ONLINE DATABASES

BY CAROL TENOPIR

Unlikely Partnerships

THIS MAY, TWO NEW partnerships were the talk of the National Online Meeting in New York. Dow Jones Interactive announced a merger with Reuters Business Briefing, and Northern Light formed a partnership with the Commerce Department's National Technical Information Service (NTIS) to produce usgovsearch, a "one-stop federal government information resource."

In the first case, this brought together competitors; in the second, government joined with the private sector. Although quite different, both partnerships create unlikely—and unexpected—relationships. While they provide convenient mergers of unique content, they also generate a fair amount of controversy.

Dow Jones and Reuters

Many business online searchers in the United States are regular users of Dow Jones Interactive (djinteractive.com), while their European counterparts rely on Reuters Business Briefing (www.reuters.com). Surprisingly, there is little duplication of both customers and content between the two services; Dow Jones Interactive gets over 95% of its revenue from the Americas and Reuters gets 80% of its revenues from Europe and Asia. The merger of these two rivals was unexpected, and probably unwelcomed by their competitors, but will make global business news information more accessible for business customers. Pricing remains uncertain.

The new Dow Jones Reuters Business Interactive (DJRBI) will be a service for business professionals and professional searchers, particularly in corporations. It is not aimed at consumers; indeed, consumer publications in the component databases such as the *Wall Street Journal* (*WSJ*) Interactive Edition or real-time financial newswires from both Dow Jones and Reuters are not in the merged service.

Over the next year, Reuters and Dow Jones business content will be combined, so searchers can get access to business news sources from North America, Europe, and Asia all in one place. The focus will be on desktop, web-based delivery, but Dow Jones promises that its Windows version for expert searchers will be supported for at least another year as the new service is developed.

Eventually, DJRBI will include the archives to *WSJ*, Dow Jones Newswire, Reuters news wires, and over 7000 international business news and information sources. Interfaces will be available in at least 11 languages, with content in 20 languages. Pricing will probably follow the Dow Jones model of subscription-based, enterprisewide pricing, perhaps with document charges, as opposed to the Reuters pay-as-you-go model.

Competitive edge?

DJRBI's main competition will continue to come from market leaders Lexis-Nexis and Dialog. Both Dialog and Lexis-Nexis are already international in scope; Dialog's DataStar service focuses on European business information. Dow Jones is a solid number three in business information, and this merger should increase the competition. Already, Dialog has issued a press release that pooh-poohed the value of news-only services since "news is increasingly a commodity, available for free on numerous Internet sites."

Reuters replied by pointing out that Dialog is a mere aggregator of other company's sources, while Reuters and Dow Jones create valuable branded content. All of the companies may overstate their cases, but the brands of Reuters and Dow Jones do carry some weight. Except among information professionals, Dialog lacks a recognizable brand name.

While search engines may be even more competitive than business information companies, they are generally outside the traditional online arena. But Northern Light combines traditional online with a web search engine.

Web search engines canvass the web for the searcher; traditional online systems bring access to proprietary research content, often through controlled vocabulary indexing. Although almost all-commercial online services have fee-based web versions, they require different search strategies and different indexes to find specialized content. Northern Light's innovation: combine web access with access to fee-based and proprietary content and also combine relevance ranking with controlled subject searching. This bridges the gap between the web and online searching.

Northern Light Technology (www.nlsearch.com) started as a division of Dataware, became a separate company in 1995, and launched service in August 1997. The web search engine marketplace was already crowded, so Northern Light attempted to differentiate itself by targeting information professionals and emphasizing research. Calling itself "the world's first Internet research engine"—a claim that would be hard to prove—Northern Light has about a dozen staffers with MLS degrees focusing on classification and indexing as well as research content.

Northern Light's innovation

Northern Light offers 1) a web search engine that crawls, indexes, and classifies web pages, and 2) a Special Collections service that provides indexing and full-text articles from journals and magazines. (The new usgovsearch provides access to both, plus government sources.)

As of January 1999, Northern Light claimed to be the largest free web search engine, with over 120 million web pages indexed. (Alta Vista is the nearest competitor with over 100 million.) Northern Light seems to be growing steadily, adding over seven million web pages in less than a month in January, while some rivals—such as Hot Bot, Lycos, or Excite—have slowed down.

Northern Light's web search engine offers some unique value-added features



Carol Tenopir is Professor at the School of Library and Information Science, University of Tennessee at Knoxville. Her E-mail address is tenopir@utkux.utk.edu

that appeal to information professionals and makes it more like traditional online systems. In addition to the default statistical search engine, Northern Light offers Boolean AND, NOT, and OR, nested Boolean, phrase searching, and truncation. Field searching is offered in searching special collections documents but can also be used to restrict a web search to text found within a URL. Searches may be restricted by date, language, country, subject, or document type.

Limits on searching

There are some limits, however. Although exact phrases can be searched, more sophisticated word proximity features are not offered. For example, searchers can enter the narrow phrase "online catalogs" or increase recall drastically with the Boolean phrase "online AND catalogs," but they cannot specify that the word "online" occurs within four or fewer words to retrieve phrases such as "online public access catalogs" or "online library catalogs."

When a searcher inputs a query, each term or phrase is matched with a vocabulary of approximately 15,000 controlled subject terms, 30 document types, or geographic terms. Retrieved results are mapped to those controlled terms and relevance ratings are assigned based on how closely each retrieved document or site matches the controlled terms. Results are displayed in relevance ranked order, as is customary, but the addition of query and document word matching to a controlled vocabulary should improve precision scores.

Notably, Northern Light produces a more focused search with Custom Search Folders. These automatically classify each site in a retrieved set into folders that represent different subject aspects or document types. The software analyzes words in the retrieved documents, matches words with a proprietary thesaurus, and places them into categories. For a search on "parental leave," for example, custom search folders included family leave, arbitration, severance pay, affirmative action, and several more.

NL's special collections

Besides crawling the web, Northern Light provides access to a growing collection of fee-based materials, thereby competing directly with services such as Dialog, Lexis-Nexis, and Dow Jones for the professional searcher's dollar.

In an agreement with Bell & Howell's ProQuest Direct, Northern Light provides access to full-text articles from over 5400 journals and magazines—though most major newspapers are missing. The Special Collection is strong in trade and industry publications, local newspapers, and business-related topics. Of special interest to business searchers, Northern Light also offers access to Investext, Find/SVP, and Market Research Reports. Indexing of the full articles is done automatically.

> Public criticism about fees led Northern Light to offer usgovsearch free to schools and libraries

Northern Light's unique and low-cost pricing structure has dismayed competitors and secondary services. Searchers access the indexing for free from the Northern Light search service. A search can be restricted to just web sites, to just the Special Collection, or to a combined search of both, all for no cost. Combined search results are separated at the folder level.

Charges are introduced only when full articles are selected from the output list. Most journal articles cost only $1–$3, far less than searching, retrieving, and downloading or printing the same articles from Dialog or Lexis-Nexis.

Usgovsearch debuts

With the May 1999 debut of its usgovsearch service (www.usgovsearch.com), which offers access to NTIS indexing and a web search engine that reaches thousands of government web sites, Northern Light both gained attention and criticism.

Usgovsearch indexes approximately 20,000 government and military web sites, representing over 3.8 million web pages. It includes the NTIS bibliographic database from 1964 to the present with full text of journal and magazine articles already present in Northern Light's Special Collection. Government journals in the Special Collection include *Defense Daily* and *FedNet Government News* mixed with the 5400 other titles. NTIS technical documents can be ordered online through usgovsearch, but they will be delivered on paper or microfiche. According to Northern Light, only about 25% of government sites are available through other search engines. NTIS and Northern Light aim for 99.9%.

Northern Light indexes federal web sites by automatically crawling all of the .gov or .mil sites. In addition, NTIS has supplied Northern Light with a list of other government-based sites, including some .com sites not usually found by web search engines. Although these sites are available for free through other search engines, usgovsearch brings all U.S. government sites together in one search engine.

Controversy erupts

Controversy erupted as soon as the joint NTIS–Northern Light announcement was made. Responding to criticism from some advocacy groups that government information on the web should be free, the U.S. Department of Commerce announced that the new product was being put on hold. According to Northern Light, however, service was never suspended.

Though for several weeks the service was free, on June 14, Northern Light began charging subscriptions for individuals or companies. Initially, individual pricing was $15 for a day pass, $30 per month, or $250 for an annual pass. However, according to company spokeswoman Susan Stearns, the criticism had some effect: NL reduced the day pass fee from $15 to $5 per day and provided the service free to schools and public libraries "to help NTIS and the government deflect the public outcry."

Enterprise accounts for organizations and corporations are charged $1000 per year for up to five simultaneous users and on a sliding scale for additional simultaneous users. The subscription fee includes access to all of the Northern Light indexing and web services, but there are additional charges for full-text delivery.

Stearns said, "The joint venture between NTIS and Northern Light remains in place. NTIS asked Northern Light to take the lead in the project. NTIS has reduced its visibility. NTIS asked to remove the link from its web site to usgovsearch and to remove their logo on the site. NTIS has not pulled out. They have only asked Northern Light to modify the physical appearance of the site."