



## EXECUTIVE NOTES

Due to the ongoing commitment and support of nearly 2,000 participating institutions and 230 participating publishers, JSTOR enjoyed a banner year in 2003. The archive now contains more than 13.4 million pages of scholarly literature. There were over 130 million significant accesses of the archive—up more than 60% from last year. Library participation grew by about 25%, and publishers now contribute some 460 titles to the 11 available collections.

These are pleasing results, but we continue to be challenged by standards that are more difficult to measure. How many scholars still do not have access to the content under our stewardship? I think of the developing world as a key bellwether for JSTOR's progress. There are many special problems with providing scholarly literature to the developing world. A primary difficulty is that, even when institutions have Internet access and skilled users, there are still too many nations where insufficient bandwidth and sporadic electricity effectively bar scholars from meaningful use of online resources.

On these difficult terms, JSTOR's progress in 2003 offers both encouragement and challenge. More than 50 of the poorest nations still do not have JSTOR access in a single institution. But there are also signs of increasing interest from scholars and librarians in developing countries. From 2001 through 2003, usage in developing nations has increased by 280%. That is more than double the usage growth rate of US institutions, and exceeds usage increases in other developed countries.

We understand the high value of electronic resources to scholars and we continue to press forward with efforts to expand JSTOR access in developing nations, working closely with philanthropic foundations and groups of institutions to bring them access and training. There is much work to be done, and this will continue to be a focus during 2004.

*Michael P. Spinella*

## METASEARCHING JSTOR

JSTOR has a dual mission: to be a trusted electronic archive of important scholarly literature, and to extend access to that archive as broadly as possible. In the last few years, we have implemented initiatives guided by both parts of our mission. In terms of access, we have initiated accessibility improvements to our website and search engine, we have provided more seamless access to citation management software, and we have facilitated article-level linking with publishers, libraries, scholars, and commercial resource providers. We realize that students and scholars want to get to information as easily as possible, and so we are committed to making access to the JSTOR archive as convenient as possible.

As part of this ongoing effort, we have been following the recent developments in metasearching (broadcast searching, or federated searching). As you may know, metasearch tools are available that will provide a single point of access to the electronic content available at a particular library, allow a user or library to search disparate electronic resources at the same time, and then aggregate the results sets. In early 2003, the National Information Standards Organization (NISO) announced an initiative to develop guidelines and standards for the metasearching environment. In October 2003, JSTOR participated in a NISO-sponsored workshop on metasearching, and has since been asked to join the committee drafting a set of standards for metasearch protocols of service and support.

While JSTOR does not offer a "broadcast search" function itself, the archive is the target of many metasearch engine requests. Presently, JSTOR has no formal agreements in place with any of the metasearch engine providers or implementers. This is not an optimal situation, for many business and user support reasons, and we are working hard to rectify this. We believe that metasearching can be beneficial, especially for the undergraduate community, and we fully support our participants' efforts to implement metasearch services.

---

However, we (as a community) should not be comfortable with the idea of sacrificing quality for the sake of expedience. Our work with the NISO Metasearch Initiative committee is focused on improving the quality of the metasearching efforts to date, to offer an example of effective partnership agreements in these endeavors, and to help create reliable service guidelines between all parties so that end-user support is as seamless as possible.

At the recent JSTOR Participants Meeting at the ALA Midwinter Meeting in San Diego, Bruce Heterick, JSTOR's Director of Library Relations, outlined several of the technology, user support, and policy considerations that are important in supporting metasearching.

#### TECHNOLOGY

The current technique employed by metasearch tools—issuing HTTP requests to JSTOR, retrieving an HTML page with the number of “hits,” then “screen-scraping” that number from the HTML page to include in a larger, aggregated result set—is neither sufficient nor desirable as a long-term strategy. HTML was designed as a display protocol and was not really built for the robotic data parsing that metasearching requires. As a result, quality suffers. In addition, currency is very difficult to maintain because each time JSTOR changes its search interface, the metasearch engine must make the same changes in order to continue to retrieve the HTML properly. The current authentication methodologies (userid/password, IP) are also problematic because resource providers cannot currently differentiate individual users from the metasearch engines. The metasearching requests originate from the same IP range as the normal searching activity from an institution, and because the metasearch engines do not identify themselves, it is very difficult to distinguish use, especially as that applies to evaluating usage statistics.

