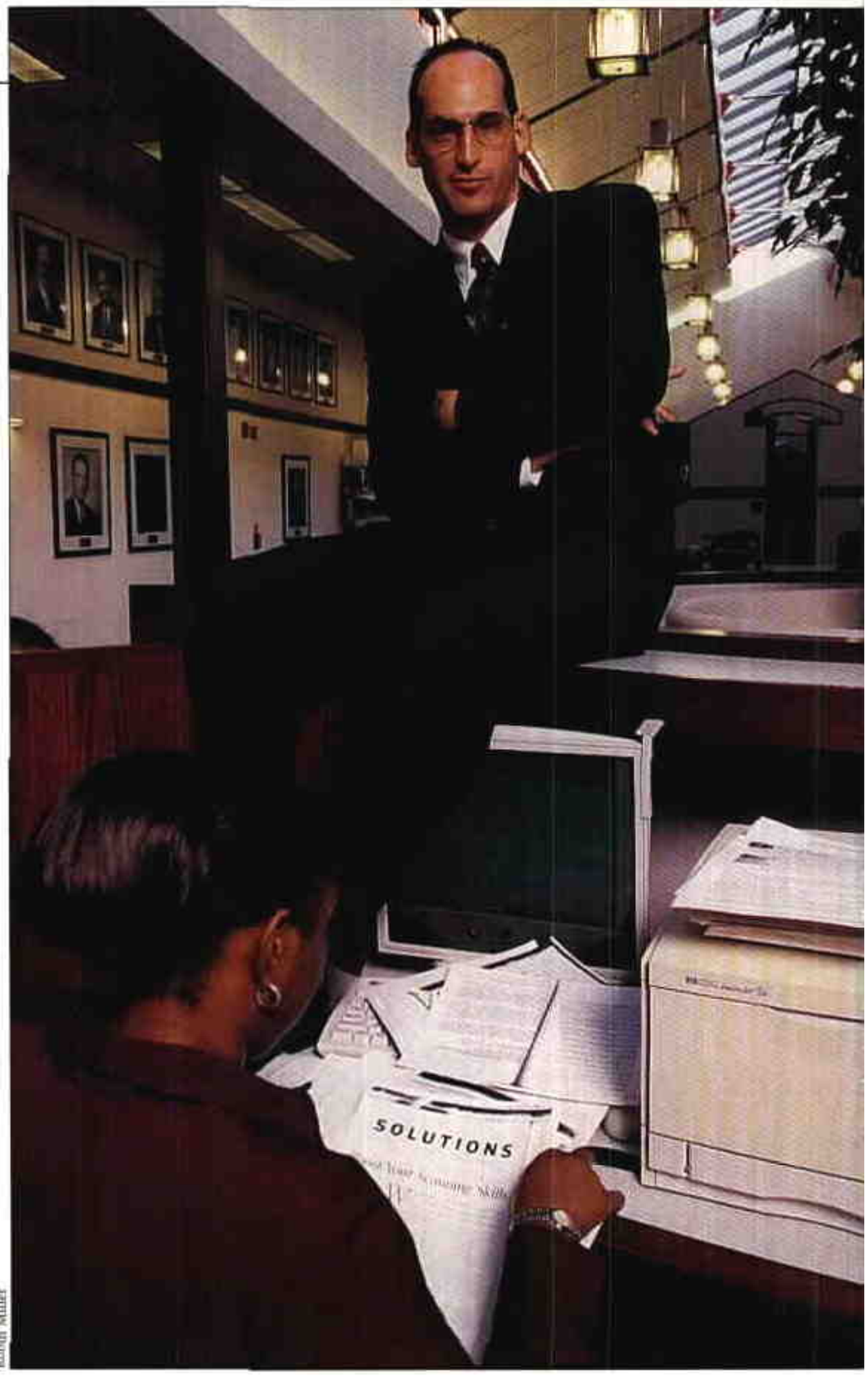
Is MORE Always BETTER?

When quality is the goal, access to everything may not be the user's best bet

by Steven J. Bell

At one time I subscribed to "the more full text, the better" philosophy. It's hard to think otherwise. Creating access to information for education and research discovery is a fundamental mission of a library, and database aggregators and e-journal publishers have heartily contributed by making vast numbers of full-text journals available in electronic format. It's certainly advantageous to offer library users instant access to a needed article, rather than pointing them in the direction of the interlibrary loan office. Lately though, as the title counts of journals in aggregator databases escalate, I've been wondering if more is indeed better when it comes to full-text journals.

