Journal and Article Locator (JAL): Federated Access to Electronic/Print Journals and Article Full-Text

Michael A. Norman
Mary C. Schlembach
Wendy Allen Shelburne
William H. Mischo

Introduction

Over the last two years at the Charleston Conference we have given detailed presentations regarding the University of Illinois at Urbana-Champaign (UIUC) Library's efforts to offer comprehensive access to its electronic resources collections, including e-journals, e-books, A&I services, reference tools, primary source collections, government publications, etc. We quickly determined that one pathway was not the best solution to provide successful access for our users and our overarching goal is to provide as many avenues as possible for users to discover the Library's electronic and digital resources.

Users employ various search methods to find what they need for their research – be it searching the online catalog, the Library's home-grown electronic resource management system called the Online Research Resources (ORR), using the Discover (SFX) service to follow OpenURL links in A&I databases, or our custom metasearch system (Search Assistant), which allows users to do federated searching across selective of journal article databases, our online catalog, digital asset management systems, and the state-wide consortium universal catalog.

This paper presents an overarching view of our e-resource management activities and specifically describes the UIUC Library Journal and Article Locator (JAL) system. The JAL pulls together a number of disparate search environments into one unified display without having to search each one independently of the others. With JAL, the Library provides one-stop access to both print and electronic journals at the title and article level. Resource integration has become the key component to providing an all-encompassing search experience, giving users access to the full range of the Library's collections. This year, our program will give a detailed view of how Journal and Article Locator integrates various resources into one search environment and outline what is involved with setting up this system.

User Behaviors

Use studies show that library users—in particular undergraduate students—greatly prefer one-stop access to information resources. This is why Web search engines are the most common starting point for undergraduate student searching (De Rosa, 2005; LibQual, 2004). Students want to use search systems that provide resource integration and return results with links to relevant and reliable information resources.

It can not be understated that for users, links to full-text content are paramount to their research. Until the late 1990s, the online catalog and/or alphabetical listings of journal listings were the first stops for users looking to determine what serials and monographs the UIUC Library had in the various research disciplines. Today, the online catalog has been marginalized due to lack of or continual maintenance of full-text links, the inability to access journal literature (at the article-level), and the rise of full-text monographic delivery systems.

Because of the difficulty of topical searches in full-text systems (page rank works for Web pages, but not full-text books), it is now even more important to provide users with quality metadata, comprehensive knowledgebases, and good access mechanisms. To provide a successful search and discovery environment, the ability to combine full-text and traditional access to print materials is critical to allow researchers the ability to discover the full range of available literature.

Different Discovery Mechanisms for E-Resources at UIUC Library

Presently at the UIUC Library our different discovery mechanisms for electronic resources embody eight different systems:
- the OPAC (Voyager) with links to both e-journals and e-books
- the Library's online research resources (ORR) – a local electronic resources management (ERM) database which has an e-resources directory combined with a search engine, A to Z title list, and subject classification
- the Discover SFX service
- the Search Assistant custom federated search system
- WorldCat including WorldCat, Open WorldCat, Worldcat.org – all with title level access
- Google Scholar – full-text articles and books
- Grainger Portal: resource integration using metasearch: topical search, specific journal title, specific article in journal
- Journal and Article Locator (JAL) – journal search and discovery module and part of Grainger Portal