From a searching perspective, one of the unique features of JSTOR is lost in the current implementation of metasearch engines. When users search the archive in the native JSTOR search interface, they are required to identify the discipline(s) in which the search should be executed. This feature is lost in a broadcast search to JSTOR, as the metasearch engine has no way to determine which discipline(s) to search. Instead, the metasearch engine simply searches all disciplines in JSTOR. This is inefficient for many reasons, and has at least two unintended consequences: skewing user statistics, and adversely affecting system performance.

#### SUPPORT

Providing quality user support is a critical concern as this juncture. At the moment, if JSTOR receives a support question, for instance, stating that “response time is slow,” we may not be able to effectively diagnose the problem because we are not aware exactly “how” JSTOR is being searched at the institution. Is it in the native JSTOR interface? Is it via a stand-alone metasearch engine, and is that engine hosted locally or remotely? Or is the search originating through an integrated library system that has licensed a metasearch engine and incorporated it into its own metasearching services? Understanding the interplay of these services at any institution

is imperative in providing support. It is important that participating institutions notify JSTOR when these types of services are implemented. As well, it is important that JSTOR have support arrangements in place with the appropriate metasearch engine providers and integrated library system providers so that support can be as seamless as possible for the library and its constituents.

In addition to understanding the impact of metasearching on system performance and user statistics, we also need to understand the impact of metasearching on other types of use of the archive. For instance, how will this type of federated searching affect browsing and printing in the archive? If metasearching begins to significantly alter user behavior in relation to the other functions in JSTOR, we will need to track those changes. Alterations in user searching will impact how JSTOR deals with future research and development on the user interface.

#### POLICY

At the NISO Workshop in October, the question arose as to whether the libraries which had implemented metasearching tools had reviewed their license agreements with the resource providers, prior to implementation, to confirm that metasearching was permitted under the terms and conditions of those agreements. Only a handful indicated that this step had been taken. JSTOR's Archive License Agreement, for example, does not specifically address metasearching at present. Given these new circumstances, it seems appropriate to add language to the License to clearly state what type of metasearching will be supported going forward. In addition, as JSTOR has no current agreements in place with any of the metasearch engine providers, nor the integrated library systems that offer this in their products, it will be important to build the necessary partnerships to allow each party to codify the appropriate business relationships and service level agreements necessary to provide quality end-user support.

There are also questions with regard to the costs that may be incurred in support of metasearching. We need to be able to determine the real system costs (e.g. servers, increased network traffic) as well as the associated support costs. If those costs are significant, as a not-for-profit organization, we will need to outline how we plan to recover those costs.

#### GOING FORWARD

Throughout the first months of 2004, JSTOR is working to define an organizational strategy for metasearching—one that addresses technology issues, interface issues, authentication issues, and support issues. In addition, we will work closely with NISO, and the Metasearch Initiative, to help define standards in this area which will influence this emerging community in a positive, high-impact way. JSTOR will begin to build business relationships with the companies and organizations providing metasearching tools to our participants, and we will start to look closely at how metasearching is affecting user behavior in JSTOR. Finally, we will work over the next few months to build cost models that we can share with the community regarding the attendant costs of supporting metasearching.

The input of our library and publisher participants will be critical as we proceed with this work. Please contact us at [jstor-info@umich.edu](mailto:jstor-info@umich.edu) should you have any comments.

---

STUDY COMPARING THE NON-SUBSCRIPTION  
COSTS OF PRINT VS. ELECTRONIC PERIODICALS

Most academic and research libraries are in the midst of what may ultimately be seen as a transition of formats from print to electronic for various parts of their collections. A key factor driving this transition involves libraries' attempts to efficiently utilize their limited funds.

Eileen Fenton, Executive Director of JSTOR's Electronic-Archiving Initiative, and Roger Schonfeld, Coordinator of Research at Ithaka, initiated a study to learn more about how this transition of formats may alter libraries' operations and their associated costs, focusing specifically on periodicals. The study is part of the planning for the Electronic-Archiving Initiative (<http://www.jstor.org/about/earchive.html>), which aims to preserve digital-format scholarly journals. Also involved in the study were Ann Okerson (Yale University), Donald King (University of Pittsburgh), and Kevin Guthrie (JSTOR Board Chairman).

