

Metasearch Infrastructure Project: Recent Achievements & Current Status

Roy Tennant, roy.tennant@ucop.edu • March 1, 2006

The Metasearch Infrastructure Project seeks to develop a robust set of tools for crafting tailored search interfaces to diverse information resources for specific audiences and/or purposes. For additional background information, see the project web site at <<http://www.cdlib.org/inside/projects/metasearch/>>. Roy Tennant is the Project Manager and Mike McKenna, michael.mckenna@ucop.edu, is the Technical Lead.

For appropriate background to this update, see the previous updates available from <<http://www.cdlib.org/inside/projects/metasearch/overview.html>>.

CDL Developments

Metasearch Software

Mike McKenna has made significant progress with integrating MetaLib with the CDL Common Framework (see <http://www.cdlib.org/inside/projects/common_framework/> for more information), and was recently joined by Erik Hetzner, who is focusing on the client code.

Concurrent with this work has been work by Margaret Low and Steve Toub to generalize the Common User Interface (CUI) component of the Common Framework so it serves the needs of the metasearch infrastructure as well as the Digital Preservation Repository and other services that will rely on the Common Framework. Once their work is done Roy Tennant will be working with Steve, Erik Hetzner, Heather Christenson, and Mike McKenna to build the search portal required by the National Science Digital Library grant.

Preliminary work was done to develop an interface component that allows librarians to automatically load an RSS feed into the entry page of a search portal. Also, work has been progressing on a batch OpenURL lookup before displaying search results, so we can display direct links to full-text when available. David Loy, Margery Tibbetts, and Mike McKenna are involved.

Metadata Harvesting

David Loy finished work to expose harvested metadata records to searching via the SRU protocol (see <<http://www.loc.gov/standards/sru/>>), which will enable integration of harvested metadata along with, or beside, licensed database results. Additionally, a utility that was created in association with our harvesting work to normalize dates in harvested metadata was released to the library community (see <<http://www.cdlib.org/inside/diglib/datenorm/>>).

Timeline

Our goal is to have a prototype search portal focused on earth sciences available in May for usability testing with undergraduates. This is in partial fulfillment of our NSDL grant, and will serve to provide us with important early feedback on what should be changed before getting campuses more closely involved. During the Spring we will also be refactoring the

back-end infrastructure, with the goal of making it as easy for campuses to deploy portals as we can.

Campus Developments

Campus staff (Elaine Adams from UCLA, Christy Hightower and Ann Hubble from UCSC, Memorie Yasuda from UCSD, and Mitchell Brown from UCI) continued to advise us on the prototype Earth Sciences NSDL portal. The UCLA European Studies portal team (Terry Ryan, Cindy Shelton, Sarah Watstein, Kati Radics, David Yamamoto, Carlo Medina, Luiz Mendes, Darrow Cole, and Bo-Gay Tong Salvador) was requested to wait until we have gained experience with our first portal and have refactored our infrastructure. Present plans are to make our nascent infrastructure available to the UCLA team this summer for setup and testing.

Campus staff will also be helping us with the usability testing prototype(s) in May. Felicia Poe and Jane Lee will be coordinating with campus staff to find appropriate facilities, find students to participate, etc.

The “SmartStart” portal team, which includes members from UCSC (Christy Hightower and Catherine Soehner) and UCLA (Diane Mizrachi and Pauline Swartz), will also likely be reconstituted to plan a portal aimed at undergraduates and/or those who want “a few good things.”

Outside Developments

Roy Tennant spoke at the NISO Metasearch Workshop in September 2005, and chaired the ARL MetaLib Portal Implementor’s Group (ARMPIG) Meeting in November 2005. Other presentations on CDL metasearch work included talks at Access 2005 in Edmonton, LITA in San Jose (with Mike McKenna), and the Charleston Conference in Charleston, SC.

Mike and Roy continue to work with other MetaLib X-Server users and ExLibris staff to identify, prioritize, and implement enhancements to the X-Server infrastructure.