



JST's J-STAGE and journal@rchive, SciELO, Microsoft Academic Live

My picks for this issue include two open access scholarly databases from the Japan Science & Technology Agency (JST) and the Scientific Electronic Library Online (SciELO) digital archive from Brazil. These two resources illustrate what is offered by a central agency in one of the world's most developed countries and in one of the low income, but fiercely developing countries. The pan is Microsoft's Academic Search service, which serves up way too little, way too late. Often the information provided is erroneous and/or misleading. The implementation is a slapdash job at best. More importantly, it demonstrates painful incompetence and carelessness by one of the once best-known software companies with no shortage of financial resources at its disposal.

“ Open access
journal aggregators
from Japan, Brazil,
and Microsoft are
picked and
panned.



the picks

J-STAGE AND JOURNAL@RCHIVE

As a semigovernmental organization, JST provides a wide variety of databases and services to promote research and development in Japan [www.jst.go.jp/en]. These include such indexing/abstracting and full-text databases and directories as STN, J-EAST (JST English Articles of Science and Technology), and READ (Directory Database of Research & Development Activities). Some of them are available only for Japanese researchers and/or subscribers, while others are only in Japanese. However, two full-text databases of mostly English language articles, conference papers, and reports—J-STAGE and journal@rchive—offer open access to nearly 200,000 full-text documents. The majority of the documents are publications from 320 different academic societies of Japan. J-STAGE is the more current, with journal@rchive covering back volumes. The cut-off dates vary by journal, rather than a more usual across-the-board, fixed date.

Life sciences, medicine, and pharmacology publications represent the largest topical groups, but there is substantial coverage of serials in chemistry and physics. The remainder are general science journals.

Both databases, which were significantly enhanced in the spring of 2006, are prime sources for any research concerning R&D activities in Japan. The pace of digitization is quite obvious when you look at the update log of JST. It shows that in the first 5 months of 2006, it added more than 50 journals to the databases with various retrospective coverage.

A search on the phrases *brain tumor* or *brain tumour* from the J-STAGE database retrieves 1,073 articles in English.

The results of my topical test searches were impressive. Searching for *brain tumor* brought up 1,178 hits. Using the British spelling *tumour* added 20 more hits. Limiting to English rewarded me with 1,073 hits—all but four of them from peer-reviewed journals published in Japan. When I extended the search to the *journal@rchive* database, the total number of hits from the two databases went up to 1,876. Unfortunately, in the archive, there is no possibility to restrict the search to peer-reviewed English-language documents, nor to some specific fields such as the title or abstract.

JST has developed a truly comprehensive linking system. The cited references in many of the articles are linked through CrossRef, Chemportal, and PubMed to the cited primary documents or its abstracts in various digital archives and indexing/abstracting databases. These, in turn, typically offer open access to the bibliographic data and abstracts, and sometimes to the cited article. Occasionally links to all the issues of a journal appear. This is the case with *DNA Research*, the highest impact factor journal of Japan (published by Oxford University Press on behalf of the Kazusa DNA Research Institute). This journal is entirely open access from the first issue to the most current one through HighWire Press, the digital facilitator of Oxford University Press. It is quite a complex cooperation, and a showcase for the power of using DOI for linking. Thankfully, end users don't need to know anything about this, as the linking to the cited references is totally transparent.

SCIELO

Most people would use three terms in association with Brazil: soccer (football), carnival, and Copacabana. Some might add "beach volleyball" and "Ipanema." I was no exception when I went to Brazil 10 years ago, even though I

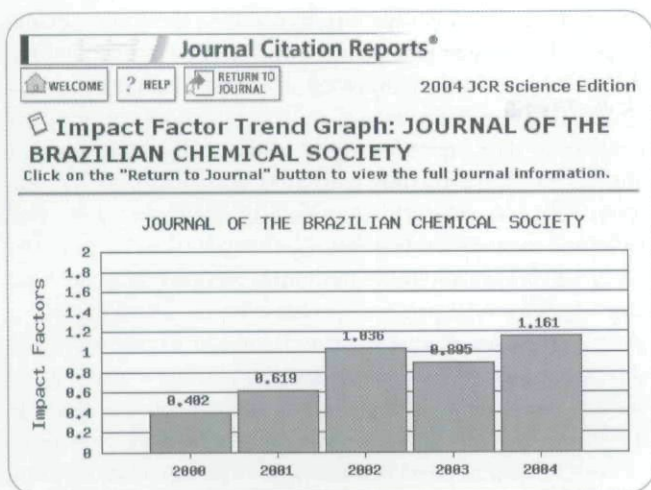
was familiar, through working for UNESCO, with the highly successful software and database development projects of BIREME, the Latin American and Caribbean Center on Health Sciences Information. By now, it is quite apparent that the hundreds of soccer and volleyball players who start their day very early on the Copacabana beach, and the girls on the neighboring Ipanema beach, got to their offices, computers, and labs after such refreshing and invigorating morning exercises to design and implement their long-term ethanol and digital conversion projects. It paid off.