The Voyager OPAC has links embedded into MARC bibliographic records for both e-journals and e-books. There are currently over 300,000 full-text links populating the online catalog (See Figure 1). These links are maintained on a semi-regular basis. A computer program using PERL scripting and the cataloging software MARCEdit has been created to continually upgrade and overlay bibliographic records with the most recent changes. The e-journal records are embedded with URLs that link into the Library's electronic resources management (ORR) to alleviate duplication of maintenance. The various feeds updating the ORR help keep the links in the OPAC as current as possible. This was one of the Library's first applications to successfully create integration of two discrete systems, allowing users to link back and forth between the OPAC and ORR.
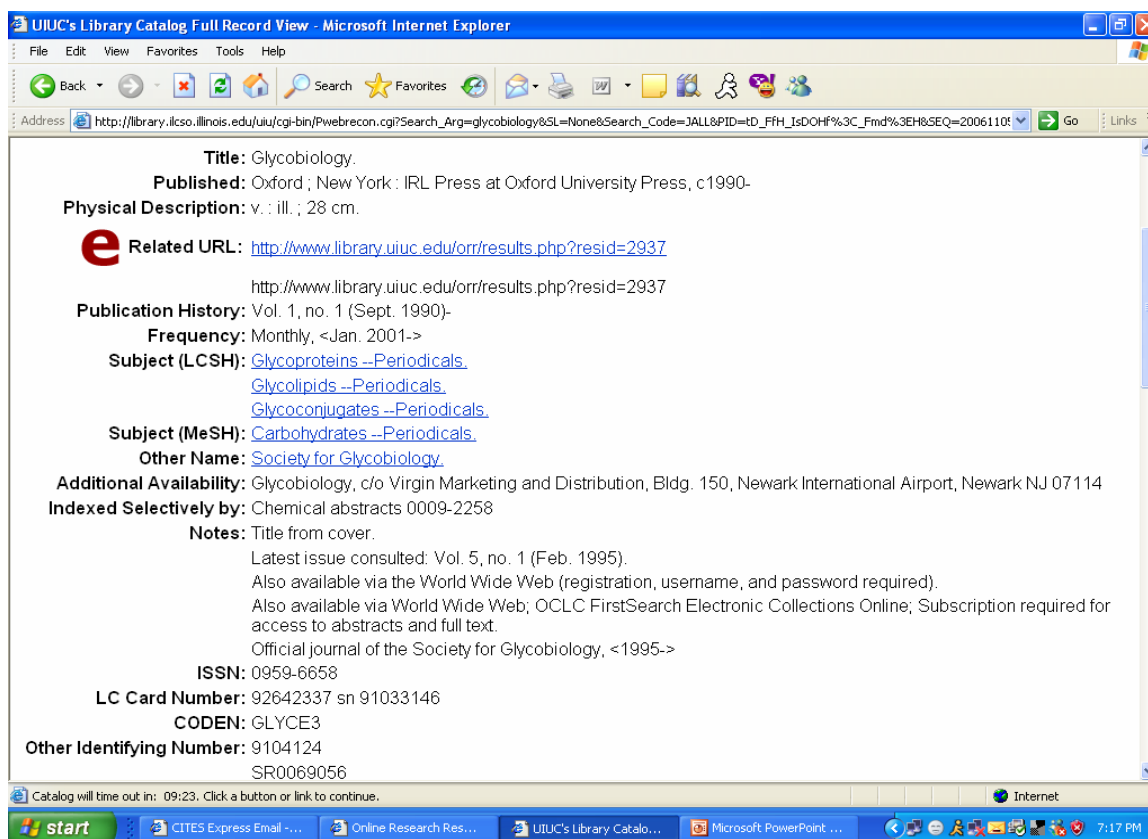
Figure 1 – Online Catalog (Voyager)

The Online Research Resources (ORR) is the locally developed hybrid Electronic Resource Management (ERM) and e-resources directory which combines a powerful search engine, A-to-Z list, and subject categorization of all the Library's electronic resources. The ORR is a SQL Server database containing multiple interconnected tables with various data about each resource (title). PHP and PERL scripts are utilized to push out information to the public interface making the most current and accurate information and URL link possible about a particular serial title or database. The knowledge base sitting behind the ORR is populated with data from a weekly feed from TDNet, and data from the Voyager online catalog, EBSCO, Ulrich's, and several other sources. Currently, there are over 40,000 unique titles (resources) in the ORR with nearly 80,000 different instances (iterations) of these titles. Important data populating the knowledge base includes: main title, abbreviated title, variant/alternate titles, title change history, all ISSN numbers associated with a work, URLs, whether it is peer-reviewed, e-coverage, print coverage, current publisher information, subject categories, OCLC number, and current usage statistics. See Figure 2 and to view the ORR in action, see http://www.library.uiuc.edu/orr/

Figure 2 – Online Research Resources (ORR)

In October 2005, the Library went live with its Discover service, using the SFX server and knowledgebase, to provide OpenURL link resolution to full-text journal articles, books, and conference proceedings. Users encounter the Discover service outside of the usual UIUC Library web setting. Instead of searching with the online catalog or the ORR, users discover citations of articles, book chapters, conference proceedings, and reviews within an A&I database and then link out to the full-text article without coming across the digital domain of the Library. Local link resolving was one of the first endeavors to place our services outside the Library web site, where, often, users are conducting their search and discovery work. It has worked out well and the Discover service use is continuing to increase on a monthly basis. (For screen shots of Discover service, see Figures 3 and 4)

Figure 3 – Discover in Engineering Village
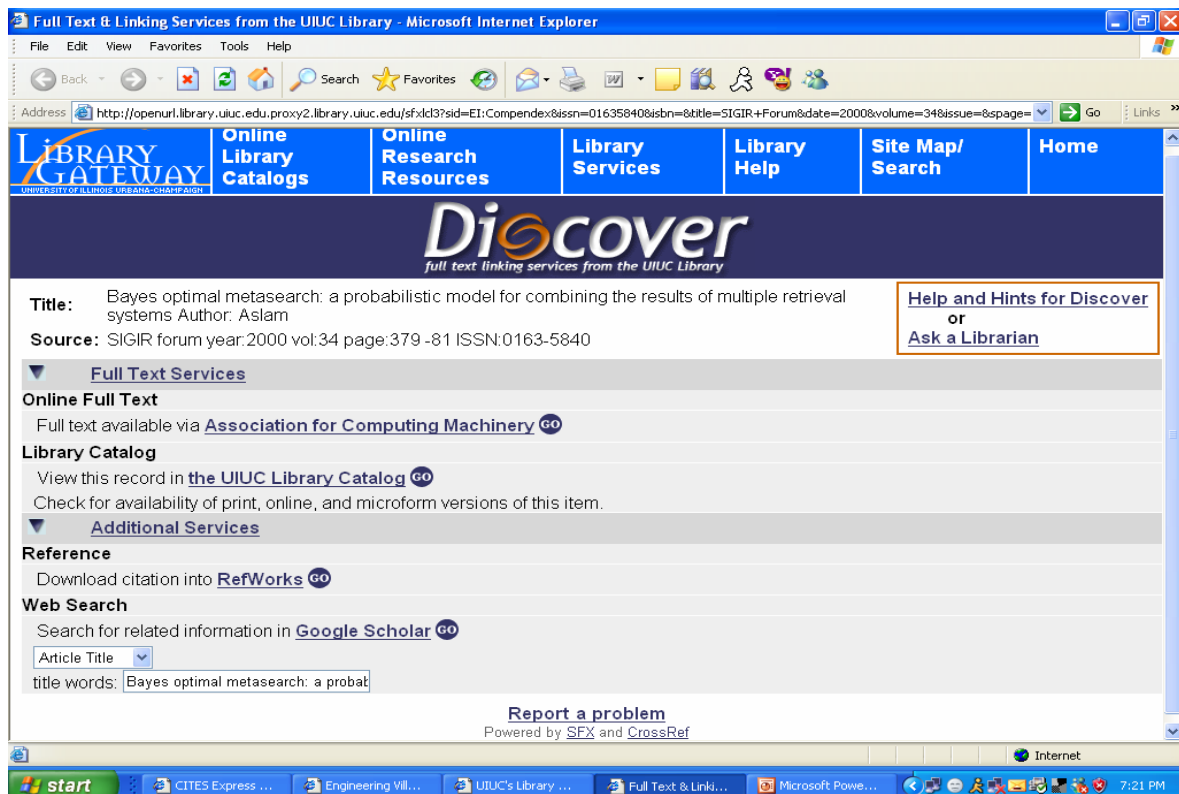
Figure 4 – Discover Results page

To offer federated search across multiple databases and catalogs, the Library unveiled Search Assistant in the summer of 2006. For users, results can include hits from Ebsco's Academic Search Premier, Gale's Expanded Academic ASP, OCLC's Periodical Abstracts, major newspapers from around the world, several business databases, History A&Is, the UIUC online catalog (Voyager), the state-wide consortium universal catalog (iShare) and several other full-text databases. Metasearch is not an easy task. Each database or catalog has its own native interface which may or may not allow easy extraction of results lists. To include "every" possible database and/or catalog available became too cumbersome and created extensive delay in bringing back results. Additionally, databases that had simultaneous user restrictions could not be added as other potential users could be locked out. Search Assistant has been added to the Library's web portal and is establishing its own distinct place in providing library users a way to do interdisciplinary research across various subject areas. (See Figure 5)
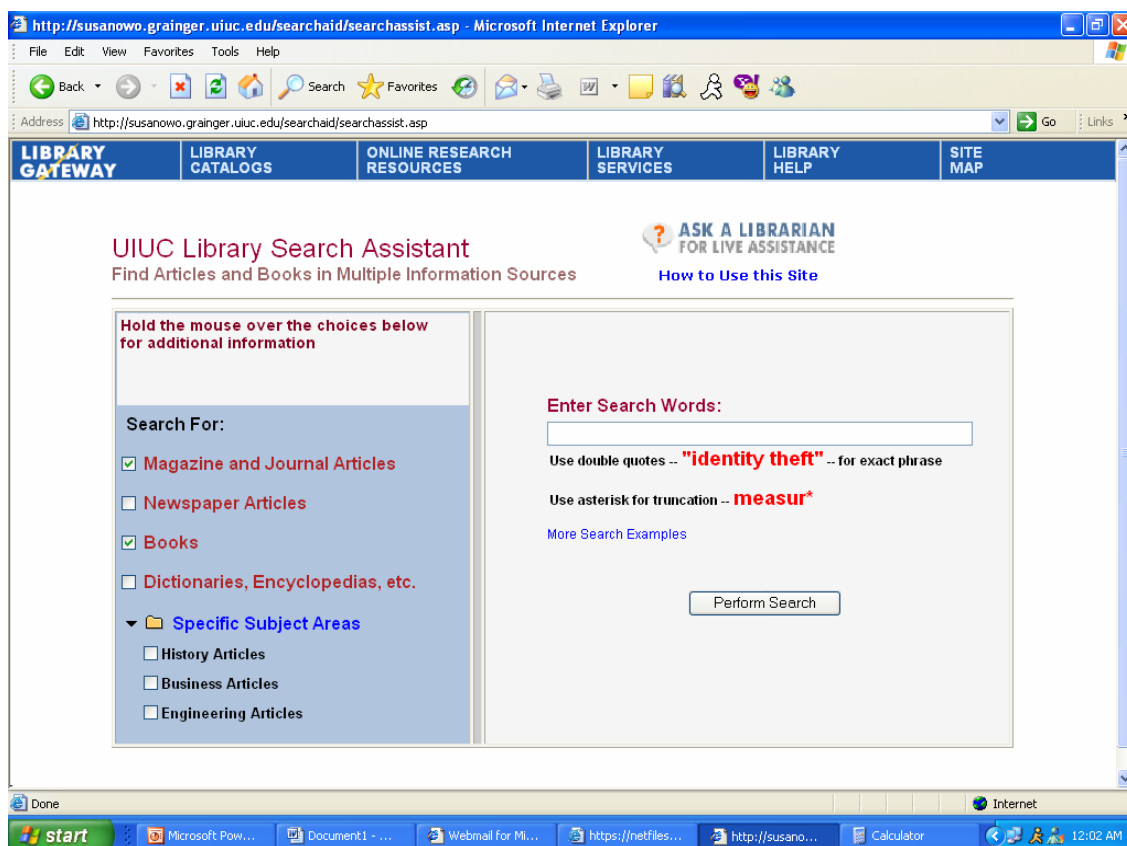
Figure 5 – Search Assistant

Today, many Library users start their research outside any of the Library's systems or catalogs. They start the research process in a more global setting such as Google, Yahoo, or OCLC's Worldcat (WorldCat, Open WorldCat, or WorldCat.org). The most heavily used source to link out to full-text articles in the Library's Discover service (SFX link resolution) has consistently been Google Scholar. To enhance these global Web sites, the Library has embedded its link resolving services and full-text holdings in Google Scholar, Microsoft's Live Search and OCLC's various WorldCat iterations. (See Figures 6 and 7) Thousands of links per month lead from these global search sites back into the

Library's various systems (Voyager, ORR, Search Assistant and Grainger Portal) or directly into the full-text article databases or publisher's sites.

Figure 6 – Discover service in Google Scholar

Figure 7 – Full-text link in OCLC's WorldCat

Use Statistics for UIUC Library Search and Discovery Systems

The UIUC Library has voracious users. In the past year, the Library Web site was visited 51 million times and users downloaded over 5 million full-text articles. They also retrieved an estimated 194 million citations to articles from the Library's collection of journal indexing databases. Discovery of these citations and references led our library users to determine if the content was available digitally. Over the past year, the UIUC Library's search and discovery systems use numbers included:

- Online Catalog – There were 124,000 link outs to the ORR and full-text access; average was over 10,000 link outs per month
- ORR – The system had over 4.5 million clicks over the past year and averages out to over 375,000 clicks per month
- Discover (SFX) service – From October 2005, there have been 314,000 OpenURL requests with 219,000 click throughs to full-text content; we are averaging over 50,000 requests per month
- Search Assistant – Since its beginning in June 2006, there has been over 100,000 user-entered searches. (February 2007) We are on target to have over 300,000 searches over the first year
- Grainger Portal - 50,000 search arguments have been captured since September 2006 (when it went live)
- WorldCat – In 2006, users clicked on 42,000 title-level full-text links within OCLC's WorldCat service in FirstSearch
- Google Scholar – 67,326 OpenURL requests and 53,134 click throughs to full-text content using the Discover (SFX) service over the past year

Library users prefer full-text access. Using the Library's search and discovery various systems, millions of citations and bibliographic references were found and millions of articles, book chapters, conference proceedings and other digital content was downloaded for research use. But, still users need to have the research skills to navigate several different sources to complete a literature review on a given topic. While some integration has occurred within the different systems, the need to create a single search portal that could bring back results from the Library's different systems is quite obvious.

Library users prefer using the most concise and efficient way to find full-text content. Searching one system rather than multiple ones (when the same search results can be achieved) is the optimum resolution to providing timely and successful access to full-text collections. The goal of integrating these different resources into one overarching search portal, pulling the best coverage information from each source led to the creation of the Journal and Article Locator.

Journal and Article Locator (JAL)

The overarching goal of the JAL is resource integration. Library users have difficulty identifying both print and electronic journal holdings due to the need to traverse multiple

systems. At UIUC, the ORR A-Z List, the local link resolver, the Voyager online catalog, and the CrossRef database provide complementary access to electronic and print journal resources. These different access tools augment each other and provide widespread access to journal content only when utilized in aggregate. In this scenario the user must perform separate searches within all of these systems in order to assemble a comprehensive picture of the all of the resources available to them. Integrated searching of these sources is necessary for the user to locate inclusive information about available resources.

To address these interoperability issues, the Journal and Article Locator utilizes metasearch or federated search technologies over the ORR, the local OPAC, the local link resolver server, several A&I service databases, and the CrossRef metadata database to provide integrated and unified access to both electronic and print journal holdings and full-text articles.

The JAL is part of a suite of resource integration tools that are made available within the Grainger Engineering Library Portal (See Figure 8). This Portal emphasizes integrated and extended access to resources and is designed around identified user information needs and information seeking behavior. The Portal users a tab-like portlet approach with four specific search capabilities, offering the user the option of looking for: articles, papers, books on a topic; a specific journal title; a full-text specific journal article; or a specific book or conference proceeding.
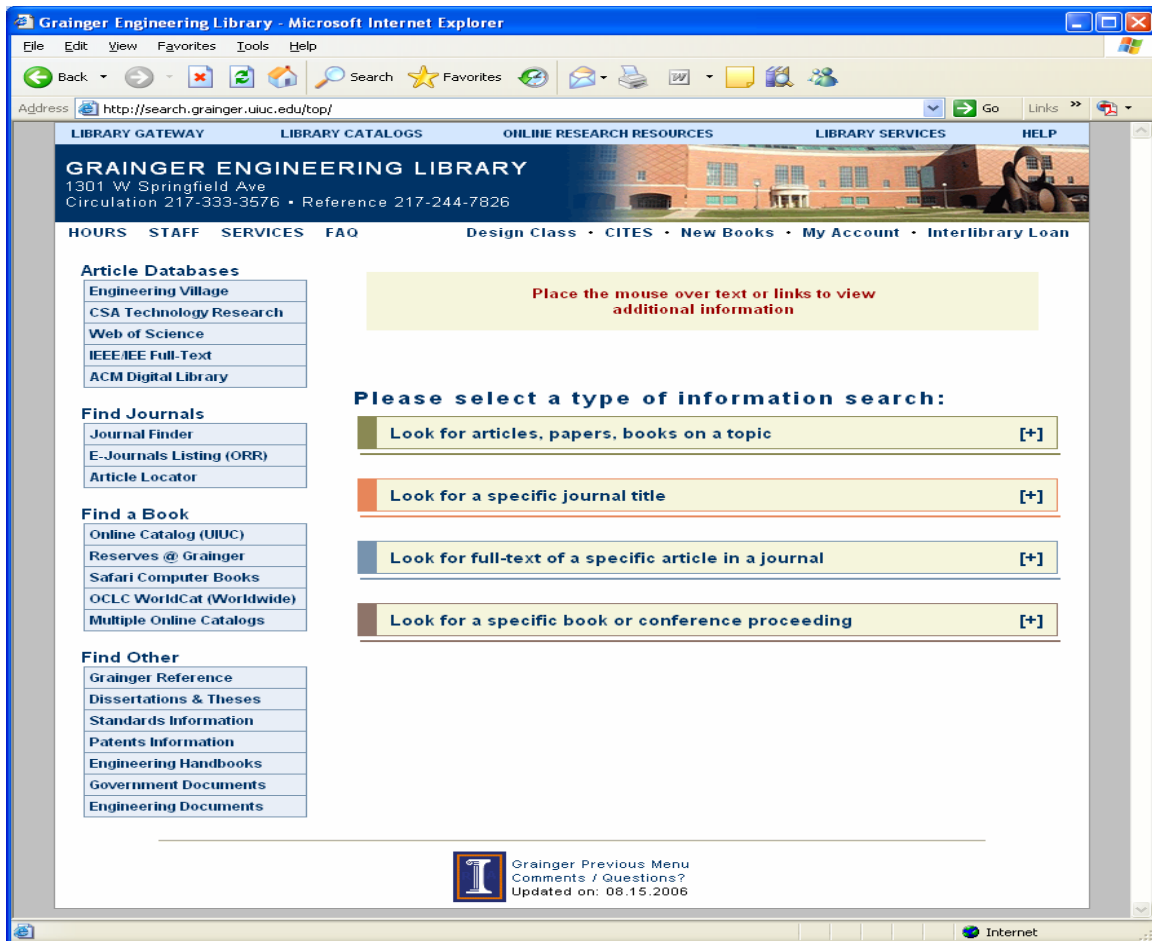
Figure 8 - Grainger Engineering Portal

The JAL operates behind the second and third portlets within the Grainger Portal and as a separate application from within the Library's Main Gateway menu (see Figures 9 and 10).
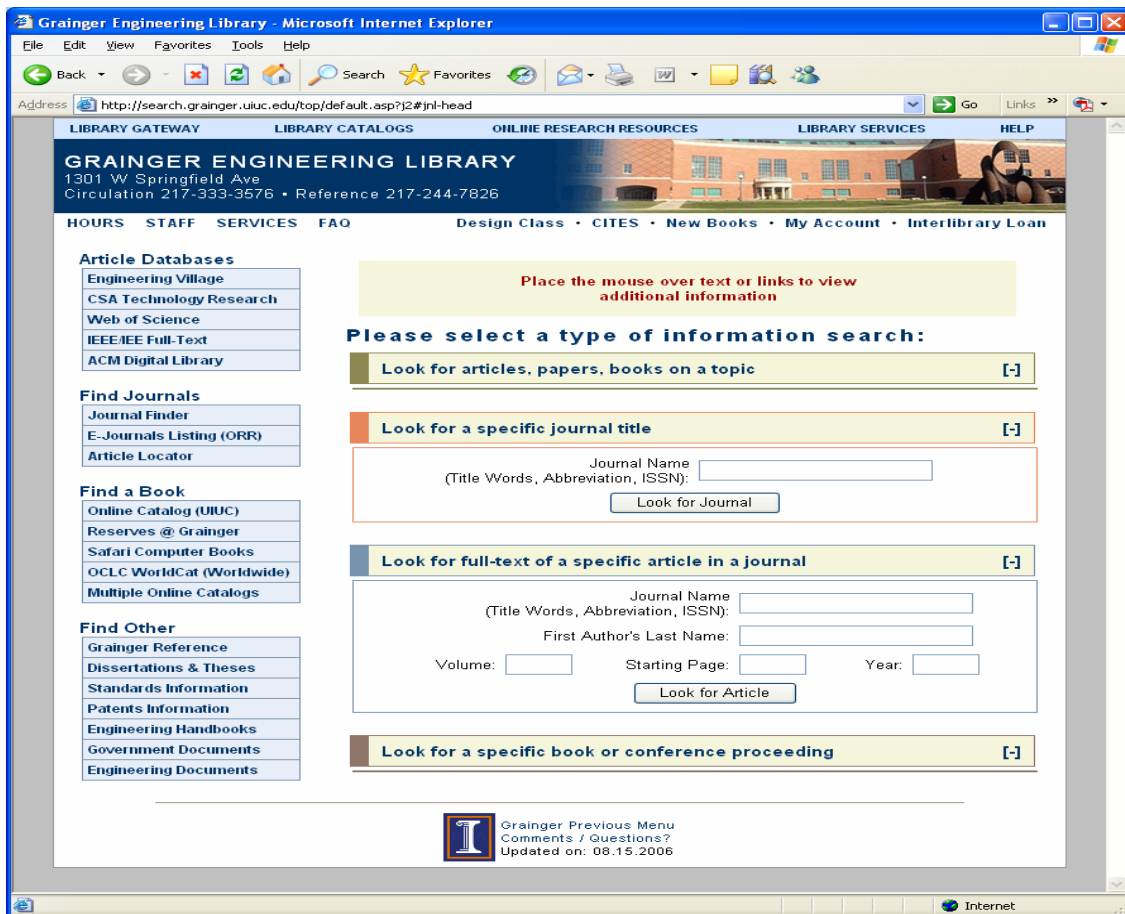
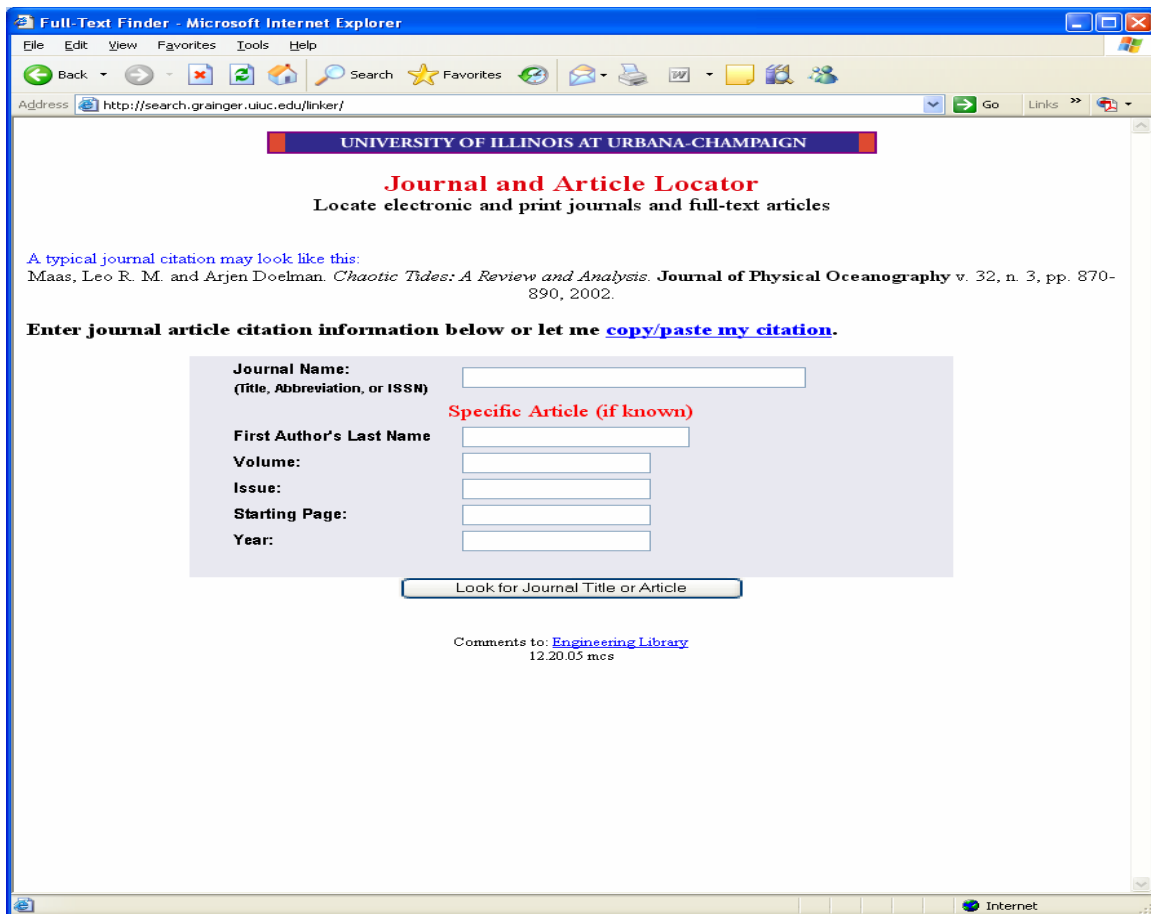Figure 9 - Journal and Article Locator within Portal

Figure 10 - Journal and Article Locator Separate Module

With this growing usage of e-journals, there is evidence that users require more effective search and discovery tools that can provide enhanced access to both e-journal and print journal collections. Data on user behavior gathered in the course of redesigning the University of Tennessee Libraries' Web site showed that searching for information about journals and locating journal articles were the most difficult tasks for users to successfully complete (Mack et al 2004; Maloney and Bracke, 2006). A Web usability study at Western Michigan University revealed that users had significant difficulties finding journal and newspaper articles on assigned topics and were confused about the role of the OPAC vis-à-vis specialized indexing services (Cockrell and Anderson 2002). User testing at the University of Rochester has focused on the phenomenon of "serial failure"—the failure of academic libraries to facilitate students' access to journal articles. The Rochester group argues that addressing these failures will prove to be a survival issue for academic libraries (Bowen et al 2004).

The fundamental approach utilized in the JAL is to employ metasearch or federated search techniques to simultaneously search the OPAC, the ORR, our SFX system, several A&I service databases, and the CrossRef database in order to collect and display full-text and title information from these complementary resources. The JAL performs a broadcast search over the above resources and provides users a custom display of print, electronic,

and, when available, specific article full-text. The resulting display provides users with an integrated single record presentation which is so important for successful user access.

The JAL interface, in standalone form or from within the Portal, provides users with a user search form to enter the required journal title or ISSN and additional full-text article information—if known or desired. The JAL constructs an OpenURL from the user entered data and simultaneously searches the multiple resources using prescribed algorithms. Using the user-entered metadata, the JAL will perform up to eight asynchronous, simultaneous searches and collect the results for customized display. The simultaneous searches include:

- a search of the CrossRef database to extract an article Digital Object Identifier (DOI)
- two ORR searches: a start-of-title match search and a keyword with truncation search
- two Voyager OPAC searches: a start of serial title match search and a keyword with truncation search limited to serials
- two A&I service searches if the user entered author, journal title and year
- a search in the SFX local link resolver.

The JAL follows a flexible software algorithm. If the user has entered full-text article data--such as author name, volume, date, issue, and starting page—JAL will check the journal title in ORR for UIUC licensing and subsequently search the CrossRef database in an attempt to retrieve a DOI. If the DOI is retrieved from the CrossRef metadata database, the user is immediately redirected to the publisher full-text site. This direct link to full-text saves the user from working through additional interface screens and performing multiple clickthroughs. This is consistent with the fact that the Web generation demands the easiest path to full text. If the DOI is not found, the JAL software proceeds with processing the journal title or ISSN and performing the federated search over the OPAC, ORR, A&I services, and SFX systems.

The JAL performs a database lookup on serial title abbreviations and substitutes the full title when found. It also performs journal title abbreviation substitution at the word level, and substitutes common misspellings (e.g. *quaterly*) with the corrected term (e.g *quarterly*). In addition, the software inserts truncation symbols after each word that is longer than 3 characters in both the OPAC and ORR searches.

In cases where more limited information (such as journal title only or journal title plus year) is entered, the search results from the ORR, OPAC, A&I, and SFX are integrated into a single display framework based on our Discover SFX format. The search results will incorporate elements from all three resources into the display, except for the cases where zero results are retrieved. Duplicate results between the two ORR searches and duplicate results between the two OPAC searches are not displayed.

As an example taken directly from our transaction logs, Figure 11 shows a search in the "Look for Full-Text of a Specific Article" option of the Portal for the journal *Applied*

