UNIVERSITY OF CALIFORNIA

SANTA BARBARA - Library

MAP AND IMAGERY LABORATORY-LIBRARY

Project Alexandria

(NSF-ARPA-NASA Digital Library Initiative 1994-1997)

Project Alexandria will develop a digital library providing easy access to large and diverse collections of maps, images, and pictorial materials. The main output of the Project will be a distributed database that permits users from across the country to view and retrieve items that were previously difficult or impossible to access. Although the actual collection of items that are accessible by users will be stored at libraries dispersed throughout the USA, it will appear to users as if the materials were held at a single, local library site.

The library will also provide a full range of electronic library services, such as an electronic reference desk, that will help users to take full advantage of the library holdings. From a librarian's point of view, items that were once viewed as cumbersome and perishable will now be easily managed. A long-term goal of Project Alexandria is to integrate the access that Alexandria will provide to these materials with access both to more traditional text materials and to other multi-media materials.

Users ranging from school children to academic researchers to members of the public will be able to view and retrieve materials from this library by electronic means. In particular, they will be able to search for maps and images on the basis of the information contained in them, as well as the places they refer to. One may, for example, imagine a school-child who accesses the library through its user-friendly interface and finds a map of Amelia Earhart's last flight, so that she can print it out as part of a project report; or a business person who obtains a map of parking availability in a certain area of town; or a scientist who retrieves satellite images that contain pictures of recent hurricanes coming ashore over Texas.

The Project involves a consortium of university, library and industrial partners and is centered at the University of California, Santa Barbara (UCSB). The UCSB component of the team involves such groups as the UCSB Library's Map and Imagery Laboratory, Department of Computer Science, the Department of Electrical and Computer Engineering, the Center for Remote Sensing and Environmental Optics and the National Center for Geographic Information and Analysis. These groups together represent a major body of expertise concerning

spatially-indexed information. Other university partners are the State University of New York at Buffalo (SUNY) and the University of Maine at Orono. Library partners include the Library of Congress, the University of California Division of Library Automation, the library of SUNY at Buffalo, the library of the US Geological Survey, and the St. Louis Public Library. Industrial partners include DEC, ESRI, Conquest and XEROX.

The project involves two distinct phases. During the first six months of the Project, the Alexandria team will build and place on-line a prototype system using software developed for geographical information systems by ESRI. This prototype will enable many sorts of users to access collections of digitized maps, images, and airphotos relating to Santa Barbara, Ventura and Los Angeles Counties. It will also enable the Alexandria team to fine-tune its design for the graphical user-interfaces that will facilitate library access. During this same period, the team will develop detailed specifications for the main testbed system. During the remaining three and one half years of the Project, the main testbed will be developed and tested. The testbed will take the form of a distributed digital library with components at UCSB, SUNY (Buffalo), Library of Congress, USGS, and St. Louis Public Library, as well as at other interested libraries. This system will allow users from around the country to access the collections of digitized maps, images, and pictures as if they were held locally rather than at widely-dispersed library sites.

The facilities at any of these library sites will include various combinations of four basic library components. These components include a user interface supporting simple access to each of the library services by means of textual and visual query languages, as well as electronic browse capabilities; a catalogue component providing rapid and appropriate response to user queries, particularly those involving content-based search; an ingest component that permits librarians and systems managers to incorporate new items into the library collection, using procedures that include digitization, reformatting and the extraction of catalogue information; and a storage component providing storage capability for, and high-speed access to, large collections of spatially-indexed items. The particular set of components at any library site will depend on the particular needs of the site and the Library as a whole.

From the users point of view, the graphical user interface is the most crucial of the four components of the Alexandria Library. The diversity of users of the Library requires that interfaces be simple to use and intuitive for accessing and retrieving spatial and non-spatial information. The functionality of the interface component involves support for text-based and visually-based query languages that permit users to express, in simple and frequently visual terms, queries concerning the availability, characteristics, and content of information that satisfies their requirements. The interface will support the visualization and browsing of information in which users express an interest, and will have information sets transmitted to users upon request.

The core of the Library is an electronic catalogue system that allows the system to search for items that are requested by users. The catalogue contains its own database of information about the library collections and efficient mechanisms for searching this database. Some of this metadata about the library collections describes both the contents and the locations of the various maps, images and pictures in the Library while another component of the metadata consists of reduced versions of images and maps that may be browsed quickly by users.

The ingest component allows libraries to extend their collections of holdings that may be accessed by Alexandria by digitizing maps, images, photographs and other graphical materials. Such ingest involves not only the use of high-performance scanners but also the extraction of metadata for the catalogue and the formatting of the data so that it may be stored and accessed in an efficient manner. In particular, the Project will be using a "hierarchical" decomposition of maps and images based on a technique known as "wavelet transformations". After making this transformation, an image is represented as a set of images at different levels of resolution in such a way that the different sub-images, when "added up", give the original image. The sub-images can be used for browsing and for extracting information about the contents of the image.

Apart from the design, development, and integration of the four basic components, a variety of research and development issues will be addressed and resolved during the four years of the Project. A first issue involves the design for the graphical user interfaces, which will have to be thoroughly evaluated in terms of whether they meets the full range of user requirements and whether they are easy to use. The Project has many mechanisms for ensuring that inputs from many different user groups will be incorporated into the final system. A second issue concerns network aspects of the transmission of information between the library sites and the users. The use of the Internet will be very important in this regard.

A third issue is that the system must be compatible with various standards that are used by libraries and other organizations in describing and transmitting the information of digitized maps and images.

A final and very important issue for Alexandria is whether it "scup" as the collection of library sites and the size of its digitized collection of items increases. In other words, will the library continue to work effectively from a users point of view as it grows very large. The Alexandria Project will use various mechanisms to ensure that the system will scale up. One such mechanism, for example, is to have much of the computational processing for the catalog and ingest, storage and ingest components performed in parallel. Hence the Project team will be exploring ways to make many workstations perform their computations in parallel. They will also be investigating other mechanisms related to the hierarchical decomposition of information into wavelets and the use of appropriate "models" of the information.

For More Information Contact:

Terry Smith

Computer Science

smithtr@cs.ucsb.edu

805-893-2966

Mike Goodchild

NCGIA

good@ncgia.ucsb.edu

805-893-8049

Joe Boisse

Univ. Librarian

boisse@library.ucsb.edu

805-893-3256

Larry Carver

Library, MIL

carver@sdc.ucsb.edu

805-893-4049

Mary Larsgaard

Library, MIL

mary@sdc.ucsb.edu

805-893-4049

Project Alexandria Home Page