



# Smart Libraries™

Formerly Library Systems Newsletter™

50 East Huron Street, Chicago, Illinois 60611-2795, USA

February 2005 Volume XXV Number 2

## Webfeat Patent Roils Metasearch Market

With ever-growing collections of electronic resources, libraries desire technologies that simplify the search process. One of the hottest technologies around today involves applications that provide a single interface to search multiple resources. Called by various terms—metasearch, federated search, or broadcast search—this type of technology helps libraries and other organizations provide searchers with easy-to-use interfaces so they can search many different electronic resources simultaneously.

Thus sales of metasearch applications, link resolvers, and related products have grown to become a significant part of the library automation econ-

omy. Because almost all of the library automation companies now offer a metasearch product, libraries can choose from a number of them. But at least two companies—Webfeat and MuseGlobal—develop metasearch technologies as their core business.

The technology developed by these two companies underlies the metasearch products offered by many library automation vendors. Dynix, The Library Corporation, EOS International, and Follett Software Company each integrate technology from Webfeat. Those using Muse Global include Sirsi, Innovative

*See Webfeat on page 3*

## Google Signs on to Digitize Scholarly Books

In mid-December 2004 Google announced a series of agreements for major digitization projects involving five research libraries, including Harvard, Stanford, Michigan, Oxford in the United Kingdom, and the New York Public Library (NYPL). Over the next decade, millions of scholarly books will be scanned and indexed by Google. All of the books to be scanned through this project will be fully searchable. Books in the public domain will be fully viewable online, free of charge, by Web users.

It is not yet clear if this large collection of scholarly e-books will be searchable as a separate Google collection, or if the indexing will be blended into Google's overall index to the Web. It also is unclear if the books will be presented cover to cover, or if only the pages pertinent to a user's search will be presented.

Google sees this as a major expansion of their existing Google Print program. Google Scholar, a related program launched earlier in 2004, attempts to capture and index scholarly information already in digital form on the Web.

*See Google on page 4*

### IN THIS ISSUE

Webfeat Patent Roils Metasearch Market  
PAGE 1

Google Signs on to Digitize Scholarly Books  
PAGE 1

Firefox Blazing Along  
PAGE 2

Bob Walton Stops Song on Dynix Seat  
PAGE 5

Libraries Up Use of Webinar Software  
PAGE 6

Microsoft Summit Focuses on Info Access  
PAGE 6

RFID Gets Tagged  
PAGE 7



#### Receive *Smart Libraries* via e-mail

Subscribers who would like an e-mailed version of the newsletter each month should forward their e-mail address and ALA identifier (the 7-digit number printed on the top line of the address label that appears on page 8 of your newsletter) to jfoley@ala.org. Type "e-mail my Smart Libraries" into the subject line. Issues will be e-mailed in addition to your print subscription and at no additional charge.

ISSN 1541-8820

## Firefox Blazing Along

With about 95 percent of the market, Microsoft's Internet Explorer (IE) has enjoyed its longtime status as the most widely used Web browser. Though Microsoft has worked hard to ensure IE's security through a constant stream of security patches and updates, IE's widespread use, as well as its intrinsic security vulnerabilities, has made it a constant target for criticism.

Microsoft's browser also has a reputation for not always following the latest standards for HTML, XHTML, and CSS. In its favor, IE comes tightly integrated with the Windows operating system, allowing fast and easy connections between applications and the Web. But this tight integration also has worked against it, since vulnerabilities in the browser thus can lead to compromising the whole computer system.

For those looking for an alternative to IE, Firefox 1.0 is here. Since its release on November 9, 2004, this browser, built as Open Source Software by the Mozilla Foundation, already has attracted about five percent of all Web users. By mid-December 2004, Firefox had been downloaded and installed on more than ten million computers.

While IE suffers from a bad case of features-and-functions bloat, Firefox takes the quick and nimble approach. Its browser-only approach results in a small application that's fast to download and easy to install. But most favorable of all, it renders Web pages quickly. It has a tabbed display that allows a user to open up new pages behind a tab and thus toggle between pages handily. Yet the differences between IE and Firefox lie in nuances—the layout of icons, menus, and windows are so similar, those familiar with IE will have no trouble using Firefox.

