



## AnaVist, Book Review Digest Retrospective, and Past Pans

**M**y picks for this issue include STN's AnaVist and H.W. Wilson's Book Review Digest Retrospective. AnaVist is STN's information analysis and visualization tool that smoothly integrates with concept search capabilities when searching the CAPlus, WIPO/PCT patent, and USPTO full-text databases. Book Review Digest Retrospective is H.W. Wilson's collection of book review excerpts ranging from 1905–1982, its prequel to Book Review Digest. Instead of having a single pan as I normally do, in this issue I review what happened with some of the databases that I panned in the past 2 years. Did they get better, remain the same, or decline even further? The answer is highly discouraging.



### the picks

#### STN ANAVIST

I must admit that I am no expert in searching STN. I used it a long time ago, and even then only when I needed some special features of the STN implementation of Science Citation Index, such as the occurrence count of search words in the retrieved records. However, a recent demonstration of STN caught my interest. STN generously supplied me with a password and I went off to explore the landscape represented by its search results. The most appealing feature of AnaVist is that it integrates the search and analysis/visualization process. There are several analyzer/visualizer programs, such as Tableau, Omniscop, and HistCite, and even free software, such as CiteSpace, which processes bibliographic data very well once the result set is saved in tab-delimited, CSV, or RIS formats and imported into the analyzer/visualizer program.

AnaVist can import search results for analysis and visualization from several STN databases, but its specialty is the mapping of customizable data elements from CAPlus, the USPTO full-text patent database, and the WIPO/PTC international patent database. AnaVist doesn't require the user to go through the export/import process. This becomes attractive when you realize, in the analysis/visualization process, that the set you created in the independent search process needs refining, filtering, and exporting/importing time and again. It is like stopping at busy intersections, getting off your bike, walking it across the intersection, and getting on it again versus smoothly gliding through the intersections.

In AnaVist you can search, then swiftly analyze and visualize the results in an iterative fashion, without losing momentum. The integrated concept search engine, which dynamically enhances the search terms with synonyms, lacks some of the features of STN Express. However, it is useful for creating a relevant set appropriate

for analysis and visualization. Examples are producing assignees/affiliations grids, researchers/publication years matrices, and contour and cluster maps to show the lay of the land in the research area defined by the query. There are many alternatives for user-chosen preferences, and navigating through the options is intuitive. The bird's-eye view plane of quadrant is a good starting point for zooming in for details.

I would like to see some additional features, such as the highlighting of matching query terms. I'd also like more databases to be made searchable individually with the results deduplicated, then analyzed and visualized in AnaVist (such as the Science Citation Index, INSPEC, and MEDLINE).

### BOOK REVIEW DIGEST RETRO

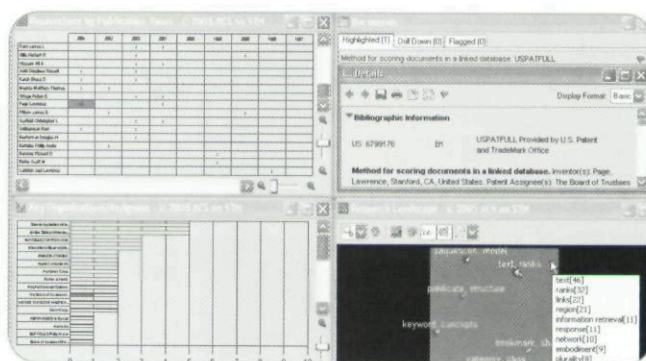
I have always thought that H.W. Wilson takes an excessively shy approach to public relations and advertising its products. It resembles the blushing modest attitude of a choir girl from Saint Andrew Elementary after a perfect rendition of "The Star Spangled Banner." This is not only retro but also inappropriate, considering that other old industry hands embellish with increasing vigor their database descriptions with non-existent features. Novice PR staff members pen announcements about some irrelevant "enhancements" that start with the phrase, "We are extremely excited."

Finally, H.W. Wilson got its PR approach right. At the end of September, I was heartened to read a mass e-mail that offered free access to BRD Retrospective: 1905–1982 until the end of October. It was perfectly designed; it stood out without being insolent; it had enough information to get the recipients' attention; it had quick links for additional information about the source journals covered and the most prominent reviewers, including Saul Bellow, Virginia Woolf, or Kurt Vonnegut; and it had a simple signup procedure.

