Against the Grain

Volume 22 | Issue 3 Article 30

June 2010

Random Ramblings -- Bigger is Not Necessarily Better

Bob Holley
Wayne State University, aa3805@wayne.edu

Follow this and additional works at: https://docs.lib.purdue.edu/atg
Part of the Library and Information Science Commons

Recommended Citation
Holley, Bob (2010) "Random Ramblings -- Bigger is Not Necessarily Better," *Against the Grain*: Vol. 22: Iss. 3, Article 30.
DOI: https://doi.org/10.7771/2380-176X.5569

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Biz of Acq
from page 54

Byte 181. It sounds like the title of something **Isaac Asimov** or **Thomas Disch** might have written, or perhaps the location of some electronic doomsday scenario. I expect there are billions of Byte 181s out there doing good work — allowing doughnuts to be sugared, tires to be treaded, roses to be planted. And now we had our own Byte 181, but it was not doing anything helpful and productive like sugaring, treading, or planting.

Our Byte 181 is a number two ("2") and lives at the nexus of the ILS/Oracle divide telling Oracle that unequivocally, without question, forevermore, the directive from the script in which Byte 181 lives guides Oracle to output the second ("2") vendor address. The problem is the code is absolute — the script tells Oracle to ignore the fact that the addresses may be tagged as active or inactive and, regardless, always plug in the second vendor address.

When new addresses are entered into the Oracle file, they get added chronologically in a list, and none are deleted (for auditing reasons), with the result that what we have in our ILS and what resides in Oracle can be, and usually are, completely different. The second address in our ILS, the one we want to use, might be address number 19 in Oracle, and yet the script points inexorably to number 2 — "Take that one," it says, which might be hopelessly out-of-date or might be a correspondence address. So, in the script, Byte 181 tells Oracle to skip lightly over everything else and print address two from its table, and voila! — the birth of our vendor address problem.

For me, working in Technical Services, the investigation into the problem with Oracle is emblematic of everything we currently are doing; we have workflows that suffer from serious constraints, and we have to examine each one to determine how we can streamline, remove, or replace the constraint and make the workflow more understandable, transparent and manageable. However, unlike Byte 181, residing happily in a binary world and performing the same logical, albeit frustrating, thing every time, the bad stuff in Technical Services does not always happen for the same reason, with the same predictable results.

Byte 181 is the exemplar for those nagging little problems where workflows intersect, the place where communication breaks down, where there is no resident expertise to know how to fix things requiring countless meetings with ever-changing players. Byte 181 is our shorthand for the process of teasing out the part of a procedure that bogs down throughput.

It was now early in 2010. Things started working. Checks were getting printed, vendors were getting paid, glitches were being reported, a new collaborative, cross-disciplinary group was primed and ready to notify IT when and if things went awry. You would think that we would be happy, that we would find our laurels, wherever they were, and rest on them. But we became aware of something.

Oracle just went through an upgrade. Stay tuned.

Random Ramblings — Bigger Is Not Necessarily Better

Column Editor: **Bob Holley** (Professor, Library & Information Science Program, Wayne State University, Detroit, MI 48202; Phone: 313-577-4021; Fax: 313-577-7563) <aa3805@wayne.edu>

My favorite public library ever was the **Clifton Branch** in Cincinnati. It was the summer of 1967, and I had just graduated from college. After a difficult year with a full-time job and a full classload, working 40 hours per week on a summer job seemed liked vacation. With time to catch up on my reading, I made frequent visits to replenish my stock of books. The **Clifton Branch** had only one room with a very limited selection. But this selection was perfect, since the branch served mostly the members of the nearby university community. Except in the children's area, I could have selected my books blindfolded and would have been happy to read around 80% of my random selection. I'll now fast forward to a few years later when I was a student in library school at **Columbia University**. The professor proposed to the class that having one unified list of all the serials in the world would eliminate the need for other lists with its universal coverage. I raised my hand to disagree and made the point that smaller libraries could easily make do with a specialized list more tailored to their interests. I argued that a small public or school library would have no interest in scholarly resources or foreign language materials. I also pointed out that the comprehensive list would be too expensive to purchase in print format and would require frequent revisions. (Such a list would make more sense today in a digital format.)

