One-Stop Search: Finding Full-text Information for Social Science Journals

Rutgers University has made this article freely available. Please share how this access benefits you. Your story matters. [https://rucore.libraries.rutgers.edu/rutgers-lib/51511/story/]

This work is an **ACCEPTED MANUSCRIPT (AM)**

This is the author's manuscript for a work that has been accepted for publication. Changes resulting from the publishing process, such as copyediting, final layout, and pagination, may not be reflected in this document. The publisher takes permanent responsibility for the work. Content and layout follow publisher's submission requirements.

Citation for this version and the definitive version are shown below.

**Citation to Publisher**


**Citation to this Version:**


**Terms of Use:** Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

*Article begins on next page*
One-Stop Search: Finding Full-text Information for Social Science Journals

Triveni Kuchi*

Abstract

Users often express confusion and frustration in trying to locate the full-text availability of a journal and have to check multiple resources and interfaces on a library’s Web site. The challenge was to seek an interim practical solution, which would address this need effectively. The Social Science Journals Database and Search Engine (Soc-dbase) project demonstrates a low-cost one-stop search solution that can be easily and quickly adopted and implemented. The project involves creating a single Web interface to search a database of selected social science and sociology journal titles that include full-text online availability information. This paper presents the design and creation of the social sciences journals database that can be searched to find a journal’s full-text availability at the Rutgers University Libraries or on the Web. The database is accessible from: http://digital-projects.rutgers.edu/socsearch/.

Keywords

ColdFusion, Database, Full-text, Journals, Microsoft Access, Search engine, Social Sciences, Sociology, Web-database

* Triveni Kuchi is Social Sciences/Instructional Services Librarian, Rutgers University Libraries, Kilmer Library, 75 Avenue E, Piscataway, NJ-08854. Email: kuchi@rci.rutgers.edu.
**Introduction**

“Isn’t there one place I can type in the title of a journal to find out if it is available in full-text? Do I have to check a whole host of databases just to determine full-text availability of a journal?” asked one of the faculty members present at a sociology colloquium rhetorically.¹ Many, at the colloquium, joined in by voicing their own experiences of tracking full-text availability. The author has observed users quite frequently express such frustration about locating journals in the library. Similar experiences of users’ exasperation with regard to searching multiple catalogs to find information were also noted in a study of faculty responses to changes in library technology (Starkweather and Wallin 1999, 648). Users with such experiences will prefer convenience of access to quality of information, leading them away from library resources to less substantive resources.

The dissatisfaction that users experience stems primarily from the variety of search interfaces presented on the library Web site for locating these resources and the extent to which these interfaces are unconnected. Ideally, users should be able to search through a simple interface across all the electronic resources available in a library. In reality, this convenience seems far-fetched because such a facility presupposes integration of electronic resources within the library. Any proposition to install such an integrated system faces a number of issues with vendors who participate in the system, technicians who develop the system and libraries that adopt the system. Several hindrances, including feasibility, effort and cost, stand in the way of developing as well as adopting this new technology. In spite of these difficulties, libraries should not settle for the status quo of offering search services with unconnected interfaces for scores of databases on their
Web sites. They need to seek ways to support the simplicity of finding journal information from a variety of different sources.