Not surprisingly, the claims of H.W. Wilson were also true. Indeed, there are 300,000 books covered (OK, 299,778 as of Oct. 1, 2005), spanning almost 80 years, and indeed, "virtually every book has at least one substantial review excerpt"—my test showed 298,917 (99.7 percent). The list of sources is very impressive, ranging from the best newspapers and magazines to the best reviews-only journals, such as *Choice*, *Horn Book*, and *Booklist*.

The only point for which the publisher seems to "gild the lily" is the sample list of reviewers, which includes William Faulkner, Raymond Chandler, and Leonard Bernstein, who have only one or two reviews apiece. One might, however, consider the trade-off of quantity and quality. Faulkner is as good in his review as in his novels.

It would be useful to offer searching by review source in a field-specific index. Searching by such source titles as *Choice*, *Booklist*, and *History* in the ALL index is inefficient, as it yields too many records when these words are not the journal names. You can avoid this by using the ISSN of the review source (0009-4978 for *Choice*), but it is not an obvious step. I would love to see an unusual sort feature, a list by decreasing number of reviews, as the more sources reviewed the book, the more likely it is that the book is a worthy one. It



Bird's-eye view of quadrants in STN AnaVist

**FREE trial through October 30, 2005!**  
Access to over 1,500,000 reviews covering over 300,000 books

"A reputation built on many years of excellent service."  
*Reference Services Review*

"Welcome at any academic or large public library."  
*Reference Reviews*

"Excellent tool for obtaining reviews on...significant popular, academic, and professional titles."  
*Online*

**New!**  
**H.W. Wilson's Book Review Digest Retrospective: 1905-1982**

Use it FREE at [www.hwwilson.com/trybrd.htm](http://www.hwwilson.com/trybrd.htm)

**Book Review Digest Retrospective** provides excerpts from and citations to reviews of adult and juvenile fiction and non-fiction, plus valuable bibliographic information and book summaries. Reflecting the ideas that shaped the 20th century, coverage in this new resource comes from nearly eight decades of H.W. Wilson's acclaimed *Book Review Digest*.

Attention-getting e-mail alert about Book Review Digest Retrospective

Records: 1  
In: Book Review Digest Retro  
Link To: WilsonLink

100%  1 [Remarque, E. M. Road back](#). Little, Putnam, [1931]. 344 p.  
Review(s):  
[Atlantic Bookshelf](#) (Oct 1931)  
[Booklist](#) v. 27 (July 1931)  
[Bookman](#) (New York, N.Y.) v. 73 (Aug. 1931) [with excerpt]  
[New York Herald Tribune](#) (May 10 1931) [with excerpt]  
[The Boston Transcript](#) (May 9 1931) [with excerpt]  
[Catholic World](#) (1866) v. 133 (Aug. 1931) [with excerpt]  
[The Christian Century](#) v. 48 (July 8 1931) [with excerpt]  
[The Open Shelf](#) (Nov. 1931)  
[Forum](#) (New York, N.Y. 1896) v. 86 (July 1931) [with excerpt]  
[The Nation](#) v. 132 (May 27 1931) [with excerpt]  
[The New Republic](#) v. 67 (May 20 1931) [with excerpt]  
[New Statesman and Nation](#) v. 1 (May 16 1931) [with excerpt]  
[The New York Times](#) (Early City Edition) (May 10 1931) [with excerpt]  
[The Outlook](#) (1893) v. 158 (May 13 1931) [with excerpt]

List of review citations with excerpts in Book Review Digest Retrospective

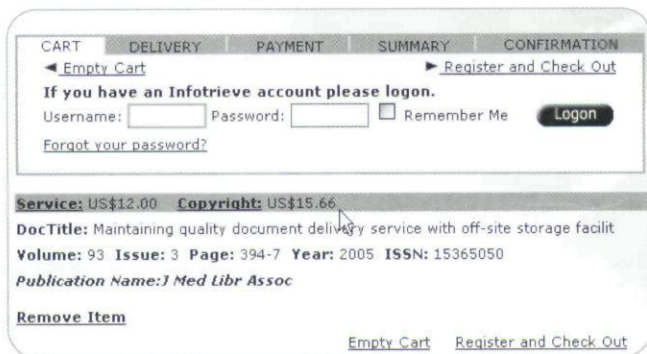
would allow finding the most reviewed books published in the 1970s, or written by Anna Quindlen, or about the chosen subject. You can get extremely excited about *this* database.



