

péter's picks & pans



Péter Jacsó
University of Hawaii

The picks include two databases that are equally good or better substitutes for two services that were, or are likely to be, discontinued. One is (in spite of its deficiencies) the BizTech Library, because it provides a lot of bang (50 articles) for the buck (the \$7.95 monthly subscription fee), with the full text of articles from nearly 5,000 professional and scholarly journals. The other pick is the Energy Citations Database from the U.S. Office of Scientific and Technical Information (OSTI) that, except for one feature, is far superior to the PubSCIENCE database. The pan is Yahoo! for abandoning subscribers of its Research Document Collection.



the picks

BizTech Library

Energy Citations
Database

Yahoo!

BIZTECH LIBRARY

Ziff Davis Media Inc., the largest publisher of information technology and specifically personal computer journals, has been struggling for years with how to have paying customers for the online versions of its own journals and ones licensed from online aggregators. I have been using ZDNet's BizTech Library [<http://cma.zdnet.com>] since Ziff-Davis created CMA. (Computer Magazine Archives—the name still appears in its URL, though it's also known as ZDNet TechInfoBase. ZDNet is now owned by C|NET.) I will use it even more, now that it offers a far larger collection of many library and information technology journals, something not apparent from its name.

Its history is too convoluted to recount here. Suffice it to say that in its latest incarnation, there's an archive of about 370 information technology journals (drawn from Gale's Computer Database Plus), and more than 4,000 other journals of general and specialist interest (contributed by divine's Northern Light Special Collections, which in turns pulls articles from Gale Group, ProQuest, and Janet Matthews databases, as well as various newswires). It also tells you a lot about how little ZDNet knows about what it offers in this collection, as it still brags about having more than 400,000 articles. Make that 4 million and then some.

This implies that there is significant duplication and redundancy in the search results, as many of the journals appear in both collections, including the titles of Information Today Inc., such as *ONLINE*, *Computers in Libraries*, *Information Today*, *Searcher*, and *Link-Up*.

With that said, you still receive a lot of bang for your monthly subscription of \$7.95, which entitles you to 50 full-text articles—that's only 16 cents a pop. This is not as good a deal as we had for the \$4.99 a month with Yahoo! Research Document Collection (also sub-licensed from Northern Light) with the same 50-item limit from the same set of journals, but who can complain? Certainly not those who have been paying for the same text-only documents at about \$3.45 an article through Dialog, with an online search fee of \$80 per hour.

ACCESS THE KNOWLEDGE OF TECH BUSINESS

The BizTech Library is an important resource for professionals searching for the latest research on business and technology trends in their industry. With ZDNet BizTech Library you gain access to thousands of specialized articles and abstracts for \$1 during the first week trial. After the one week period, if you would like to continue with the service, it's only \$7.95 per month.



NEW: BizTech Library now lets you search 4000 more publications from the Northern Light Special Collection! [Find out more.](#)

Already a member? [Log in here](#)

WITH ZDNet BIZTECH LIBRARY YOU RECEIVE:

- A timely archive of top business and IT magazines, sites and journals
- Full access to hundreds of tech and business publications
- Specialized information for individual industries
- Thousands of premium full-text reference sources from Northern Light

BizTech start page with database size underestimated by an order of magnitude.

A search for XML gives some sense of the size of the BizTech Library, since it yields 6,745 items from the core collection and a whopping 61,239 items from Northern Light. True, there are way too many newswire and news items in which the propaganda blurb about the companies outweighs the essence of the news. But even after excluding newswire from the results the total number of articles that mention XML is 47,862.

As another telling example, the exact phrase query "copyright infringement" yielded 1,367 and 8,191 items, respectively, from the two portions of BizTech Library. The results set can be reduced by restricting the search to specific start and end dates, company and product names, and—in the case of journals with distinctive names, such as *Reference and User Services Quarterly*—to a specific journal. The journal limit feature is very poor. You cannot enter a journal title, but can only choose from a pop-up menu. It lists the journals of the smaller subset, not those in Northern Light. Even if you believe that you picked a specific journal, like *ONLINE*, from the list, you did not. Every journal that includes the word "online" will be retrieved, often crowding out and pushing to the very bottom of the results list the items that are from *ONLINE*. I was not pleased when I want-

ed to find an article from *ONLINE* about copyright and received hundreds of articles from *MacWeek Online*, *PC Week Online*, *Online Newsletter*, and *Electric Vehicle Online Today*.

With such deficiencies you may wonder how the BizTech Library ended up on my pick list. By virtue of the appealingly low cost and the good selection of journals and magazines. If you use your monthly budget of 50 items intelligently and do not display the full text of articles that you can find for free in the FindArticles database [www.findarticles.com], you have an excellent return on investment. With a little brainwork, ZDNet could greatly improve the software to make

BizTech Library an unconditional pick for many customers who are lamenting the loss of the very similar Yahoo! service from a practically identical content source.