Findings from this study were presented at the Coalition for Networked Information (CNI) Fall Task Force Meeting in December 2003 (<http://www.cni.org/tfms/2003b.fall/index.html>) and an article-length version of this study recently appeared in the January issue of D-LIB Magazine (<http://www.dlib.org/dlib/january04/schonfeld/01schonfeld.html>).

The study will be discussed in detail during JSTOR's Participants Meeting at the ALA Annual Conference this June. A full version of this study is forthcoming in a publication from the Council on Library and Information Resources (CLIR).

.....

MOVING WALL "FLIP" INVOLVES  
MORE THAN A "FLIP OF THE SWITCH"

Libraries with a mission to archive material for future generations are charged not just with taking care of older literature, but also with taking the steps to preserve new literature as it is published. For journals, this not only means cataloguing and shelving new issues as they come through the door, but also, with the passage of time, binding issues into volumes to protect them from wear and tear and arranging for their storage and retrieval. Preserving journals in the JSTOR archive involves a similar practice. For each publication, JSTOR digitizes the back run one time, but with each passing year new volumes are added to the archive with a "flip" of the moving wall. Selected by publishers and typically ranging from three to five years, the moving wall defines the gap between coverage in JSTOR and the most recently published volume of a journal. For example, for journals with three-year moving walls, JSTOR presently displays content through the year 2000 issues. The moving wall allows JSTOR to provide our library participants and their constituents with a reliable archive of material, while not jeopardizing our participating publishers' current content revenue streams.

Adding content to JSTOR with each moving wall "update," however, involves more than a simple "flip of the switch." The sheer amount of material released every year as part of the moving wall update is quite substantial. Our first update, completed in January 1998, comprised just 40 volumes. The January 2004 moving wall update included 1,157 volumes and 288,554 pages—the equivalent of re-releasing the complete back run of the *Proceedings of the Royal Society of London* (1854–1905), plus the entire back runs of all seven art and architectural titles added to the archive in November 2003. In terms of physical library space, this represents more than 30 linear feet. Despite the scale and complexity of the moving wall update, our goal is for it to be completely seamless for all of the libraries, faculty, students, and researchers using the archive. This requires a great deal of planning within JSTOR.

The process starts in late spring/early summer with our Production unit ensuring that we have the printed versions of all of the journal issues that are to be released. We then concentrate on getting any issues not already digitized ready for scanning and indexing. Approximately four to six weeks later, once this process is complete, JSTOR's Production staff checks the resulting files for accuracy and completeness. One area among many that increasingly requires special attention is encoding and transliteration of non-Latin alphabetic content in the JSTOR metadata.

With the accuracy of the electronic files verified, they are released to JSTOR internal users for further testing. Testers in our User Services unit look over a sample of titles and articles to make sure everything appears properly in the JSTOR interface, that our journal coverage notes have been appropriately updated to include the newly released journal issues, and that those new issues are actually accessible through our public user interface. Once internal testing is complete, our Systems group actually copies into public view all of the new content. This involves mirroring the new content to a number of servers distributed in the United States and the United Kingdom. Careful coordination is required, so as not to interfere with user access in the different time zones represented by JSTOR participating libraries.

Each moving wall update is a significant event that is paramount in our efforts to build and maintain the scholarly archive. Moreover, as time goes by, the annual moving wall update has become even more important to our participating institutions. As one librarian noted in our 2002 Bound Volume survey, "JSTOR gives us flexibility for the future." We see this in library activities and in the cost savings libraries are beginning to realize from this effort. A growing number of librarians have begun relying on JSTOR and the moving wall to make decisions about binding and shelf space on a going-forward basis. When we surveyed librarians in 2003 about their binding and preservation practices, 33% of respondents said that because of JSTOR, they had already stopped binding recent issues, and 23% indicated they had moved bound volumes of titles included in JSTOR to remote storage. As budgets tighten, JSTOR's secure and accessible archive provides library administrators with reliable choices in making decisions about how best to utilize their limited monetary and spatial resources.

The results of the 2003 Bound Volume Surveys are available at: <http://www.jstor.org/about/bvs2003.html>

NEW COLLECTIONS FROM JSTOR OFFER  
BREADTH, DEPTH AND FLEXIBILITY

Over the past several months, JSTOR has announced three new multi-discipline collections. These collections, outlined in the table below, will broaden the number of disciplines covered in the JSTOR archive, and add depth to the areas of study already covered.

COLLECTION	DATE TO BE COMPLETED	# TITLES AVAILABLE AS OF 2/20/2004	MINIMUM NUMBER OF TITLES AVAILABLE (WHEN COMPLETED)
Arts & Sciences III (A&S III)	December 2005	72 titles	120 titles
Arts & Sciences IV (A&S IV)	December 2006	44 titles	100 titles
Arts & Sciences Complement	December 2008	6 titles	150 titles

In the process of adding these collections, we have attempted to offer maximum flexibility to the wide range of libraries and institutions that JSTOR serves. There are 11 collections currently available from JSTOR, including five multi-discipline and six discipline-specific collections.

For some institutions, the discipline-specific collections (Business, Music, and Language & Literature) may be more appropriate for their particular needs. For others, the option to "migrate" from the discipline-specific collections to the

multi-discipline Arts & Sciences collections is important. Since December 2003, 174 institutions have migrated from the Language & Literature Collection and 79 from the Music Collection to the Arts & Sciences III Collection. Since January, 32 participants in Business have moved to the Arts & Sciences IV Collection.

MULTI-DISCIPLINE COLLECTIONS	DISCIPLINE-SPECIFIC COLLECTIONS
Arts & Sciences I*	General Science*
Arts & Sciences II*	Ecology & Botany*
Arts & Sciences III*	Language & Literature
Arts & Sciences IV*	Business
Arts & Sciences Complement*	Music
	Mathematics & Statistics

Institutions interested in licensing all of the content available in the JSTOR archive at this time need only participate in the collections marked with an asterisk in the table above. In 2008, when the Arts & Sciences Complement is complete, institutions that license all of the Arts & Sciences collections will have a single archive of over 600 journals in the humanities and social sciences. The multi- and interdisciplinary nature of the JSTOR archive will enable students, researchers, and scholars to search across a wide range of topics, enhancing the archive as a valuable educational and research tool.

If you would like more information about participating in any of the JSTOR collections, please contact us at: [jstor.participation@umich.edu](mailto:jstor.participation@umich.edu)



149 Fifth Avenue  
New York, NY 10010

JSTOR FACTS  
January 1, 2003-December 31, 2003

Total accesses.....132,069,903  
Searches performed.....30,642,020  
Articles printed.....15,937,227  
Pages viewed.....61,867,755

Total issues available.....91,397  
Total full-length articles available.....968,346  
Total articles.....2,888,827  
Total pages currently available.....13,448,592

Number of participating institutions.....1,941  
Number of countries with participants.....79  
Number of participating journals.....465  
Number of journals available online.....360  
Number of participating publishers.....227

COMMENTS

Thank you for the announcement....It's wonderful to have these back issues available for scholars.... We're delighted with the way *Black Music Research Journal* looks on your Web site.

Marsha J. Heizer  
Associate Director  
Center for Black Music Research  
Chicago, IL

I could check the site this morning and I was really impressed by the results. Your team and you have done a great job in a very few months; the pages are clean and very legible and the word indexation is reliable. This will be shown to the board in a few weeks with an online demonstration; I'm sure they will be very impressed....Congratulations! Great job!

Laurent Guillo  
Revue de musicologie  
Société Française de Musicologie  
Bibliothèque nationale de France  
Département de la Musique  
Paris, France

The email you sent me giving me access to the *Journal of the Society of Architectural Historians (JSAH)* online is the most remarkable thing! I immediately signed on to your site and I had no trouble accessing *JSAH*. I feel like opening a celebratory bottle of champagne! Thank you.

I forwarded your email to our current *JSAH* editor and she was amazed and delighted as well. You provide a wonderful service for the entire academic world and we're honored to be a part of it....Thanks again for sending such good news.

Pauline Saliga  
Executive Director  
Society of Architectural Historians  
Charnley-Persky House  
Museum Foundation  
Chicago, IL

Thanks for sharing this good news [Arts & Sciences III] with us. Clearly JSTOR's success is due to the fact that you anticipate needs of the scholarly community and move forward to meet them.

Dr. Gail Paste  
Shakespeare Quarterly  
Director  
Folger Shakespeare Library  
Washington, DC