**Problems and Needs**

Studies show that libraries have been attentive to users’ information needs and are trying to match their expectations. Electronic resources have now become a standard feature on library Web sites. A study analyzing content of 150 library Web sites from four English-speaking countries (including United States of America, United Kingdom, Canada and Australia) found that all of the Web sites provided links to online databases (Still 2001, 161). Libraries, as a regular part of their services, have been directing users to potentially useful sources of information online from their Web sites. A recent article from the Chronicle of Higher Education states, “students and faculty members turn to online library materials before printed ones” (Carlson 2002). It mentions a study contending that students and faculty members at various U.S. colleges and universities have been increasingly seeking online materials for fulfilling their information needs. The users’ demand for more online resources is also reflected in the budget allocations of libraries. According to Association of Research Libraries’ statistics, 106 universities have spent almost $132 million on electronic resources in 2000-01. “The average percentage of the library budget that is spent on electronic materials was 16.25% in 2000-01, nearly five times as much as in 1992-93…the vast majority of these expenditures increases have gone towards the purchase of electronic serials and subscription services” (Association of Research Libraries 2000-2001).
The trend of increased online resources has brought with it the problem of disseminating them in a way that is convenient to the user. Users, especially novices are confused by multiple search interfaces. They are overwhelmed with the number of alternative sources to look up just to find full-text availability status. With too many journal-locating resources scattered across the typical library Web site, finding a particular journal with full-text often becomes a time-consuming and tiresome process. Most of the time, each database needs to be searched separately. Users at the start of their search do not know which of the library’s online resource has the full-text for a particular journal. A user attempts to find information of the full-text journal by a tedious process of elimination, going through the library catalog or a series of lists of links to electronic journals or search engines offered by each database. This diligent drill may not always end in a successful search. Moreover, users are not informed of the alternatives available when the information is not found. Negative results lead to dead-ends, especially since the various resources are not connected. When users, especially students face dead-ends after substantial search efforts, they can get discouraged enough to turn away from library's resources. According to Vassiliadis and Stimatz, “if students cannot locate the information they need on the library’s Web site, they will go elsewhere” (Vassiliadis and Stimatz 2002, 339). In a focus group study of undergraduate, graduate and faculty responses to changing methods of information seeking, Young and Seggern found that the expression `ease of use’ was among the most often mentioned items on the participants’ list of words used to describe characteristics of information and access (Young and Seggern 2001, 163). Libraries will see their collection of electronic resources used more often only if such resources are conveniently accessible and retrievable through intuitive and uncomplicated interfaces.
Conceptually, both libraries and Web search engines serve the same function of providing the means viz., a searchable facility for finding information. Among the Web search engines, Google has increasingly gained popularity among Web users. Its strength mainly lies in maintaining better search algorithms for its search engine and keeping the search interface simple. It has a basic search box on its main page for users to type in search terms. When users submit search requests to a Web search engine such as Google, the results are quickly returned on the very next screen. The users can then browse and select the document or Web site by clicking on a link. Users approach the library’s Web site for finding journal information with this kind of experience and mental framework and, therefore, expect a quality of output similar to search engines while searching library’s electronic resources (Antelman 1999, 176).

Libraries need to adapt a search system that allows searching across multiple databases and brings about results similar to that of Web search engines. To achieve this, libraries will need to install a system with sophisticated technology capable of integrating a variety of online resources. In fact, processes for linking various online resources for libraries are currently being developed and implemented, for instance by organizations such as ExLibris. The development and implementation of such a technology requires database providers to enable linking of resources they provide to the libraries. Libraries will also have to make elaborate preparations necessary to adapt existing library records to the new technology. In addition, many libraries may opt to wait for the new and advanced technologies to stabilize and mature with regard to performance before deciding to adapt them. Furthermore, the cost of this new technology is expected to be high due to its complexity and newness.
In the current environment of budget cuts, most libraries are generally inclined to wait before they can better afford such a costly undertaking. This seems sensible considering that there is a general tendency for prices on technology products to decline over time. Libraries may decide to wait before adapting this technology in view of the requirements, newness and cost of such technologies. Nevertheless, libraries need not let users wait until a long-term advanced integrated solution can be implemented. They should look for alternative solutions to incorporate simplicity into their search systems that allows users to retrieve full-text journal information across a multitude of online resources. To be of much practical value this solution should be capable of being acquired and setup at a minimal cost and time, built and implemented easily, used effectively, conveniently and reliably. The author sought an interim solution to address the problem of finding full-text availability information for sociology and social science journals across multiple electronic resources from a single search interface bearing these criteria in mind. Alternative popular technologies were reviewed for this purpose.

**Technologies reviewed**

The library Catalog or Web-based online public access catalog (WEBPAC) is a primary source for looking up information about books and periodicals in libraries. Exploring the possibility of incorporating full-text journals information into this source seemed to be a logical starting point for solving the problem of finding full-text journals from multiple online sources. The WEBPAC is regarded as a central database representing a large part of a library's collection. This hub-like feature makes it a vital point of access for all types of library resources. It includes information about books, journals and other collections such as videos, sound recordings, etc. Prior to the
availability of online journals, information about journals in the WEBPAC typically consisted of print subscription information. With the increase in the number of online journal subscriptions, full-text links are now being added to the WEBPAC. Keeping this development in mind, the author looked into the feasibility of including full-text availability information of sociology journals into the WEBPAC. The author learnt that technical services and systems departments of libraries face several problems in the process of integrating full-text information for journals in WEBPAC (Day 2000):

- Large number of journal titles require cataloging and updating.
- Online full-text journals exclude relevant cataloging information, which is otherwise usually available in a print copy.
- Database vendors do not provide ISSNs (International Standard Serial Number) for their journal titles.
- Vendors do not offer Machine Readable Cataloging (MARC) record formats that could be conveniently incorporated into the WEBPAC.
- WEBPACs do not reflect accurate holdings when electronic resource vendors add or drop journal titles without prior warning or intimation.

