kwan, Terry Lichvar)**

Faculty in business schools write "working papers", which are preliminary reports of their research. They then receive critiques about their work and revise their papers accordingly. Producing these papers and making them available to the scholarly community using traditional print methods has been an expensive, time-consuming proposition for the academic community. The purpose of this project is to develop a system for submitting, editing, and making available the full-text of working papers of faculty from the Marshall School of Business accessible from departmental web sites as well as being included in a combined database.

### **4. CNTV Moving Image Archive (Sandra Joy Lee)**

The USC Moving Image Archive houses USC student film and video productions from the 1930's to current productions of the School of Cinema/TV. Due to the School's prominence, there is considerable interest in both current and older student films. The purpose of this project is to provide the database of student films via the Web and to provide the capability to view stills and movie clips.

### **5. EDGAR Corporate Data (Caroline Sisneros)**

The U.S. Securities and Exchange Commission collects and makes available corporate reports (e.g., 10Ks, 10Qs, 8Ks, etc.) through its EDGAR web site <http://www.sec.gov/edgarhp.htm>. The material is available in ASCII format and is difficult to use due to the length of the documents and the extensive tabular data included in them. The purpose of this project is to make this information more readily useable.

### **6. General Library FAQ's (Chris Ferguson)**

Academic libraries are rapidly expanding electronic communications with remote clientele, including suggestions, ideas/comments, and reference & research assistance. Most interfaces to services of this sort (including the University

Libraries' "Talk to the Library") encourage the user first to consult a FAQ or "previously asked questions" area to minimize duplication of questions and answers. These FAQs often are quite labor-intensive to construct and to maintain, however, especially when the in-coming questions arrive at a high rate vis-a-vis persons available to generate responses. The purpose of this project is to develop an automated tool to manage these inquiries.

### **7. Hancock Library Photographic Archives (Jean Crampon)**

The Hancock Library of Biology and Oceanography includes an archival collection of photographs, slides, and films. Material from this collection is used by both scholars on campus and worldwide. Most of the Hancock materials are still under copyright, but the copyright is owned by USC in most cases. The purpose of this project is to provide access to these materials via the Web, to provide access to medium-quality images for students, and to offer a mechanism to order high quality images for a fee.

### **8. Hancock Museum Virtual Tour (Jean E. Crampon)**

One of the hidden treasures at USC is the Hancock Memorial Museum. One way of using this Museum for educational purposes is to provide a virtual tour of the Museum or parts of its collection. Although other museums and libraries have mounted virtual tours, it would be useful to have a template to prepare and edit a virtual tour. The purpose of this project is to develop a virtual tour generator for the museum which might be adaptable for other locations as well.

### **9. Interlibrary Loan Data & Interface (Joyce S. Toscan)**

In August 1997, Web pages were created to allow USC students, staff, and faculty to request interlibrary loans electronically. There are three forms: one for books, one for dissertations, and one for photocopy requests. Although the data is collected via the Web, traditional paper-based techniques are still used to process the requests. The purpose of this project is to develop a system which will handle the data after it has been received for processing. The system should have good error/data checking; provide the capability to generate statistical reports daily, monthly, quarterly, and annually; and process requests through the standard interlibrary loan utilities.

### **10. International (French) Cross-Cultural Teaching Model (Danielle Mihram)**

USC offers courses on French culture which could benefit from a Web-based system which includes multi-media information about contemporary France. The purpose of this project is to build model which would: 1) link resources between the USC campus and USC courses in Paris, 2) include archives of materials such as video clips, film shoots, exhibits, dictionaries, etc., 3) include a collaborative component allowing instructors to input material either from Los Angeles or from Paris, 4) provide machine translation, 5) provide real time communication between individuals in France and a class in USC Los Angeles, and 6) be portable to other courses and languages.

### **11. ISLA Forms Development for Data Ingest (Li Hunt)**

ISLA has a need for an easy-to-use form generation process. The purpose of this project is to develop a tool for creating a form with different fields based upon the

collection which is being readied for input. This form should be easy for the "owner" of the collection to tailor the form to the specific characteristics of the collection.

### **12. ITV Course Materials (Julie Kwan)**

Many engineering courses are videotaped by USC's Instructional Television Network (ITV); these courses are attended by students both on-site and off-site. There are two types of products: videotapes of class lectures and handouts. Videotapes are frequently used by students to review course material. Handouts are produced in high volume and delivered to off-site students by courier. If materials were available electronically, students could more easily review course lectures and retrieve course handouts regardless of their physical location.

### **13. Korean-American Museum (Ken Klein)**

Resources for Korean-American studies are just now being seriously collected by various institutions in Southern California, namely, the Korean American Museum, Radio Korea, UCLA, and USC's Korean Heritage Library. The materials are in varied formats: published books and magazine articles, manuscripts and documents, photographs, audiotapes, videotapes, and museum pieces of all types. What is needed is a way to bring all of these materials together in one searchable catalog so that they may be retrieved by the different institutions and accessed by the broader research and museum community.

### **14. Latin American Pamphlets (Barbara Robinson)**

The Boeckmann Center collection includes about 5,000 pamphlets from the 19th and early 20th centuries. These pamphlets are fragile and in need of conservation. They have significant historic value since they were a popular form of publishing in Latin America at the time. The purpose of this project is to convert these materials to digital format and to provide full-text search capability. Although this collection deals with Latin America, the project is relevant to fragile materials published in any language.

### **15. Lion Feuchtwanger Archive (Marje Schuetze-Coburn)**

The Feuchtwanger Memorial Library in the Special Collections Department contains several hours of audio recordings (reel-to-reel and cassette copies) of the German writer Lion Feuchtwanger reading from his works during the 1950s. The purpose of this project is to digitize these audiorecordings, make them available on the Feuchtwanger website <http://www-lib.usc.edu/Info/FML>, and provide the text of Feuchtwanger's readings (both in the original German and English translations) on the screen simultaneously. Since the quality of the recordings is sometime poor, this project might also need to consider ways to digitally enhance the sound.

### **16. Los Angeles Regional History Photographic Archives (Dace Taube)**

Photographic archives are increasingly being digitized but to efficiently access them requires that each image be linked to a searchable text record, the creation of which is an intellectually demanding process. In addition to photographic archives there are other types of digital images to be accessed: art collections, graphics, and even "scrapbook" photographs on individuals' home pages. The problem: to devise a method of accessing digital images which is not text-based.

### **17. Maps for Instruction (Julie Kwan)**

Maps are important resources for many undergraduate courses from history in the humanities to geology in the sciences. Students need access to maps for instruction not only in the classroom but outside as well when they are studying. The purpose of this project is to provide digital access to maps as part of any collection of electronic reserves. The digital presentation should allow independent viewing of the map legend as well.

### **18. Medieval Manuscripts (Ruth Wallach)**

Medieval manuscripts are important to historians, not only for their content, but for their physical structure (binding, drawings, markings, hand-written notes in page margins, etc.) They are unique items with very few copies of any particular manuscript available throughout the world. The purpose of this project is to represent manuscripts digitally, not only their content, but their format and binding, so that they can be accessible to scholars both locally and worldwide.

### **19. Networked Consultation Support (Judy Truelson)**

USC's libraries are moving to a 24/7 service delivery model with increasing usage of services and collections from remote locations. Students working at home have difficulties calling for consultation if they have only one telephone line which they use for both voice and data communications. The purpose of this project is to develop a model for how reference and research consultation might be provided to remote users from librarians and other staff who are staffing service points or are in their offices. The system should allow for communication between the consultant and the student as well as provide access to the information systems which will resolve the student's information problems.

### **20. Online Catalog Search Strategies (Phil Reese)**

Online library catalogs typically provide very structured approaches to searching for materials of interest. These search systems are considerably different from the search engines which are used on the Web. The purpose of this project is to develop a model for searching the online catalog using search engine types of searches but which would display the results sorted by category, author, subject, keywords, title, etc. Would this be useful to the campus users or does the structure imposed by conventional online catalog search pay better returns?

### **21. Planning Documents (Robert Labaree)**

The VKC Library collection includes a number of planning documents which include city and neighborhood plans, municipal codes, environmental impact documents, and planning documents related to specific projects, such as airports, highway construction, and greenways. The content of the documents varies tremendously, but are generally textual in nature with detailed maps and sometimes statistical charts imbedded throughout. These materials are heavily collected at universities with urban planning schools, city and state government libraries, and some federal agency libraries, such as the Department of Transportation.

### **22. Serial Control Records for the Integrated Library System (Chris Sterback)**

The USC libraries are just beginning to implement true serials control in Sirsi's

Unicorn Collection Management System. To do so, library staff need to create serials control records for the over 15,000 journals currently received. To date, most other libraries have created these records manually; manual record creation however would take years to complete. The purpose of this project is to develop a program which would provide for the automatic creation of at least a skeletal record which might need some modification, but which would considerably speed the creation of these records.

### **23. Statistical Charts (Marianne Afifi, Royd Muroaka)**

USCInfo data is currently collected automatically and stored electronically in tables on a monthly basis. The purpose of this project is to develop a tool to take this data from tabular format and provide the capability to: 1) select particular types of data, and 2) view the results graphically. This would give library staff the capability to analyze data comparatively and to look at differences in numbers during different times of the academic year, etc. Ultimately, such a tool could be applied to any type of tabular data available on the USCWeb and might be very helpful for analyzing data relating to other issues.

### **24. Stereoscopic Slides (John Ahouse)**

The Special Collections Department owns a collection of stereoscopic slides from the early years of this century. The slides, which were viewed through a stereoscopic viewer, depict a variety of scenic images from around the world. A popular form of entertainment at the time, they continue to have research value for their image content and quality. The purpose of this project is to make these images available through the Web and to make them viewable stereoscopically.

### **25. Technical Reports Archive (Charles Phelps)**

Engineering faculty usually first publish research results as a Technical Report; they may subsequently publish the entire report or a portion thereof as a journal article. Traditionally, Schools of Engineering and their associated departments sent out lists of newly published technical reports and supplied a copy of a complete report to anyone who requested one. The purpose of this project is to develop a database of Technical Reports including a mechanism to provide full-text reports to the reader, provide email updates about newly issued reports, and to link with the national database of computer science technical reports, [NCSTRL](#).

### **26. TV Show Files (Don Thompson)**

The Cinema/Television Library owns the Academy of Television Arts and Sciences Library. Included in this collection are several file cabinets of press releases called "Show Files". The press releases cover the period from the mid-70's to the early 90's, and list the cast and crew for such shows as MASH, Cheers, Hill Street Blues, Family Ties, and Rosanne. The purpose of this project is to digitize these files and to develop a software system that includes a database of digitized images without manual keying the information. The system should be searchable full-text.

### **27. Virtual Business Reference Assistant (Deborah Holmes-Wong)**

Libraries supporting professional schools are facing challenges of providing services to students who are located at a distance from the library with an ever-

shrinking staff. The purpose of this project is to develop a system to help students answer routine and repeated questions through the use of discipline specific knowledge-bases and intelligent agents. The system should track questions asked and responses given so that staff and end-users can evaluate accuracy of the responses and use them to improve the system. A video-conference link or other service may be offered during specific times of the day as a back-up, but this service would be the front end and link to many of Marshall's online subscription services.

**28. Virtual Education Reference Assistant (Linda Weber)**

The School of Education has a large number of off-campus students who are in nine instructional centers around California, Hawaii, and Lincoln, Nebraska. The purpose of this project is to find new ways to use information technology to help students get answers to their questions without requiring a librarian. This model might involve intelligent agents to search designated sites, especially the U.S. and California Departments of Education sites which include full-text data and statistics. The system should track questions and answers so staff can evaluate and improve answers. If possible, the system should screen and automatically add new education Web sites to the answer database.