### Web Site Reworks?

In addition, Firefox shows a high degree of compliance with Web standards. Unfortunately, many Web sites have been specifically tuned for IE's idiosyncrasies, with pages coded in nonstandard ways. The domination of IE has allowed Web site designers to focus on how the pages look in that browser



rather than ensure they adhere to the standards. As Firefox and other browsers gather more followers, those who manage Web sites may need to rework pages to ensure compliance with standards.

The release of Firefox brings up two issues for libraries: Should they migrate their public and staff computers to Firefox? And are all library Web pages up to snuff for those that choose to use Firefox instead of IE? Many operational, logistical, and technical considerations will influence those decisions differently for each library. As both users and providers of Web-based resources, libraries cannot ignore this significant change in browser demographics.—*MB*

Contact: Firefox/Mozilla:

[www.mozilla.org/products/firefox/  
tabbed-browsing.html](http://www.mozilla.org/products/firefox/tabbed-browsing.html)

Desktop Pipeline article by Scot Finnie:

[www.desktoppipeline.com/53700233](http://www.desktoppipeline.com/53700233)

Washington Post article by Rob Pegoraro:

[www.washingtonpost.com/wp-dyn/articles/  
A47146-2004Nov13.html](http://www.washingtonpost.com/wp-dyn/articles/A47146-2004Nov13.html)

Webfeat from page 1

Interfaces, VTLS, Mandarin Library Automation, and Endeavor Information Systems. Ex Libris and Fretwell-Downing Informatics each developed its own metasearch technology.

This metasearch realm brings to the fore a number of technical issues that require cooperation among the stakeholders to resolve. The National Information Standards Organization (NISO) established a Metasearch Initiative in April 2003 to facilitate the development of standards, practices, or guidelines that will be of mutual benefit to the technology developers as well as to the community of publishers of information resources these technologies deliver. To date, though, the NISO Metasearch Initiative has not resulted in the development of a formal standard.

In this context of a rapidly expanding segment of the industry, Webfeat announced it has been awarded a patent covering several key components of metasearch technology. United States Patent and Trademark Office awarded Webfeat patent number 6,807,539, dated October 19, 2004. The patent focuses its claims on the session management and authentication components of a metasearch system, but also describes the broader methods of searching multiple resources or databases. The final patent application was filed on September 23, 2003; an earlier version of the patent application was filed April 27, 2001.

Todd Miller, founder and president of Webfeat, says his company was the first to solve the most difficult aspects of metasearch—that of the developing technology that performs authentication into restricted databases and managing the session. “The field of competitors in

the market today did not exist at the time of the original patent application,” explains Miller. “The claims in the patent are specific to authentication and session management and are not especially broad. The original patent application made broader claims that were pared down in what was awarded.”

Having the patent will bring benefits to Webfeat; the company can proceed with pursuing new opportunities to license its technology. Prior to the patent award,

---

**In this context of a rapidly expanding segment of the industry, Webfeat announced it has been awarded a patent covering several key components of metasearch technology.**

---

Webfeat had licensed its technology to a number of other companies, including Serials Solutions, Thompson Scientific, and Infotrieve AFX (Article Finder eXtreme) as well as the library automation companies listed previously.

Not surprisingly, others in the industry voice a strong concern over the award of the patent. In the short time elapsed since the patent award announcement, none of the companies with competing technologies have announced any specific plans in reaction it. While there is speculation that one or more of the competing interests may challenge the patent formally, none have taken specific action. Doing so would be a long and expensive process.

Though they do exist, patents are not commonplace in the library arena. OCLC, for example, owns seven patents issued between 1983 and 1996. More recently, a 2003 patent application related to OpenURL caused, and continues to cause, concern.

Among Webfeat’s competitors there seems to be an almost universal belief the patent *should not* have been issued and that it will be disruptive to the industry. Most of the comments regarding the patent focus on the issue of prior art; such comments assert many projects and systems have utilized the technologies, the ones claimed and awarded in the patent, for at least the last decade. And because patents are transferable property, there’s also concern about how the patent might be used should its ownership be transferred. For example, if Webfeat is acquired or sells the patent, it could be used more aggressively than Webfeat itself might have intended.