Physics with author *Hindeleh* within volume 4 from 1971. The results, shown in figure 12, indicate that there are a number of word matches for the title in both the ORR and OPAC. However, in this case the citation in Compendex/INSPEC is the most relevant because the article actually appears in the journal title *Journal of Physics D (Applied Physics)*,
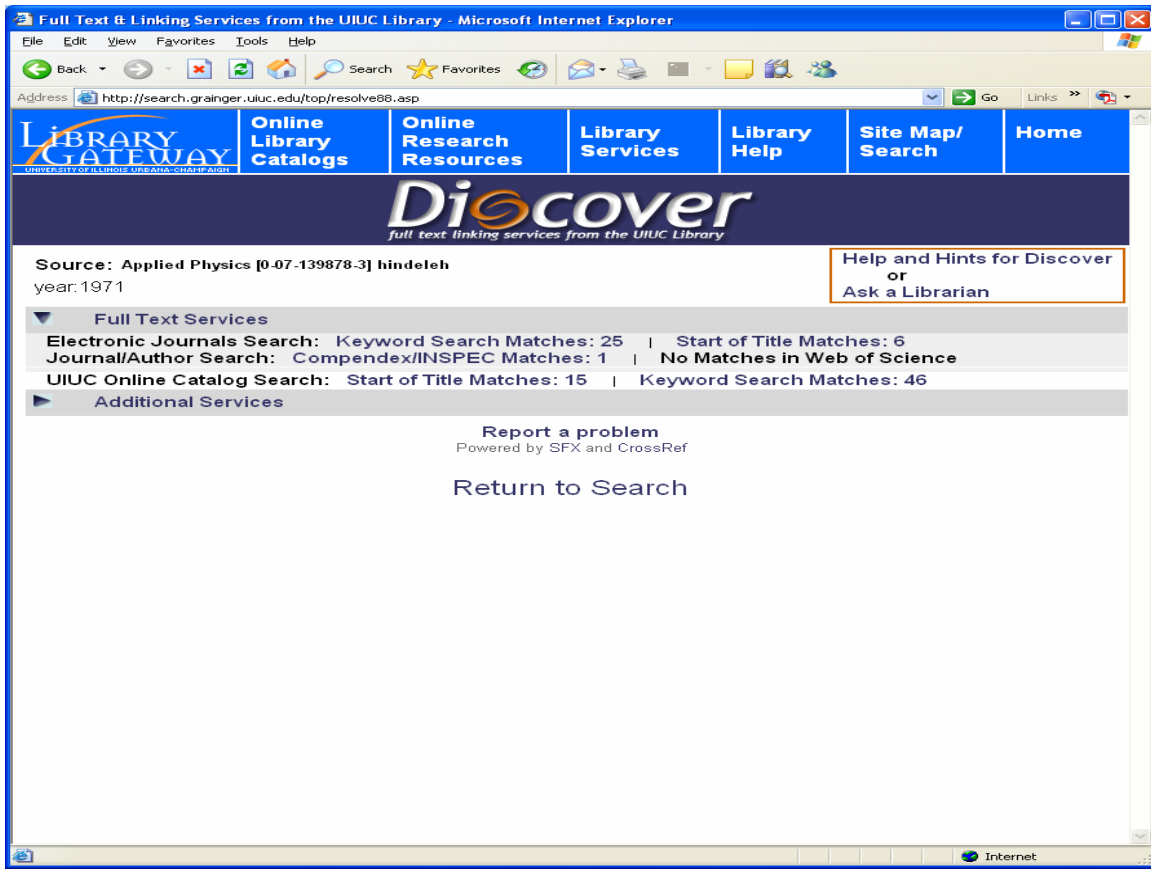

Figure 11 - Locate Article

Figure 12 - Locate Article Results

The JAL performs the functions of the SFX Citation Linker plus it:

- integrates into Grainger Portal functions
- performs CrossRef lookup for DOI and, after ERM TITLE check, goes directly to the publisher site
- normalizes user-entered queries by resolving abbreviations into full titles
- provides access via the OPAC to print only titles, and
- better resolves incomplete article citations than SFX
- Performs broadcast searches on multiple A & I Services and OPAC.

The JAL mimics the behavior of a reference librarian by performing multiple searches in relevant information resources—only it does them simultaneously. The JAL search strategy algorithms are designed to address several common journal retrieval scenarios that have been identified by an analysis of user and librarian search strategies and the retrieval characteristics of the four resources.

Both the Portal versions and standalone versions of JAL are instrumented with all user search arguments, normalized search strings, and system click throughs being captured and stored in transaction logs. We are in the process of monitoring the portal approach in evaluating the effectiveness of JAL. The transaction logs reveal interesting user

behaviors, they show successes and failures, and provide us with possible modifications and change ideas. The JAL is a dynamic system.

It is clear that the JAL represents a successful approach to providing enhanced access to the library's print and electronic journal literature. As expected, the users greatly prefer this one-stop access approach. We will continue to enhance the JAL and other Portal search algorithms, with an eye to better integrating access to the journal literature with access to OPAC data, digital asset management systems, full-text content from mass digitization efforts, and harvested metadata from Open Archives Initiative for metadata Harvesting (OAI-PMH) discovery systems.

Bibliography

Bowen, Jennifer, Judy Briden, Vicki Burns, David Lindahl, Brenda Reeb, Melinda Stowe, and Stanley Wilder. 2004 *Serial Failure*. Charleston Advisor 5(3), p. 48-50.

Cockrell, B. J., and Jayne Elaine Anderson. 2002. How *Do I Find an Article? Insights from a Web Usability Study*. Journal of Academic Librarianship 28(3), p. 122-132.

De Rosa, Cathy, et al. 2005. *Perceptions of Libraries and Information Resources*. OCLC. http://www.oclc.org/reports/2005perceptions.htm.

LibQual. 2004. *Survey Results. Executive Summary.* http://www.libqual.org/documents/admin/ExecSummary%201.3.pdf

Mack, Thura, Maribeth Manoff, Tamara J. Miller, and Anthony D. Smith. 2004. *Designing for Experts: How Scholars Approach an Academic Library Web Site*. Information Technology and Libraries. 23(1), p. 16-22.

Maloney, Krisellen and Paul J. Bracke. 2004. *Beyond Information Architecture; A Systems Integration Approach to Web-site Design*. Information Technologies and Libraries. 23(6), p. 145-152.