ENERGY CITATIONS DATABASE (ECD)

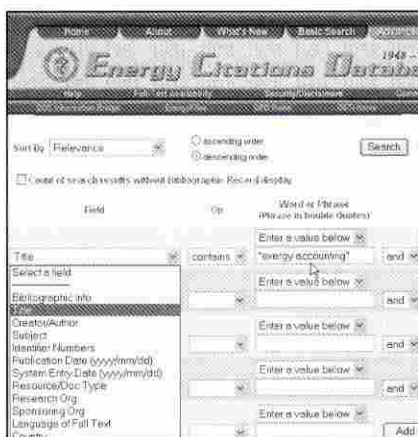
The free Energy Citations Database [www.osti.gov/energycitations] is a subset of the commercially available Energy Science & Technology database, available on Dialog as File 103, on STN as ENERGY, and on SilverPlatter's CD-ROM ETDE Energy Database. Different versions are available on STN, Dialog, SilverPlatter, and MISC on CD-ROM and online. The NISC online version is by far the most comprehensive version. ECD has over 2 million records, 50 percent more than PubSCIENCE. In actuality, it's almost twice as large if you discount the very large number of duplicates and triplicates in PubSCIENCE, which are minimal in ECD. One of the reasons for the much larger size of ECD is that it laudably includes records not only about journal articles, but also about books, conference proceedings (55,250), conference papers (235,200), technical reports (602,700), and patents (107,000). These numbers are approximations because document types are assigned at two levels—a journal article may be the primary document type and a conference paper the secondary.

As for the content of the records, ECD has descriptors, assigned from the Energy Thesaurus, for every item. PubSCIENCE also has descriptors in records taken from the ES&T data-

The screenshot shows the BizTech Library search interface. At the top, it says "BIZTECH LIBRARY" and "SEARCH OVER 400,000 ARTICLES". A search bar contains the text "xml". Below the search bar, it displays "61,239 items found" and "Sorted by Relevance". The main content area is titled "Extended Library Results from Northern Light" and lists three search results:

- XML Global Signs Global Joint Marketing Agreement with Sun Microsystems -- New Y... announced that it has signed a joint marketing agreement with Sun Microsystems [SUNWW] which will market and sell a wide range of XML Global's ... **Phillips Publishing International** Added 09/16/2002
- XML Server Software Market Continues to Grow, Driven by Integration Demands from... FRAMINGHAM, Mass., Sep 12, 2002 /PRNewswire via COMTEX/ -- The worldwide XML server software market will sustain its momentum over the next five years as the ... **PR Newswire**: Added 09/12/2002
- XML Application Firewalls Secure Expected Explosion of XML and Web Services Traf... MOUNTAIN VIEW, Calif., Sep 9, 2002 /PRNewswire via COMTEX/ -- According to a recent report by ZapThink, a leading industry research and analysis firm focused on ... **PR Newswire**: Added 09/09/2002

Whopping number of articles mentioning XML.



The index entries for many of the data elements are also browsable, and far more search criteria can be used on the appealing search template.

base, but the records submitted by publishers who partnered with PubSCIENCE almost never have descriptors. The few publishers' records that do have descriptors represents 0.6 percent of those records, and the descriptors are not taken from the Energy Thesaurus. OSTI does not touch these records, just shoves them into PubSCIENCE *en masse*.

When both the ECD and the journal publisher have a record for the same article, the former is generally bumped in favor of the latter from PubSCIENCE. This is a very bad strategy, as ECD records are typically far more rich in content, including not only controlled vocabulary descriptors but many extra value-added pieces of information, such as the sponsor's name, the identifying numbers of a report, language, and research organization. Presumably, this practice will cease if PubSCIENCE is discontinued, as OSTI has proposed.

ECD also has a far better layout than records in PubSCIENCE. Examples are in my PowerPoint file presented at the Hawaii Library Association conference [www2.Hawaii.edu/~jacso/hla/HLA-3.ppt]. When the nation's librarians mourned the imminent demise of PubSCIENCE, they were apparently unaware of its serious deficiencies. They also hadn't appreciated the proficiencies of its sibling, which was treated by OSTI's PR machine like a stepchild in a Grimm tale.

The fame of PubSCIENCE is that purportedly it can instantly link the user to the source documents if the user has a subscription for the digital version of the journal. Nice PR talk swallowed hook, line, and sinker by many. Even in theory, such links would apply to less than a third of the PubSCIENCE database, so I don't really buy this argument against the far superior ECD.

In practice, a large percent of the publisher-submitted records have no URL, or no appropriate URL, for the articles. Many of the URLs take you to the publisher's home page, something you could do through a well-organized bookmark system. Other times you are getting to a wrong document because the link is not unique or unambiguous. True, there are good links in PubSCIENCE, while the ECD records have no such links at all, but the records can be enhanced with links to articles, patents, and other documents for a tiny fraction of the cost wasted on PubSCIENCE.

As for the software, ECD runs circles around PubSCIENCE. It is intuitive and powerful. The index entries for many of the data elements are also

browsable, and far more search criteria can be used on the appealing search template. The results in ECD can be sorted by title, date, or relevance. The record content and layout is better in ECD. In PubSCIENCE there is no rhyme or reason for the sequencing of the results, as is clear from the scattered listing of duplicates and triplicates that should line up next to each other like twins in a Sears photo shop poster.

It is quite telling that the phrase "alternative fuel" in the title occurs in only 66 records in PubSCIENCE, with a single publisher-submitted record. The number of unique records of the 66 items is less than 40 because of the many duplicates and triplicates. How many of those records have links? One.

There are 202 records in ECD with "alternative fuel" in the title and hardly any real duplicates. As I wrote in my *Information Today* column and illustrated in depth at the Hawaii Library Association, there is no reason to weep over the loss of appropriation for PubSCIENCE, whose excessive propaganda almost completely overshadowed the development of ECD. I see one advantage of keeping PubSCIENCE available—using it for education, like medical schools use the cadaver from a disease-ridden person. It can be contrasted with the X-rays, blood samples, and vital statistics of a healthy person, or in our case, the ECD database. I am doing just that in a systems analysis course in which we redesign the highly redundant bibliographic information system of OSTI, which hides some gems. So I suggest that U.S. citizens write to your repre-

Title	Energy accounting of energy and material flows in steel production systems
Creator/Author	Costa, Marcio Macedo, Schaeffer, Robert, Worrell, Ernst
Publication Date	2000 Apr 01
Report Number(s)	LENL-47306
DOE Contract No.	AC03-76BFD0099
Other Numbers	R&D PROJ: 434001 ; TRN: AF200129%*131
Resource	Journal Article
Doc Type	
Resource Relation	Energy, the International Journal, VOL. 26, ISSUE: 4, Journal Publication Date: April 2000
Research Organization	Lawrence Berkeley National Lab., CA (US)
Sponsoring Organization	Environmental Protection Agency (US)
Subject	36 MATERIALS SCIENCE; ENERGY PRODUCTION; STEELS; ENERGY MATERIALS
Related Subject	ENERGY ENERGY EMISSIONS
Descriptions/Abstract	No abstract prepared.
Country of Publication	United States
Language	English
Format	np
OSTI Identifier	OSTI ID: 783515
System Entry Date	2001 Aug 30

Author	Costa, Marcio Macedo, Schaeffer, Robert, Worrell, Ernst
Title	Energy accounting of energy and material flows in steel production systems
Journal Title	Energy, the International Journal
Date	Apr 01 2000
Journal volume	26
Issue number	4
Identifier	783515
Subject	energy energy emissions
keyword	

Comparing the results layout from ECD and PubSCIENCE for the same record shows better content and layout in ECD.

sentatives in Congress to earmark money for enhancing ECD, and go use it even as it is now. You will be glad to have done so.



the pan

YAHOO!

Yes, I know, I chose the Research Document Collection of Yahoo! [www.yahoo.com] in the July/August 2002 issue as a pick in this column. I maintain that it was a good idea, very well integrated into Yahoo!'s search results. The \$4.95 cost for the monthly subscription entitled you to display and save 50 articles from qualifying documents (which numbered about 4,000). It was too inexpensive to be true. I could have lived with higher prices and still been happy.

What I resent is that after 6 months, Yahoo just pulled the plug without even bothering to notify the users. I was in the middle of a class when I wanted to demonstrate it and did not want to believe my eyes when the Research Document Collection results did not show up. Yahoo! had the e-mail of every user and could have alerted them that the service would be ditched, not just walk out on us as a yahoo (in Jonathan Swift's *Gulliver's Travels* "crude and brutish person" sense of the word) from a failed relationship.

Yahoo! has diluted its once excellent directory by throwing into it worthless sites—check out the mix of the good, the bad, and the useless in the site listings under the encyclopedia section in the directory. Now this latest move indicates to me that Yahoo! may be determined to earn its Swifitian yahoo name by its acts and attitude. I hope it is not going to drop unceremoniously the good, quality reference works it licenses.

Péter Jacsó [jacsó@hawaii.edu] is Associate 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.

Information Today, Inc.

invites YOU
to subscribe to

NewsLink

A **FREE** weekly e-mail newsletter, NewsLink is designed to highlight the information that both users and producers of information products and services need to do their jobs as effectively as possible. By subscribing, you will receive a full-length issue on the first of every month, as well as brief weekly updates every Monday.

Visit

www.infotoday.com

to subscribe!

NewsLink



Information Today, Inc.

143 Old Marlton Pike, Medford, NJ 08055