The library automation industry has a long history of developing open standards as the basis for competing products in the same arena. These standards provide opportunities for interoperability and protect libraries from any given company’s proprietary technology. But these standards do not preclude vigorous competition. It is yet to be seen how this patent will affect the development of standards related to metasearch and whether there will be any significant business implications between Webfeat and its competitors because of this patent.—*Marshall Breeding*

Contact: NISO Metasearch Initiative:  
[www.niso.org/committees/MS\\_initiative.html](http://www.niso.org/committees/MS_initiative.html)

Webfeat press release:  
[www.webfeat.org/releases/2Dec04\\_Patent.htm](http://www.webfeat.org/releases/2Dec04_Patent.htm)

Google *from page 1*

## Library Logistics

Michigan and Stanford plan to work with Google to scan nearly all of their print collections. At Michigan, that amounts to approximately seven million volumes that will take six or more years to complete. If the Google scanners work around the clock for six years, they would need to scan 3,200 book each day, or 2.2222 books per minute. Harvard plans to begin with a pilot project that will digitize approximately 40 thousand volumes of the 15 million volumes held at Harvard University Library. The test volumes will be drawn at random from the Harvard Depository, an off-site storage facility for infrequently used scholarly books.

According to the BBC News World Edition article, approximately one million 19th-century volumes from the Bodleian Library at Oxford will be scanned.

Google plans to use a non-destructive scanning process. An article in the *San Jose Mercury News* notes, "Google is using its own, secret scanning and digitizing technology that it says will not harm older, delicate books." Once the contents of a volume have been digitized, the participating library will decide what to do with the physical volume.

Google has reached separate agreements with these five research libraries, and it is bearing the direct costs of this cluster of digitization projects. Google plans to install proprietary, high-speed scanning facilities at Harvard, Michigan, Oxford, and NYPL. Because the Google home office is only a few miles from the Stanford campus, volumes from Stanford will be shipped to Google for scanning.

Several of the media reports published on the heels of this announcement indicated the average cost to scan a book, in the context of a massive project such as this, would be approximately \$10 per book.

It is unclear how the five scanning projects will be coordinated to avoid redundant effort. Ten bucks per book doesn't sound like much, but multiply that by thousands, hundreds of thousands, or millions of potentially duplicate books held by two or more of the five collections, and you're talking mega money.

For example, both Michigan and Stanford hold at least one print copy of William Wallace's 1890 book, *The Life of Arthur Schopenhauer*. If the Google scanning beavers at Stanford scan the book on a Tuesday, how will the Google scanning beavers

in Ann Arbor—as they pull the book from the shelves on Wednesday—know that and prepare to scan it? Multiply this question by millions of books shelved and pulled in different orders at the five libraries, and you have a logistical nightmare on your hands.

Actually, Wallace's biography of Schopenhauer is a bit of a trick example. Evidently, a reprint of the 1890 book was published in Kitchener, Ontario, Canada, in 2000. The reprint was digitized and became part of the ebrary digital library, which already is available to Stanford-affiliated users. Because the reprint already has been scanned, albeit by another for-profit company using a different scanning method, should one (and, ideally, only one) of the original print copies be scanned by the Google scanning beavers? If we assume that all this knowledge necessary to make an informed decision *can* be delivered on demand to the various Google teams, they will need to make a rapid decision, because *tempus fugit*, and the rate of 2.2222 books per minute must be maintained.

In addition, it is not yet known *when* these scanned scholarly books will begin showing up on the results page of your Google "Greek vase painting" search.

If this project proceeds at the scope and pace outlined in December 2004, the impact on many facets of academic library operations could be profound. If the complete texts of more than ten million out-of-copyright scholarly books are available online worldwide, to everyone, many academic libraries may



choose to maintain print collections only for those works for which copyright is still in force. To free up space for other uses and to avoid construction costs, regional print repositories may become swamped with donations from universities and colleges culling their collections.

## A Changing World

“This is the day the world changes,” said John Wilkin, a University of Michigan librarian working with Google, as quoted by the BBC News World Edition article.

In the article, Wilkins goes on to say, “It will be disruptive because some people will worry that this is the beginning of the end of libraries. But this is something we have to do to revitalize the profession and make it more meaningful.”

The competition for large digital academic library collections (tentatively defined as those containing at least one million e-books) may not take long to heat up. Within 24 hours of Google’s big announcement, the Internet Archive, a not-for-profit organization based in San Francisco, announced that ten major libraries around the world had committed to combine their e-book collection into a free archive hosted by the Internet Archive.

Within the United States, the initial participants will be the Library of Congress (American Memory) and Carnegie Mellon University (Million Book Project). Other participating libraries include the universities of Toronto, Ottawa, and McMaster in Canada; Zhejiang University in China; the Indian Institute of Science; the International Institute of Information Technology; the European Archive; and the Bibliotheca Alexandrina in Egypt. The Text Archive at the Internet Archive could make more than seventy-five hundred texts available online by the end of the first quarter of 2005.—*Tom Peters*

Contact: Google’s Press Release:

[www.google.com/press/pressrel/print\\_library.html](http://www.google.com/press/pressrel/print_library.html)

Google Print: <http://print.google.com>

Harvard’s FAQ:

<http://hul.harvard.edu/publications/041213faq.html>

Harvard Depository: <http://hul.harvard.edu/hd>

San Jose Mercury News article:

[www.mercurynews.com/mld/mercurynews/10412659.htm?1c](http://www.mercurynews.com/mld/mercurynews/10412659.htm?1c)

BBC News World Edition article:

<http://news.bbc.co.uk/1/hi/technology/4094271.stm>

## Bob Walton STOPS SONG ON DYNIX SEAT

*By now, we’re used to the musical chairs of executives and staff as they move from one library automation company to another. It’s a bit more surprising when such moves take place at the board of directors’ level. Shocker aside, that’s what happened in November as Bob Walton joined the Dynix board of directors after serving as chairman of the Ex Libris (USA) board of directors from May 2003 through at least mid-2004.*

While at first blush, it might appear that Walton’s transition from Ex Libris to Dynix represents some alliance between the companies, further study seems to show this likely *isn’t* the case.

While Ex Libris (USA) had a vacancy left by the departure of Carl Grant, Walton served on the board, primarily in an advisory role. With the appointment of Dan Trujman as Ex Libris (USA) president, Walton’s work for the company was complete.

As a library automation industry veteran with significant experience in finance, Dynix also sees Walton as a valued advisor. The Dynix Board of Directors is chaired by John Ware, president and CEO of 21st Century Group, the venture capital firm with the largest equity ownership in Dynix.

Walton’s previous executive positions include that of CLSI president, from July 1990 through October 9, 1992, shortly before the company was sold to Geac; and executive VP and CFO of Innovative Interfaces, October 1992, through July 31, 1999. Since 1999, Walton has served as the VP for finance at the College of Wooster in Ohio.—*MB*

## Libraries Up Use of WEBINAR SOFTWARE

In November 2004, the Microsoft Accessible Technology Group opened its doors to 75 representatives from more than 20 nations, the World Wide Web Consortium's Web Accessibility Initiative, the United Nations, and the World Health Organization. Joined by Microsoft's indomitable leader Bill Gates, the three-day international forum, "Libraries for the Blind and Print Disabled: Moving toward a Digital Future," explored and examined the digital future of libraries serving print-impaired readers, including patrons who are blind, visually impaired, physically challenged, or learning disabled.

Held at Microsoft's Washington state corporate headquarters in Redmond, the summit was cosponsored by the DAISY Consortium, an international organization that develops standards and best practices for digital audio books.

Summit attendees focused on several areas that require additional resources and effort, including:

- Converting more library materials to digital formats. Digital formats in general promise better accessibility and general usability for print-impaired library patrons.
- Developing better and faster distribution methods for digital content that are accessible to all users.

- Providing better digital library services for print-impaired library patrons.

One major and ongoing challenge is to ensure that software, hardware, systems, and content are accessible to screen reader programs used by many blind individuals to access content. The major screen reader software programs in the United States market include JAWS for Windows from Freedom Scientific and Window-Eyes from GW Micro.

As a result of the summit, a joint commitment—involving librarians, publishers, technologists, and others—was drafted to help develop and implement a worldwide collection of accessible library materials. One plank of the commitment document is to integrate the various facets of this global effort with mainstream libraries. Rick Weingarten from the ALA Office of Information Technology Policy will be one of the leaders of the integration effort.—*TP*

**Contact:** Microsoft's Press Release about the Summit:  
[www.microsoft.com/presspass/features/2004/nov04/11-15LibrarySummit.asp](http://www.microsoft.com/presspass/features/2004/nov04/11-15LibrarySummit.asp)

DAISY Consortium's Press Release about the Summit:  
[www.daisy.org/news/news\\_detail.asp?NewsId=172](http://www.daisy.org/news/news_detail.asp?NewsId=172)

JAWS for Windows from Freedom Scientific:  
[www.freedomscientific.com/fs\\_products/software\\_jaws.asp](http://www.freedomscientific.com/fs_products/software_jaws.asp)

Window-Eyes from GW Micro:  
[www.gwmicro.com/products](http://www.gwmicro.com/products)

## MICROSOFT SUMMIT FOCUSES ON INFO ACCESS

Whether called webinar, web conferencing, online meeting, or web collaboration software, products in this software category are being used increasingly by libraries and library-related organizations of all types, both to avoid the costs of in-person meetings and conference calls, and to expand and extend public programming online.

Webinar software typically offers several key advantages over a telephone conference call:

- Voice-over-IP (VoIP) that avoids per-minute telephone charges
- Synchronized browsing, even into proprietary databases
- Text chatting, often with timestamps, hotlinked embedded URLs of pages visited, and so on
- Private text chat between two or more individuals attending a meeting or public event.
- Presentations
- Application sharing, often with whiteboard features
- Versatile, easy-to-use recording features, which often capture all facets of an online meeting or public event (sound, text chat, presentation slides, Web pages)

# RFID Gets TAGGED

The advantages and pitfalls of RFID (radio frequency identification) tagging systems—what could be thought of as really smart barcodes—continue to be hot topics in librarianship. In the November 29, 2004, issue of *Publishers Weekly* (251, no. 48, p. 9), James Lichtenberg provided an update on RFID from the vantage point of publishers and major corporations. Many of these entities are attempting to comply with Wal-Mart's early-2005 mandate to include RFID tags on boxes and pallets of merchandise.

Lichtenberg reports on Pearson Education, a publisher that's been experimenting with inserting RFID tags in hardcover and paperback textbook spines. Lichtenberg writes, "The spine was chosen as a place where the tag could not easily be tampered with or removed. Reportedly, the experiment with hardcovers was successful, but the glue and heat required for binding a paperback damaged the type of tag Pearson was using." This may turn into another version of the tattle-tape saga.

EPCglobal, the trade association (quoting from its Web site) "leading the development of industry-driven standards for the Electronic Product Code (EPC) to support the use of Radio

Developed primarily for corporate users, webinar software products—such as Macromedia's Breeze, Talking Communities' iVocalize, Microsoft's Live Meeting, and WebEx—are part of a crowded and competitive market.

In conjunction with the December 2004 issue of INFOSCOPE, Donna Cohen, an information management consultant with D. L. Cohen Information Services and an adjunct instructor for the School of Library and Information Management at Emporia State University, used Talking Communities' iVocalize to create and distribute a brief introduction to webinar software. Cohen's introduction puts to good use some of the key features of webinar software, including VoIP, presentation slides timed to appear when appropriate, screen shots, and easy-to-use recording/capturing functionalities.

As telephone and travel costs rise, and as individual libraries' travel budgets and telecommunication budget lines remain

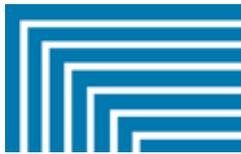
Frequency Identification (RFID) . . .," held its Fall 2004 meeting in Baltimore, which attracted more than 2,000 corporate representatives. Evidently, many corporations are taking RFID seriously now, with active plans and tests underway that will lead to full-scale deployment.

Libraries and librarians are scrutinizing RFID systems as well, not only in terms of costs and benefits, but also regarding professional and end-user concerns about possible invasions of privacy and misuse. Librarian, blogger, and RFID maven Lori Bowen Ayre recently moderated an RFID Q&A panel at the California Library Association's annual meeting. Among Ayre's vendor panelists were: Rebekeh Anderson, 3M; Ron Birchard, