LIBRARY JOURNAL.com
Visit *LJ* on the web!
www.libraryjournal.com

Remaking an Online System

ONLINE COMPANIES continually tinker with their systems. Commercial aggregators need to accommodate new databases or new features such as linking. Search engines must change ranking algorithms to subvert rogue web sites or hackers who seek to manipulate ranking. Systems need to improve performance and keep up with the competition or the current fashions in interface design.

It isn't often that a system can afford to do a total makeover, including architecture, new programming code, and new functionality. Factiva did this over the last few years; recently the H.W. Wilson Company has totally revamped its system. Ed Tallent, Boston College, reviewed the new WilsonWeb (Database&Disc Reviews, *LJ* 1/03, p. 173ff.), focusing mainly on the search features and interface. But a total redesign goes beyond what the user sees and searches. It is also a redesign of the underlying structure and methods by which Wilson indexers and editors work.

Database creation

Lucian Parziale, Wilson's vice president for information systems, said that the impetus was to handle more full-text databases better and to develop a new production system. To accomplish this, the new system uses Oracle software and includes a new editorial system (WIN), a new way to store files, and the new WilsonWeb search and retrieval system with Verity search software. The new system represents an investment of over $10 million.

The new system begins with database creation. The Wilson abstracting and indexing (A&I) and full-text products require various processes, such as journal issue check-in, A&I, and editing (and, now, full-text preparation). Traditionally, this involved shuffling paper journals and creating electronic records for sequential processes. In WIN, journal articles are scanned and converted to PDF files when an issue is checked in, so the indexers and abstracters can work at their desktops at the same time the full-text version is readied. Since Wilson indexers are in New York and the abstracters are in Dublin, Ireland, this simultaneous processing is a boon to production.

The new system represents an investment of over $10 million

In Wilson's new repository system, the content is organized by type of material—bibliographic, biographical, and monographic—not by the corresponding product as it was previously. Most of the Wilson bibliographic databases come in a variety of products (including indexing only; indexing and abstracting; and indexing, abstracting, and full texts). Now there is just one master database for each product family rather than separate databases. Through authentication, customers view only the parts for which their library has paid. This less redundant underlying organization replaces 67 separate product files with 39.

WilsonWeb

"WilsonWeb Next Generation" is the new system that users see. It was designed with input from focus groups, from which a major lesson emerged—no one system can make every user happy. In response, Wilson created customizable administrative functions. Librarians can customize the interface and search options for their customers and subgroups of customers. Buttons, names of buttons, default search screens, access options, and more can all be customized. Libraries can let all the records in a database be searched or limit searches to only those titles held by the library.

This means that WilsonWeb may look quite different in each library. One library version may have a logo on every screen, another may default displays to only the title and author for the first ten records, still another may display buttons for all the advanced search features.

The Verity search engine is a major change in the WilsonWeb system because it allows natural-language searching and provides relevance-ranked results. Experienced searchers can use Boolean logic, proximity operators, and field-restricted searching. Also, in a first for Wilson, in the full-text version of a database the complete articles are searchable. Still, Wilson's controlled vocabulary remains a major precision tool and advantage. The controlled subject terms are not only searchable, but the Wilson thesauri are invoked automatically in searching, so an incorrect term is mapped to the correct term. Term lists can be browsed and selected, plus broader, narrower, and related terms are displayed. Matches in any of Wilson's name and subject authority files produce results that have a high relevancy ranking.

Full text and linking

Full text was a major impetus for these changes, according to Parziale: "The market called for more full text." WilsonWeb full-text databases provide PDF versions of articles (the same format generated at the beginning of the process), but Wilson honchos also recognize the need to allow linking to the full texts found on other systems. WilsonLink, another component of the new system, uses SFX software to link to full texts within Wilson and other OpenURL databases and systems.

Running an online service requires a commitment to keeping up with developments. Sometimes it requires a total revamping. Since the upgrade's unveiling last year, Wilson has already made improvements in response to user and reviewer input. An online system is never completely finished.

Carol Tenopir
(ctenopir@utk.edu)
is Professor at the
School of Information
Sciences, University
of Tennessee,
Knoxville