A Digital Library in a Virtual Enterprise

Lizanne Payne, Executive Director
Washington Research Library Consortium
901 Commerce Drive
Upper Marlboro, MD 20772
301-390-2031
eap@wrlc.org

This fall (1996), WRLC added locally-created digital collections to its shared system known as ALADIN (Access to Library and Database Information Network), joining the union catalog of seven libraries' holdings plus locally-mounted article citation databases. WRLC staff designed a Web-based user interface which groups these disparate online resources supported by many separate systems into a coherent digital library shared by the member universities (see http://www.aladin.wrlc.org). Simultaneously we began to deploy "electronic library workstations" in the member libraries, preloaded with client software to present a centrally-manageable, unified user interface.

In designing a shared Web-based digital library, we have had to determine which technologies are appropriate for various kinds of locally-created digital collections, and where various electronic resources should fit in the structure offered to searchers.

Digital Library Technologies

Terminology: Digital library collections are electronic representations of materials which are linked by some organizing principle, where individual items (digital objects) must be retrievable independently based on a variety of attributes. Preparing a digital collection, like a physical one, requires grouping and adding descriptive text both to the collection and to individual items to facilitate future unknown retrieval paths.

Digital collections consist of metadata, secondary data such as finding aids, and representations of primary data. These different data types reside at different levels in the system hierarchy and require different storage and retrieval mechanisms. Earlier generations of information retrieval systems such as online catalogs were designed to handle text-based metadata. The supreme value of new technologies including the World Wide Web lies in the ability to store, retrieve, and display primary data as well.

Design principle: Digital collections should be accessible through the same retrieval mechanisms used for physical collections and, in addition, should be directly-addressable as Web resources.

Thus ideally, searching ALADIN for information on Robert E. Lee would retrieve all locally-available information including monographs, special collections finding aids, and local digital images, plus article references and cataloged Internet resources.

Data taxonomy: A photograph of Robert E. Lee owned by a WRLC member library could be represented through the following electronic records, in order from the most primary to the most secondary:

- Stored digital image (digital object)
- Object descriptive record
- Digital collection record
- Finding aid or research guide (secondary data)
- Finding aid descriptive record

Structure and Contents of the ALADIN Digital Collections

The plan for WRLC's ALADIN system calls for these records be stored in the following databases:

1) Object descriptive record and stored digital image (digital object)
OCLC SiteSearch system (Image Support, Z39.50 server, WebZ Z39.50 to HTTP conversion). SiteSearch provides database maintenance features and Z39.50 compatibility. Object records are searchable through a Web browser

through WebZ. Objects and object records are grouped into separate databases or collections. Objects and collections can have independent URLs for direct Web access.

2) Digital collection record

MARC record in the online catalog containing the collection URL in the 856 tag. Searchable through any Z39.50 client or Web browser through WebZ.

3) Finding aid or research guide (secondary data)

HTML or SGML files with links to URLs for the collection and individual objects where identified. The finding aid itself could have a URL.

4) Finding aid descriptive record

MARC record in the online catalog pointing to a printed finding aid or online finding aid with URL, if available.

The goal of this multi-layered approach is to integrate these different record types across different systems and databases so that the searcher can reach the appropriate digital object regardless of the starting point of the search and without needing to know in advance that the digital object exists.

Lizanne Payne
Executive Director
Washington Research Library Consortium