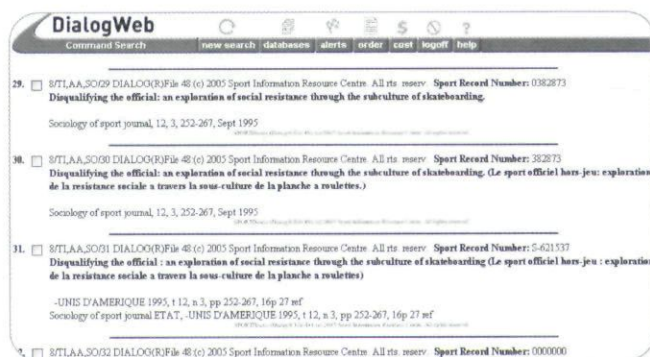
the pan

### SOME PANS OF THE PAST

It tells you something about the indifference of content providers and online services that they are not bothered to correct obvious, large-scale, and, in some cases, expensive errors—expensive that is, for the end users. I like to believe publicly exposing these deficiencies should lead to corrections,



How much an open access article costs through Infotrieve



Quintuplicate records from SportDiscus

but this is not always the case. FindArticles, for example, remains as brain damaged as I described in the July/August, 2005 issue, but its free content makes it able to get away with nonsense software "features." I wonder about the long-term prospects for this database, since parent company LookSmart will probably be de-listed from NASDAQ by the time you read this column.

In my May/June 2005 pan, I lamented the plight of the excellent, open access *Journal of Electronic Publishing*. Columbia University Press had promised to continue publishing the journal, but instead has sat on it since Spring 2002. Luckily, the University of Michigan University Library's Scholarly Publishing Office [www.press.umich.edu/jep/] has regained the publishing rights. If all goes as planned, *JEP* is to resume electronic publishing in January 2006.

Kluwer Academic Publishing merged with (or rather was acquired by) Springer Science + Business Media, so you no longer have to put up with the painfully mediocre implementation of the expensive, but powerful, Verity software for its own archive. KAP is now available through SpringerLink and its URL [www.kluweronline.com] redirects you to www.springerlink.com.

The very expensive—and very cheap—Mental Health Abstracts database continues to disgrace the worthy Dialog stable of databases. It now has a companion in shame. The pathetic SportDiscus database gathered from a variety of databases and burdened with an ill-conceived and badly implemented thesaurus was not fixed, let alone put out of its misery. Although updated several times since my review in the November/December 2004 issue (the 200601, 200602, 200603, and 200604 updates were purportedly loaded already in early October 2005), the database as implemented on Dialog retained more than 700,000 duplicates and triplicates, about 100 percent more than the Ovid or EBSCO versions of the database. Any search will show duplicates and triplicates; some will yield quadruplicates and quintuplicates. Dialog's indifference in this matter is like the negligence of the supermarket that keeps selling spoiled merchandise even after the visit and report by a health inspector. I wonder when customers will get sick (and tired) of this incompetence and careless attitude. Thus passes the glory of Dialog.

Google Scholar, which I panned in my September/October 2004 column, has improved in some regards, such as accept-

ing phrase searching in the journal name field. However, essential problems result in highly inflated hits and the omission of hundreds of thousands of important scholarly articles. May I suggest that more Google designers spend their 20 percent free-as-you-please project time, so proudly proclaimed in interviews than are parroted endlessly by awed journalists, on Google Scholar. Since it is a free service and does serve up good enough results for the have-nots, the adulation and the large-scale omissions are likely to remain—at least until Yahoo! or MSN are also given a free pass by the publishers to their archives.

Infotrieve keeps charging a \$12 service fee and varying copyright fees for hundreds of thousands of open access journal articles in its document delivery service. Take as an example the \$12 service fee and a \$15.66 copyright fee for every article in the *Journal of the Medical Library Association*, which has all articles in full text freely available through many MEDLINE gateways from Volume 1, No. 1 (under the earlier title of the *Bulletin of the MLA*). I wonder when these guys will show up in Hawaii and start charging tourists entry fees to the ocean. In the meantime, the once excellent open access PubList database (a free subset of Ulrich's Serials Directory) acquired by Infotrieve, keeps rotting away.

EconLit, which in my May/June 2004 column I criticized for dropping working papers, started to collaborate with the RePEc open access project to add them once again. Finally, some good news, as I sorely missed the working papers. The e-psyche database, panned in March/April 2004, no longer exists. It came in as a lion and went out as a lamb—and not a day too soon. The magnificent MEDLINE Plus site with the best illustrated medical encyclopedia, as I noted in November/December 2003, unfortunately still offers the ultra amateur, primitive, and ridiculously illustrated X-Plain slide shows, which really represent X-Pain. There go my efforts of constructive and illustrative criticism of databases.

*Péter Jacsó* [jaco@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.