Although Brazil ranks 71st in terms of gross national income per capita, it is light years ahead of far more affluent countries when it comes to locally created, open access digital libraries. (Only India, ranked 129th, can brag about a similar achievement.) The SciELO project [www.scielo.org], the fruit of the cooperation between BIREME and the Sao Paulo State Research Foundation, now offers open access to more than 57,000 full-text articles from more than 250 journals. This includes 150 from Brazil, 53 from Chile, 18 from Cuba, six from Venezuela, along with some additional journals from Argentina, Uruguay, Peru, and Mexico, and 28 scholarly periodicals from Spain. I'd bet Portugal will soon join the Hispano-Lusitano circle of scholarly publications that benefit from the services of SciELO.

The majority of the publications are health science-related articles, but journals in chemistry, engineering, and botany are also well-represented. BioMed Central, HighWire Press, and especially PubMed Central certainly offer open access to far more full-text articles in the above disciplines—but they are in English and reflect the concerns, problems, and solutions prevalent in and relevant to the most developed countries.

SciELO gives a chance to present and read the findings of indigenous research and the solutions to endemic problems that are practical and affordable in the prevailing conditions

The Botany section of SciELO contains 3,577 articles from Portuguese and Spanish open-access publications.



There has been a significant increase of the impact factor of this Brazilian chemical journal.

outside the most developed countries—in Portuguese and Spanish (and sometimes in English). You would not easily find more than 3,500 open-access, full-text articles about botanical research in Portuguese and Spanish in other databases. The same is true for the 3,225 Brazilian articles on engineering in SciELO.

There is another aspect of importance of the SciELO project: It helps disseminate information about research projects carried out in Spain and some South American and Central American countries, thus raising the regional and international clout of several of the journals included in SciELO. It is a quite telling example that the impact factor of the *Journal of the Brazilian Chemical Society* increased from 0.402 in 2000 to 1.161 from 2000 to 2004. Brazil offers an excellent example for many other low-income countries (I use this fairly newly coined, politically correct term instead of poor and developing) of what can be achieved in the dissemination of relevant and pertinent information by worthy and lasting internal efforts, coupled with support from UNESCO and other international agencies.



the pan

WINDOWS LIVE ACADEMIC

As I reported in April 2006 in Péter's Digital Reference Shelf column [<http://reviews.gale.com/index.php/digital-reference-shelf/2006/05/windows-live-academic/>], Microsoft's debut of Live Academic [<http://live.academic.com>] was one of the worst I have seen for a long time. Nothing changed after the debut, even though Microsoft had had almost 18 months to figure out how to make it better than Google Scholar (G-S), to which it is often compared, mostly by bloggers who know

neither well enough, but feel compelled to write about both. Live Academic and G-S are not in the same league.

Live Academic does not do citation counting or searching; G-S does (although poorly). Live Academic does not allow fielded searching (by author and journal name, for example); G-S does (although poorly). Live Academic does not allow limiting by year or discipline; G-S does (although poorly). Live Academic covers only computer science, physics, and electrical engineering; G-S covers many more (although often poorly). Live Academic has 4 million items; G-S has many more (although never reports them reliably).

But wait, doesn't Microsoft claim to have more than 6 million items? It does, which shows how uninformed the designers are about its content, or how desperate they are to beef up its paltry 4 million records. This number is fewer than you find in the Elsevier ScienceDirect segment alone of its ever-improving and expanding open access Scirus system. Add to these the millions of records harvested from the Harvard/NASA ADS database, the arxiv.org preprint server, the digital archive of the Institute of Physics, the Scitation database of the American Institute of Physics, and the intelligent software of Scirus, and you can't really understand what Microsoft was thinking.

Coming back to Microsoft's exaggeration of the size of its database: As Live Academic does not apply a stop-word list, I could do the simplest size test by searching for the three most common words in English texts. The word **the** found 3,979,838 hits, the word **a** found 3,925,997 hits, and the preposition **to** found 3,914,027 hits. Tell me again, how many items there are in this database?

The PR announcement (masquerading as FAQ) is full of other feel-good (but false) claims and statements. It's difficult for me to rave about the side panel, which provides more details about the items over which the mouse hovers. And don't believe the claims that the "abstract for this document is unavailable." Often it clearly *does* have one, as evidenced by a click on the link that takes you to the source where a nicely presented, so-called "structured abstract," or summary, is right under your nose. I show this on an article about abstracts for fun, but it was no fun when I saw this message in about 80 percent of my test results. It was obvious to even a modestly experienced librarian like myself, that, judging from the name of the journal, the length of the article, or the snippet in the

Abstract to publication ratio for papers presented at scientific ... - Published Version (2001)
Walby, Andrew | Kelly, Anne-Maree | Georgakas, Con
... This article is cited by the following articles in **Blackwell Synergy** and CrossRef The **abstract** to publication rate for papers presented by Australasian emergency physicians and trainees at ACEM/ASEM ...
Search Web Hide Abstract

Title: Abstract to publication ratio for papers presented at scientific meetings: How does emergency medicine compare?
Abstract: The abstract for this document is unavailable.
Authors: Walby, Andrew | Kelly, Anne-Maree | Georgakas, Con

Does it or does it not have an abstract? It's unclear at Microsoft Academic Live.

result list panel of Live Academic, the item must have an abstract. Often these are clearly labeled as abstracts or summaries. Martha Stewart could not have wanted her spice containers labeled any better to look sharp in her TV show.

Then again, it is still better than the EndNote export format that *never* includes the abstract, the name(s) of the author(s), the digital object identifiers (DOI), and several important metadata elements. As for the feel-good talk about sorting, it is just that. Sort the result of the search about the query on "sorting results," and you will find that Börner, Katy is first (omitting the second letter, changing her name to Brner), followed by Wang Jianchao and Yan Yuanyuan (quite a big leap in the alphabet for 105 results), then by Wang James Ze, Grammatikakis, M., Zhou, L, and Krovetz, R. Sorting by journal name is equally inane.

The result list without borders is more hogwash. You may feel the freedom of not being sandboxed in a result list of 10, 20, or 50 items per page, but in reality, this design makes it very inefficient to look up an item and then go back to the top of the unnumbered result list and scroll down.

Listing the sources covered is to be celebrated, particularly when Google Scholar is so secretive about this and many other basic questions about its database. When you look at the source list of journals of Live Academic, however, you see

triplicates and quadruplicates of many of the titles, spelled inconsistently and/or erroneously. Then you realize that the real number of sources covered is far less than claimed.

I like Microsoft reference products, which were among my picks years ago. I greatly respected the software developers and the content specialists (cartographers, historians, art historians, etc.) of, say, the Virtual Globe and Encarta, when I had a chance to spend 2 days with them some years ago at the Microsoft headquarters in Redmond, Wash. I can't work up any respect for this sorry slapdash product. I was invited to participate in a 2-day meeting with the Live Academic group this past January, but I could not make it without missing two classes. Now I regret it, because apparently no one told the designers that this is not the way to do this type of project. Google Scholar needs a little competition, but it has nothing to fear if Live Academic remains the "competition"—and if bloggers and journalists keep praising it as elementary school children and adults in North Korea praise the Dear Leader.

Péter Jacsó [jacso@hawaii.edu] is professor of library & information science at the University of Hawaii's Department of Information and Computer Sciences.

Comments? E-mail letters to the editor to marydee@xmission.com.

LEARN HOW TO TURN SEARCH INTO RESEARCH AT WEBSEARCH UNIVERSITY!



websearch

UNIVERSITY

POWER SEARCHING WITH THE PROS


- Washington Marriott
Washington, D.C.
- September 18-19

WebSearch University is 2 packed days of nonstop, intensive instruction about how to search the Web—and find what you are looking for! Taught by an experienced faculty of well-known super searchers, WebSearch University offers a fast learning curve that teaches you how to master the strategies and build the skill sets you need to become an expert searcher.

TO REGISTER OR FOR MORE INFORMATION:

Web site: www.websearchu.com
E-mail: custserv@infotoday.com
Phone: (800) 300-9868 or (609) 654-6266
Fax: (609) 654-4309

Organized and Produced by

 Information Today, Inc.

WWW.WEBSEARCHU.COM