The NEEDS Virtual Library System

David Martin, Iowa State University

John M. Saylor, Cornell University

Bill Wood, University of California at Berkeley

The National Science Foundation has funded a number of Engineering Education Coalitions whose broad mission is to reshape engineering education. About sixty-six institutions are now affiliated through eight coalitions at a total cost in excess of $24 million per year. The aggregate output of these coalitions will have substantial long term effects on the nature and content of engineering education and practice. The work described here is intended to be a principal means of rapid, international dissemination of the many diverse technical and cultural innovations developed by the engineering education community in general and the Engineering Coalitions in particular.

We are affiliated with one of the first coalitions formed: Synthesis, which is comprised of the Universities of California at Berkeley, Cal Poly at San Luis Obispo, Cornell, Hampton, Iowa State, Southern, and Tuskeege. One of Synthesis' initial concepts was the establishment of a National Engineering Education Delivery System, NEEDS, which would be an open, multi-dimensional database of multimedia courseware modules and related materials to support engineering education. Arv Eide described "offices of the future" in which students and faculty would have on-line access to a wealth of courseware and materials to help them understand their disciplines. Faculty would be able to easily create effective multimedia presentations by "clicking and dragging." NEEDS was designed to enable this vision.

A fundamental strategy was to build our system on existing and emerging standards to ensure scalability and long term viability. We convened a expert panel of academicians, publishers, computer sages, content specialists, copyright specialists, etc., to help us decide what basic models to use and on which standards to build. After about a half-dozen meetings over a two year period, this group published version one of the NEEDS Standards Study Report. Among other things it recommended that the NEEDS database system should:

- appear to users as a unified library of courseware modules and data elements for building courseware even though it is actually a highly distributed system;
- follow the Library model of a structured, searchable catalog with links to arbitrary objects;
- be available to users with minimal hardware, software, and expertise;

- be built on a base of communication and networking standards such as TCP/IP used over the Internet; and,
- conform to Library cataloging standards and practices from the outset.

We decided to integrate with an existing university on-line catalog or catalogs in order to take advantage of the library's fundamental mission and their wealth of existing infomation science knowledge. This enabled us to:

- make use of the robust, powerful search engines associated with on-line catalogs;
- use library on-line display engines to ease the burden of training users;
- gain access to cataloging expertise;
- ensure long term viability by integration with an institution committed to information service;
- take full advantage of library standards and practice by employing trained librarians.

A prototype NEEDS system was established and tested. This system used a central server for client access and about a dozen database archive servers distributed across the country. The design of the server system was based on the principles that the functions of the (virtual) NEEDS database server should include:

- providing a transparent connection between users and on-line catalogs as well as to other data sources;
- initiating the transfer of data files from owners to users, and providing users with special access information at that time (i.e. page charges, etc.);
- keeping simple statistics including the number of times the transfer of a particular deliverable is initiated but excluding any personal information such as user names and passwords;
- the NEEDS database servers may also assist users in locating developmental multimedia databases of courseware and data elements;
- managers of the various computers serving as data storage devices are responsible for controlling access to those devices. (Initially access was controlled through the standard Telnet protocol);
- managers of the various computers to which users direct retrieved files are responsible for controlling access to those computers. (Initially access to those computers was through the standard FTP protocol.)

A key philosophical element of the system is that while users percieve that they are interacting directly with the NEEDS system, transactions actually are between the client and the owner of the data. NEEDS is simply a powerful agent which brings the information user and supplier together in a highly transparent manner; NEEDS is not responsible for enforcing copyright, use, or access restrictions.

As soon as World Wide Web (WWW) browsers were released we began to redesign our system to make use of that technology. In retrospect, this was a wise decision, since the web continues to expand rapidly. One of our implicit strategic positions has been to use only widely available client software; we see no way to reach a national or international clientele if we must distribute and maintain client software. We still maintain Telnet access to our system through a Lynx/Kermit gateway to our WWW servers, and will continue to support the most popular, standard communication software.

We have implemented completely standards based schemes for all aspects of our delivery system: entering cataloging information, search & retreival by clients and staff, user interfaces, and all data communications. The system capitalizes on the following standards: USMARC, SGML, HTML, HTTP, Z39.50, and WAIS. Z39.50 and WAIS are used to query databases, the records of which are described using USMARC cataloging rules based on on-line WWW forms filled out by authors.

Paper is unnecessary. SGML editors are used to create and maintain the database records using a DTD written for MARC. The same editors are used to create HTML documents. HTTP and FTP servers are used for access to archived data, as well as for maintenance and user interfaces to a variety of gateways. The catalog records searched contain Universal Resource Identifier information (in MARC tag 856:$u). Actual storage location URLs are substituted by an HTTP based "location" server when a user begins to transfer information s/he has located through our search interface. The system will be demonstrated and discussed in detail.