I believe that most users would like to have all needed items together in one physical or digital space with as few as possible extraneous materials to complicate finding what they want. This is why most of us have personal collections. This is also why most faculty like to have departmental libraries. I still remember the faculty member who couldn't understand why the book on ceramics in Vermont was in the art section (LC N), while the book on ceramics in Pennsylvania was in the science library (LC T). She had looked at both books and found them quite similar even if the catalogers had determined that one was over 50% art and the other over 50% technology. She would have much preferred an art departmental library where both books would have been within easy reach rather than in far distant locations from each other in two different libraries.

Many research universities have an undergraduate library for somewhat different reasons. The first is to save undergraduate students the time needed to navigate the complex research library, since the simpler undergraduate library contains most materials that they need for their assignments and facilitates effective browsing. The library can also provide services including reference tailored for this student population. A second reason is that undergraduates may not yet have sufficient information-seeking skills to understand that a research library includes source materials that represent all positions, including those in scholarly disrepute. Having the undergraduate library helps protect the sophomore from citing Klu Klux Klan propaganda in a research paper on race relations in the United States.

The digital era makes vast quantities of materials theoretically available but practically inaccessible. Most information professionals understand this concept in regards to search engines. It is impossible to look at result number 5,023 even if the user were willing to scroll through all the screens to get there. (In one test, **Google** stopped providing results after around 300 entries.) The search algorithms that put popular materials at the top may push scholarly materials to the bottom of the result stack.

I am not sure that information professionals realize that the materials that libraries offer to their users can pose the same problem of too much rather than too little. To return to the pre-digital age, major microform sets often went unused because researchers didn't know what they contained without using print finding aids. Even worse, the researcher doing a general search might not even be aware that the library owned materials in this format. I know of one faculty member who was contemplating a trip to a distant university to consult a rare item before the reference librarian at the other institution told him that the item had been filmed and was available at his home institution in a major microform set. The pre-Internet solution to this problem was a major effort from around 1980-1993, supported in part by grant funding, to catalog major microform sets and to make the records available from **OCLC** for batch loading. The sheer volume of Internet resources and their mutability make this level of bibliographic control impossible.

Search rules for large library databases can complicate access and show that more is not always better. I once needed to find a known item in **OCLC WorldCat** with a one-word title that was a common word. Since I didn't have any other bibliographic information, I typed the one word in the title search box. The search algorithm defaulted to a keyword search that retrieved thousands of items in no useful order. The reference librarians that I consulted didn't know how to solve this problem. A call to the **OCLC** help desk didn't provide an answer either. Only a year or so later, when

continued on page 56

Random Ramblings
from page 55

I spoke to an expert from **OCLC**, did I learn the proper procedures. She emailed me the rather complicated steps, which I most likely have stored somewhere but am not certain that I could ever find again.

I've already written a short article in favor of the **Google Books Project** since having all the books in the world accessible is a laudable goal. I have not, however, in my reading seen any discussion of the potential problems that opening up the floodgates of availability might bring. "The Public Access Service license will allow free, full-text, online viewing of millions of out-of-print books at designated computers at U.S. public libraries." (*http://books.google.com/googlebooks/agreement/faq.html*) From the **Google** terminal, the patrons of the smallest public library with a few thousand books will face some of the same access problems as those who use the world's largest research libraries.

What problems will these users face? First, patrons will need to learn more effective search strategies. Many will enter search terms that bring up thousands of records. The **Google** search algorithm may bring to the top of the list the books that would most interest them, but then again it may not. Some will be overwhelmed at the number of possibilities when they would have been less frustrated with a more limited number of options. Choosing breakfast cereal in a convenience store is much easier than in a mega supermarket.

Second, the rules for searching and displaying results are not clear. I pretended to be an untrained user and searched for "Mars" to see how **Google Books** would handle this ambiguous search. The **Google** results page told me that I had 173,478 hits but returned only around 190 books before **Google Books** stopped providing results. All the suggested refinements at the bottom of the first page of results referred to the planet. Searching "planet Mars," "God Mars," and "candy Mars" all had fewer hits; but **Google** showed more results before cutting off access. Finally, the French word for the month of March ("mars mois") returned the most available results of any search — around 400 books. If I'm confused as a trained librarian, think what will happen for the average user who wants books on Mars, the Roman God. I believe that readers can guess what happens when a teenager looks in **Google Books** for items on the singer "**Sade**."

The third issue is the question of reliable and useful information. Small-to-medium public and academic libraries choose the most useful items for their user community as the **Clinton Branch Library** did for me. These patrons are not interested in esoteric scholarly materials that will become an increasingly important part of **Google Books** as **Google** staff scan the collections of major research libraries. The problem may be even worse if the **Google Books Settlement Agreement** is not approved, because full-text availability will be more common for out-of-copyright materials that are older and less useful for most patrons of smaller libraries. The 1910 book on child rearing certainly won't help today's parent very much. As I said earlier about undergraduate research, the patron may also access primary sources that large libraries collect for research but that require sophisticated evaluation skills and background knowledge beyond the competencies of some small library users.

To conclude, I am convinced that one reason why libraries and librarians will survive is that they help people find the right needles in the massive information haystacks on the Internet. Before the arrival of the Internet, the problem was often too little information. Now the problem is too much information. I'm not sure that individual librarians and the profession have adjusted completely to this mind shift. Pathfinders, bibliographies, and reference sessions may retain their importance, not to find needed materials, but to screen out the garbage in an information universe where bigger is not necessarily better.

From the University Presses — What University Presses Think about Open Access

Column Editor: **Sanford G. Thatcher** (Director, Penn State Press, USB 1, Suite C, 820 N. University Drive, University Park, PA 16802-1003; Phone: 814-865-1327; Fax: 814-863-1408) <sgt3@psu.edu> *www.psupress.org*

Column Editor's Note: *This article is based on a talk prepared for the* ***Open Access Symposium*** *at the* ***University of North Texas*** *on May 18, 2010:* ***http://openaccess.unt.edu/symposium****. —* ***SGT***

I envy a commercial publisher like **Elsevier**. Its mission can be very simply defined: make enough money to pay your employees and keep your stockholders happy. Whether **Elsevier** were in the business of making widgets or publishing books and journals, that mission would remain the same. The means to achieve that end can be very complex, but the mission itself is simple and straightforward.

Not so the mission of a university press like the one that employs me. It straddles two worlds, academic and commercial, which each have imperatives unique to them that are often in tension if not outright conflict. On the one hand, and above all, a university press's mission is defined by the imperative that drives academe as a whole: create new knowledge and communicate it to the next generation of students and scholars. On the other hand, every university press must make enough money to stay viable as a commercial enterprise operating in the same business environment as any other publisher. A few can do so without the help of their parent universities; the vast majority cannot and need to be subsidized at some level (on average, 10% of their operating budget).

How these two imperatives are balanced differs from press to press, depending on pressures both from the university's administration and from the commercial marketplace. Some presses like my former employer **Princeton** have the advantage of being semi-autonomous: it is separately incorporated in the State of New Jersey, but the use of its name is controlled by a faculty editorial board and a board of trustees on which a number of university administrators sit. It receives no financial support from the university at all but fortunately has a handsome endowment, which derives from the astute management of the **Bollingen Series** taken over from **Pantheon** in the late 1960s accompanied by funds from **Paul Mellon** to see through publication of the remaining volumes, some of which (like the translation of the *I Ching* and books by **Joseph Campbell**) have been huge commercial successes. A few of the very largest presses, like **Cambridge** and **Chicago**, are obliged to turn over a portion of their earnings to their parent universities and thereby subsidize those universities in small part. At least one smaller press, **Rockefeller**, is also similarly obliged. Much more typical is the press at **Penn State**, which after more than a decade with no operating subsidy now has a subsidy at the level of the 10% average I mentioned above. Depending on how close to the margin any press operates, you may find one press feeling it necessary to raise prices on its books to satisfy the commercial imperative, while another press may feel it can afford to prioritize its goal of maximizing dissemination of its books by keeping their prices low and making them available as soon as possible in cheaper paperback editions. (Some presses, like ours, cross-subsidize between journal and book operations, the former's surpluses used to offset the latter's losses.) Overall, because of this disparity in missions between commercial academic publishers and university presses, independent studies of pricing of books have routinely showed university press titles to be priced lower, sometimes much lower, than those from commercial publishers. In this way, too, some university presses are consciously subsidizing academe in general, if not just their own universities.

Those who, like **David Shulenburger**, have been critical of the positions that university press-

continued on page 57