These problems of uncertain availability of information from vendors, inconvenient forms in which information is made available, and additional burden of maintenance on cataloging departments make it complicated to add full-text journal information to WEBPAC. Consequently, the author decided to take a different approach.

Pursuing an alternative approach, the author conceived a project, viz. Social Sciences Journals Database and Search Engine (Soc-dbase) project to devise a mechanism for searching journal
information across multiple sources. This approach considers a Web interface database solution that provides a dynamic response to search requests made to the database. It involves building a search engine and database using a suitable combination of technologies. Popular technologies that are capable of handling such a solution were explored, viz. PERL with MySQL, Active Server Pages (ASP) with Access and ColdFusion with Access.\(^5\) PERL scripting is often used effectively to connect Web-based front-ends to interact with databases. Many PERL scripts are available for free over the Internet, which can be used to create a Web-based interactive search interface. However, in order to carry out this project, some knowledge of application and Web server, scripting languages, database development, and SQL (Structured Query Language) was necessary. Periodical databases of Old Dominion University and Radford University are good examples that use PERL scripts for searching journal titles. Old Dominion University's periodical database uses PERL scripts to draw records with certain fields from their library's INNOPAC\(^6\) catalog. These records together with journal title lists from electronic resource vendors formed the database. This database was queried using MySQL housed on an Apache Web server.\(^7\) Similarly, Radford University's periodicals database employs the PERL scripts from Old Dominion University, but it uses a different database building process. Periodical records from their library's INNOPAC catalog were extracted to a text file and then combined with title lists from online resources subscriptions to create a single file or table for the database. This database, driven by PERL scripts, was made available for searching journal titles (Day 2000; Summers 2000; Hendrickson 2000).

Active server pages (ASP) was the next choice considered for a Web-based interactive database solution for use with Access database. ASP uses a server side scripting language though this is
not a full-fledged programming language such as PERL. ASP is usually coded using VBSCRIPT although it also allows the use of JavaScript. ASP code is embedded in the Web page that can respond to user queries submitted from HTML forms, access any data or databases and return the results to a browser. ASP scripts operate the Web interface by communicating with databases such as Access interactively to provide real-time results for a search. Access is popular and simple to use and has relational database capabilities to build databases. Jackson Library at the University of North Carolina, Greensboro successfully used this combination of technologies to build a Web-based database (Felts 2001).

The third combination of technology considered for the soc-dbase project was ColdFusion and Access. ColdFusion is also a Web application development tool produced by Macromedia that can be used to create dynamic Web sites. ColdFusion is similar in concept to ASP. It is available for all operating systems including Linux, Solaris and Windows Servers. ColdFusion Web applications can support XML, HTML, and other client technologies such as CSS and JavaScript. ColdFusion uses its own tag based markup language, called ColdFusion Markup Language (CFML), which are used to create scripts. These scripts execute SQL commands that interact with the database such as Access and produce results in HTML pages. The convenience of retrieving data from the database and presenting it on to a Web page on the fly has made this software very popular. Virginia Commonwealth University employed this combination of technologies to produce their online journal search page available from http://www.library.vcu.edu/ejournals/index.html (Anderson 2000).
The author compared the three alternatives and found ColdFusion to be the most appropriate choice for the Soc-dbase project. ASP lacks portability in that it works only with Internet Information Server (IIS) Web server. This Web server in turn runs primarily on Microsoft's Windows operating systems (including 2000/XP Pro/NT4) and needs additional software to use it on other operating systems. Further, less experienced developers will usually find PERL challenging to build Web applications. In fact, both ASP and PERL have a steeper learning curve than ColdFusion. Moreover, the development time for PERL as well as ASP is longer relative to that of ColdFusion. ColdFusion is an effective tool that makes building and managing Web applications easy for new developers. Ease of development, minimal programming skill requirements, availability of software in-house, familiarity with scripting techniques and database fundamentals, free online tutorials for ColdFusion and documentation from Macromedia, and support from Rutgers University Libraries provided a strong basis for the author to consider the use of ColdFusion for the Soc-dbase project.

**Description of Soc-dbase project**

The purpose of the Soc-dbase project is to help users locate full-text availability information for social sciences journals. With the help of the search engine devised under this project, users need not search numerous resources such as the library catalog, e-journals lists, indexes and databases lists in sequence just to find information about full-text availability. This search interface lets users find information about a journal in two basic steps: First, the users simply enter the title or a few words of the title in the search engine and click “Search”. In response to the search request, a list of journal titles appears. The users then choose from the list of titles and click on the link to view information about full-text availability for that journal.
The database

The database of Soc-dbase project was populated with title lists from database vendors by using SQL queries in Access as well as entering journal information manually. The author chose to combine title lists from EBSCO’s Academic Search Premier and Sociological Collection and Ovid Technologies’ Sociofile databases. These three databases were considered to cover the area of sociology and social sciences extensively. “Academic Search Premier has a solid collection across most disciplines. It is quite strong in social and behavioral science…” (O’Leary 2001). Sociological Collection “…is a comprehensive database covering information on all areas of sociology, including social behavior, human tendencies, interaction, relationships, community development, culture and social structure” (EBSCO Information Services). Sociofile covers journals from the “Sociological Abstracts” database, which is a key source of journal information in this field (Whitson 2000). These three databases were converted into tables in Excel. Detailed procedures were carried out on the three tables:

- Field names were shortened and blank spaces within words removed. This was done to facilitate running SQL commands on them in Access.
- Records that did not have ISSNs were initially removed from the table. The author then added ISSNs for records with missing ISSNs wherever they could be found from other sources.
- A ‘source’ column indicating the name of the electronic resources vendor was next added to each of the tables to be able to distinguish titles by vendor.

After being processed in this manner, the three tables were stacked as one table in an Access database. The combination of title lists from these three databases produced a set of more than 2000 core sociology journal titles, of which 700 were titles with full-text availability links. SQL
queries immensely reduced the workload of manually entering journal information into the database. In addition, full-text journal titles from JSTOR’s sociology list were added manually to this database.\textsuperscript{12} Journal titles from other databases such as LexisNexis Academic, Dow Jones Interactive, and ABI/INFORM that contain some social sciences journal information were also included in the database, especially if they offered full-text links to journals. The resulting database became the ‘final’ database of the Soc-dbase project.

\textit{Web-based search interface}

A basic interface primarily comprising the search engine was designed using CFML,\textsuperscript{13} HTML, CSS, and JavaScript for the Soc-dbase project. The objective was to keep the search interface as simple as possible so that users can transfer their Web search skills to this interface without any additional learning. Figure 1 shows top half of the search page with title, keyword and publisher options for searching. The title, keyword and publisher input boxes could be used to find a journal title by typing in the exact title, some words from the title or publisher. A tip offered on the search page informs users that a list of all titles in the database will be displayed when a search request is submitted with nothing typed in the search box. This display presents a list of journals from which users can browse and select. In addition, a full-text option is available in the form of a check box. The check box is turned \textit{on} by default, so the user can only search for the full-text titles. At present, this full-text option lets the user search for information of over 800 full-text journal titles. On the other hand, if this check box is turned \textit{off}, the search includes all the titles in the database, both full-text and citation titles. Figure 2 shows the lower half of the search page. This section consists of related links, which are other common ways of searching for social science journal information at Rutgers University Libraries.
Figure 1. Search Interface of SOC-DBASE

Social Sciences Journals Database (SOC-DBASE) with special focus on Sociology

This database provides information about Social Sciences Journals (especially related to Sociology) available full-text either online (Rutgers restricted access or free on the Web) or in print at the Rutgers University Libraries. It can also be queried for citation information for journal titles that are not currently available full-text.

Search

Title:
Keywords:
Publisher:

Limit: Full-Text Only available at Rutgers Libraries and on the Web

Search Clear
Figure 2. Search Interface of Soc-dbase continued.
Results pages

When users submit a search for a journal, a list of titles that match the search terms are displayed if the search is successful (Figure 3). Each of the titles displayed on the list is a hyperlink. The user can make a choice by clicking on the hyperlink for more information about a specific title. However, if no matches were found for a search, a message to this effect is shown. In addition to this message, links to search alternative databases for further information and a return link (to get back to the search page) are provided.
Figure 3. Level One: List of results for a search.
Figure 4. Level Two: Detailed information for a journal title

Here’s More Information for your title

**Title:** BMJ : British medical journal  
**Publisher:** British Medical Association, London  
**Full-Text Availability:** 01/04/1975 through present  
**Full-Text Link:** Academic Search Premier (EBSCO)

Related Links:
- IRIS (Rutgers Libraries’ Catalog)  
- Indexes and Databases  
- Electronic Journals  
- Sociology Subject Guide
Figure 4 shows details of a record that is displayed when a user clicks on a hyperlinked title on the results page. The details include title of journal, publisher, range of dates for which the journal is available in full-text, and links to full-text databases. The full-text link connects to a subscription database such as Academic Search Premier (Figure 5). An alternate link is also provided which leads users to another subscription database link or to a Web site that offers full-text of the journal for free.
Figure 5. Link to the Database where full-text is available for the journal selected
The alternate link points to another library database or to the Web site of the journal (Figure 6). The link to a Web site connects to the main page of such a Web site. It is worth noting here that a direct link to the full-text of the journal is not provided. Instead, Soc-dbase drops off the users at the “doorstep” of the information source, viz. the initial screen of the vendor’s database. Despite the obvious drawbacks of sending the user to a database screen to retype the journal’s search terms, the author decided to use this approach to avoid journal or issue level maintenance of links within the Soc-dbase. Updating of direct links depends largely on the timely notices from the publishers of the databases. The indirect approach is relatively low-maintenance but ensures a highly reliable arrangement for leading users to full-text journals.
Figure 6. Link that opens up a journals’ home page or Web site
Conclusions

According to Ortiz-Repiso, “the library of the future must use the maximum potency that the technology has to offer in order to manage and retrieve information in electronic form; to integrate cataloging, indexing, and online full-text databases; and to provide, in this way, universal and comprehensive access to information” (Ortiz-Repiso 1999, 75). In sync with this thinking, the author examined the possibility of integrating multiple databases and search interfaces so that it would reduce the chaos and confusion that users experience when exploring online resources on a library Web site. The Soc-dbase project was conceived as an interim solution that could be adapted within a short period of time and at a small or no cost. This project demonstrates the value of a single search mechanism capable of finding full-text information about sociology journals from multiple electronic sources. The Soc-dbase project follows an efficient approach, putting to good use resources and skills already available at an institution.

The author carried out a review of appropriate technologies to accomplish the project. Since integrating journal information into WEBPAC did not seem to be a feasible solution, a web-based database solution was sought. This database approach adapts the model of a Web search engine as an interface, which takes a request and returns results extracted from a combination of varied Web sources. ColdFusion with Access was used for the Soc-dbase project since it was found to be the most appropriate combination of technologies, given the criteria considered for the project. The database constitutes the foundation of this project. It was constructed to encompass as many titles as possible that were regarded quintessential to the field of sociology and related social sciences. A search engine was designed to interact with this database to locate and retrieve full-text availability information for the journal titles. This facility has been
operative since the fall of 2002 and is currently available at http://digital-
projects.rutgers.edu/socsearch/.
Notes

1. This remark was made at the sociology colloquium presented by the author at the Department of Sociology in the Fall 2001.

2. The study mentioned in Scott Carlson’s article is entitled “Dimensions and Use of the Scholarly Information Environment: Introduction to a Data Set Assembled by the Digital Library Federation and Outsell, Inc.” by Amy Friedlander, more information is available at http://www.clir.org/pubs/abstract/pub110abst.html.


4. Sociology and related social sciences journals will be referred to succinctly as “sociology journals” throughout this paper.

products” section. Macromedia’s ColdFusion provides scripting via CFML (ColdFusion markup language) and a server for developing Internet database applications. More information is available from http://www.macromedia.com.

6.  INNOPAC is an OPAC system available from Innovative Interfaces Inc.

http://www.iii.com/.


8.  JavaScript is an object-oriented scripting language developed by Netscape.

9.  CSS are Cascading Style Sheets, written in a language that is used to design the HTML Pages.

10. Tutorials for CFML available online from Macromedia’s Web site:


(08/30/02).

11. The Scholarly Communication Center at Rutgers University Libraries provided server space for the project, and ColdFusion software and support.

12. In JSTOR full-text for the most recent 5 years is not available.
13. For this purpose, the author consulted ColdFusion scripts by Ronald C. Jantz, Data librarian at Rutgers University Libraries.

References


<http://www.arl.org/stats/announce/sup01pr.html> (26 December 2002).


Summers, Edward. *Old Dominion University, Norfolk, Virginia.*

*Presentation at a forum on Access to Electronic Journals from Aggregator Sources, Virginia Commonwealth University.* 7 January 2000.