### Domain Models: Risk and Expectations Management

Type of Application	Simple Block Diagram	Examples	Simplifiers	Complicators
<p style="text-align: center;">Multimedia Archive</p>	<pre> graph TD     Q1[Query] --&gt; C[Catalog]     U1[Update] --&gt; MA[MM Archive]     C --&gt; MA[MM asset info]     MA --&gt; C[Update Notification]         </pre>	<p>1, 2, 3, 4, 7, 12, 13, 14, 15, 16, 17, 18, 21, 24, 25, 26</p>	<ul style="list-style-type: none"> <li>• Use standard query languages</li> <li>• Use standard or COTS search engine</li> <li>• Uniform media formats</li> </ul>	<ul style="list-style-type: none"> <li>• Natural language processing</li> <li>• Automated cataloging or indexing</li> <li>• Digitizing large archives</li> <li>• Digitizing complex or fragile artifacts</li> <li>• Automated annotation/description/ or meanings to digital assets</li> <li>• Integration of legacy systems</li> </ul>
<p style="text-align: center;">Selective Dissemination of Information</p>	<pre> graph TD     NA[New assets] --&gt; IB[Info Base]     IB --&gt; F[Filter][New asset notification]     F --&gt; OI[New assets of interest]         </pre>	<p>New library material notification</p>	<ul style="list-style-type: none"> <li>• Use of existing or standard information base</li> <li>• Well defined distribution points</li> <li>• COTS notification and event processing</li> <li>• WWW/internet based</li> <li>• Restricted interests vocabulary and filtering structures</li> <li>• Single information base</li> </ul>	<ul style="list-style-type: none"> <li>• Volatile or ill-defined interest or filtering criteria</li> <li>• Complex distribution</li> <li>• Multiple distribution formats</li> <li>• Heterogeneous information sources</li> <li>• Complex filter reasoning</li> <li>• Automated interest update</li> </ul>
<p style="text-align: center;">Data Analysis</p>	<pre> graph TD     U[Updates] --&gt; DS[Data Sources]     DS --&gt; PS[Parser / Searcher][Queries]     PS --&gt; DS[Relevant data]     PS --&gt; PDP[Pattern Detection/Processing][Parse Format]     PDP --&gt; R[Reporting][Analysis display]         </pre>	<p>11, 23, USC web-page stats</p>	<ul style="list-style-type: none"> <li>• Use of data analysis packages (statistics, etc.)</li> <li>• Implement using interpreted or script languages (e.g. PERL)</li> <li>• Data stored in an RDBMS</li> <li>• COTS reporting packages (for graphics, etc.)</li> <li>• Simple, consistent data formats</li> </ul>	<ul style="list-style-type: none"> <li>• Natural language processing</li> <li>• Highly unstructured data</li> <li>• High degree of formatting or conversion</li> <li>• Computationally intensive reporting</li> <li>• Spatial data analysis</li> <li>• Complex pattern recognition</li> </ul>
<p style="text-align: center;">Activity Monitoring and Control</p>	<pre> graph TD     A[Activity] --&gt; AA[Activity Agent]     C[Controls] --&gt; AA     AA --&gt; Act[Action]     AA --&gt; SA[Status Assessment][Event]     SA --&gt; MC[Monitoring &amp; Control][Status Reporting]     MC --&gt; SA[Policy]     SA --&gt; AA[Status]     SA --&gt; AA[Revised controls]         </pre>	<p>Room scheduling, Reserve materials management</p>	<ul style="list-style-type: none"> <li>• Standards based agent interfaces</li> <li>• Simple, well-defined policies</li> <li>• Uncoupled controls</li> </ul>	<ul style="list-style-type: none"> <li>• Real time or embedded monitoring</li> <li>• Synchronization of monitoring &amp; control activities</li> <li>• Concurrency management of activities</li> <li>• Distributed monitoring of activities</li> <li>• Natural language processing of policy</li> <li>• Policy learning</li> </ul>

**Domain Models: Risk and Expectations Management (continued)**

Type of Application	Simple Block Diagram	Examples	Simplifiers	Complicators
Reference Services and FAQ		6, 27, 28	<ul style="list-style-type: none"> <li>• COTS repository</li> <li>• WWW interface</li> <li>• Restricted query vocabulary and format</li> </ul>	<ul style="list-style-type: none"> <li>• Natural language processing of query</li> <li>• Multiple domains</li> </ul>
Data Migration		5, 11, 22	<ul style="list-style-type: none"> <li>• Single data source</li> </ul>	<ul style="list-style-type: none"> <li>• Heterogeneous or distributed data sources</li> <li>• Context sensitive conversions</li> <li>• Large number of exceptions</li> <li>• Highly relational data</li> <li>• Complex computational dependencies</li> </ul>
Special Collection Access		7	<ul style="list-style-type: none"> <li>• Fixed virtual architecture</li> </ul>	<ul style="list-style-type: none"> <li>• Non-homogenous physical architecture</li> <li>• Immersive virtual reality (VR)</li> </ul>
COTS Package Extension		5, 20, 23	<ul style="list-style-type: none"> <li>• Clean, well-defined API's</li> <li>• Single COTS package</li> <li>• Simple mappings of interface inputs and outputs</li> </ul>	<ul style="list-style-type: none"> <li>• Natural language processing</li> </ul>

**Domain Models: Risk and Expectations Management (continued)**

Type of Application	Simple Block Diagram	Examples	Simplifiers	Complicators
Distributed Borrowing		9, 19	<ul style="list-style-type: none"> <li>• Homogeneous asset sources</li> <li>• Simple asset source interfaces</li> <li>• Few asset sources</li> </ul>	<ul style="list-style-type: none"> <li>• Complex borrowing policies and requests</li> <li>• Organization politics and economics</li> </ul>
Interactive Communication		10, 19	<ul style="list-style-type: none"> <li>• Internet/WWW interface</li> <li>• COTS communication services</li> </ul>	<ul style="list-style-type: none"> <li>• High degree of integration</li> <li>• Real-time</li> <li>• Synchronous</li> <li>• Concurrency</li> <li>• Rich media (video, voice, NLP, etc.)</li> </ul>