Perhaps the more substantive question is what aspects of database search systems best serve our user communities, and whether the existing aggregator databases meet those needs. If they do not, then how can the library profession influence aggregators to make the necessary product modifications? Is giving our users another 100—or 1,000—more full-text journals going to improve the quality of their research? In this age of information literacy, what elements of journal search services are most critical to helping us impart sound research skills to our patrons? If we conclude that simply offering more and more full text may ultimately do more harm than good, then our professional challenge should be to work with aggregators to develop products that enable us to more readily communicate and teach information literacy skills.



Robin Miller

Steven J. Bell helps Philadelphia University student Christine Pate stay on top of her research.

Database users are not simply failing to make good use of system features. Inept search skills and low-quality search results are epidemic in proportion. In the February 2002 *Searcher* magazine, Barbara Quint observes that the problem of poor end-user searching is due to our antiquated thinking about good search skills:

"All of us as information professionals must recognize that end-user searchers who now totally dominate the user community and online marketplace will use a Google-style search strategy on any and every database they see. So if we want our databases to work, we must work around that basic reality. The days when training someone to use a database also meant introducing someone to online searching have passed forever. The training materials and curricula and philosophies developed in those days must be discarded."

Whether or not you agree that our search resources should be "Google-compliant" (read: dumbed down),

STEVEN J. BELL *directs the Philadelphia University Library.*

there is general acknowledgment that contemporary database systems tend to work against, not for, the end user. Quint is on the mark: We do our users a service by providing them with resources that enable them to find the best information with the fewest barriers to access. Despite the enormous convenience full-text databases provide, the trade-off for search systems that defeat the ultimate goal of connecting library users with the information they seek may be a price we no longer wish to pay.

Linking full-text quantity to database superiority may make good ad copy for some aggregators, but dumping yet more full-text titles into their systems exacerbates the search quality problem. It forces the end user to sift through even more irrelevant documents to find the few sought-after good articles. Better search interfaces that facilitate a user's ability to retrieve sets of concentrated, relevant information from a database of fewer full-text articles may provide one possible solution.

But even if the interfaces do improve, library patrons who refuse to even consider citations lacking full-text content are harbingers of a full-text–fixated generation of researchers who will readily pass up valuable information to their own detriment, simply because full-text content is not instantaneously available. How might we reintroduce to this generation the concept that quality research and information retrieval can precede the acquisition of content? Is it possible, as it once was, to develop systems that decouple the search and document retrieval stages, so searchers could concentrate on the former without being influenced by the latter?

Getting lost on a searching safari

Common sense and the laws of economics suggest there is a saturation point at which we will have given our users so many full-text publications in a particular subject area that adding more will yield little additional benefit. In economics this is known as the "law of diminishing marginal utility": The more of a specific product consumers obtain, the less anxious they are to get even more units of the same product. Eventually a consumer, or library patron, will become saturated with a resource, and the marginal utility derived from successive units will decline.

One of the comments I hear from faculty on my campus is that they and their students are feeling tremendously overwhelmed by the information options offered by the library. Making choices is difficult. We have all witnessed an end user struggling to decide whether to look at article 179 out of 657 and stop there, or plow ahead to examine article 180 and beyond. Granted, if the library doesn't subscribe to the hard copies of known items that library users seek, they do not care if a database has 500 or 5,000 full-text titles—as long it has the one they need. However, despite the upward spiral of electronically available titles, not all print titles are going to be available in an aggregator database or e-journal collection. There are innumerable requests that cannot be satisfied except through traditional document delivery channels.

Where is the harm if aggregators want to give us thousands more full-text titles? Since librarianship has always been based on "just-in-case" collection-building, why shouldn't aggregator databases operate on that principle too? Libraries with the capacity to do so routinely store millions of books, most of which go unused, just on the possibility that someone might engage in research for which a rarely used item is needed. Though the Pareto Principle (20% of the content meets 80% of the need) likely applies to the full-text content of most databases, having thousands of full-text titles available in electronic format means that at some time, for someone, one of those thousands of full-text titles may satisfy an information need.

We should no longer be emphasizing the sheer number of items in our collection.

But consider this: Our profession has spent a considerable amount of time rethinking how to measure the value libraries and librarians add to lifelong learning. We've concluded that we must focus on outputs, not inputs—what we do that improves our library users' ability to acquire and internalize new knowledge, and how to quantify it. One method is to focus on outcomes derived from our services. Just as we should no longer be emphasizing the sheer number of items in our collection—the inputs—neither should we be promoting the sheer number of full-text titles in aggregator databases.

What does it mean to have access to a gargantuan number of titles? Can we measure what impact all this full text has on the quality of research? Perhaps with more limited, yet concentrated, aggregator journal holdings we might be able to identify techniques that allow us to truly assess, through quantitative measures, the relationship between full-text journals and research output and quality.

The contemporary aggregator business model may be unable to accommodate the concept of reducing full-text content. For years the model, with few exceptions, has been based on "who has the most full text." To some extent, this has shifted the database selection process away from more traditional measures, such as quality of abstracting, indexing, and search-screen interfaces, to one in which the primary (and sometimes only) criterion is how many full-text titles are available. In some cases significant amounts of full text turn out to be mostly backfile with restricted access to current issues owing to embargoes, which fails a library's user population; it's wrong to provide several years of older content without equal access to the latest six or 12 months, just so officials can exult higher counts of full-text journals in the annual report. Besides, while aggregator database subscription costs increase more slowly than journal subscriptions, and to some extent are stable, adding more full-text titles must add to the price of the product.

Yet we, the buyers, have no say in what that additional content is. So we pay for added titles without really knowing if they add value to our collections. Price increases would be better justified if they were accompanied by significant improvements in the search systems and interfaces themselves.

It's worth making a distinction between full text in electronic-journal collections and those in aggregator databases. The former rarely enforce embargoes because the journal publisher is the direct provider; when the aggregator serves as intermediary, it must conform to publisher demands before it can make full text available. Full-text electronic journal collections also present less search-and-retrieval challenges for the end user. The electronic journal collections, because they can be acquired with subject selectivity and offer stable access, often serve to supplement and strengthen the library's print journal collection. But full-text titles in aggregator databases cannot be depended upon to serve as a collection-building device owing to the lack of content stability (the constant dropping and adding of titles), and the fact that the titles represented may not actually contain the full text of the journal (missing graphics, tables, no cover-to-cover inclusion, and so on).

Librarians might better appreciate databases that provide the core titles in a subject field.

Rather than continuing to add titles of questionable or marginal utility, aggregators could concentrate their efforts in reducing the underrepresentation of certain disciplines in full-text databases. There are significant gaps in full-text titles available online in art, architecture, a variety of design fields, and other humanities, social science, and even science, technology, and medical niche areas. Put another way, we do our users a disservice by providing a database with hundreds of trade journals that provide dozens of articles on the same topic, but that ultimately fails to provide substantive information for critical analysis. Instead, we should be providing discipline-balanced aggregator databases, of which some disciplines and specialties lag far behind. But that lack is hardly something to blame on aggregators. The cooperation of publishers will be needed to eradicate the current imbalances.

Tweaking the status quo

Is the answer a return to a library landscape dominated by indexing and abstracting services? A change of that magnitude is unrealistic and counterproductive—even though there is some possibility it could help students achieve better research skills. Unarguably, databases with full-text materials make it convenient for library users to find what they need. When properly searched, full-text aggregator databases can lead to the answers to questions that would be impossible to find with indexing and abstracting alone.

The concept, design, and execution of the full-text database work fine. What needs to change is the current aggregators' race to add more and more full-text titles. This business practice hinders more than it helps. The type of aggregator database librarians might better appreciate is one that concentrates on providing the core titles in a subject field, that adheres to a roughly set number of titles so there is no expectation of random addition to the database, and that aims for a high ratio of full text to available titles.

What the library community needs to do is examine the impact of the full-text obsession—both the growing dependency within the end-user community and the aggregators' fueling of this phenomenon. What we need to lobby for instead are "information literacy–friendly" systems and interfaces that offer:

- High concentrations of full text from limited numbers of core publications that are representative of a broader range of disciplines;
- Less dependence on "omni" searching, and more use of subject-specific modules that allow users to compartmentalize searching (some systems offer something similar now but cater to users' willingness to choose "all" rather than taking the time and thought necessary to pick the appropriate module);
- Interfaces and features that promote high-quality searching. Some systems already feature these options (field searching, peer review limit) but they need to become better integrated and more intuitive;
- Built-in control factors that force users to revise their search if the results would exceed some pre-set limit, at which point the search will not be conducted;
- Fields other than "full text" that can be assigned as the default search;
- Local control over the amount of full-text content that is available in the database, down to the title level.

Of course, no one in the aggregated-database industry would want to be the first to blink in the race to amass larger quantities of full-text titles. It is difficult to imagine any of the major aggregators boldly announcing a change in strategy to the limited addition, or possibly even the halt, of new additions while the firm reexamines its existing source titles with the goal of developing a more highly concentrated core in each discipline. It would certainly take courage, and require considerable input from user advisory boards and subject experts.

Our future success as a profession will be driven by our ability to educate and connect with patrons as their information advisers, not as gatekeepers to massive collections of aggregated full-text titles. We desperately need aggregators to serve as our partners in helping us succeed. We must work together in the pursuit of a mutually beneficial solution that achieves the goal of reversing these two troublesome trends, namely, poor quality research and full-text fixation.

Aggregators, to their credit, are good listeners. Let them know what you think. ❖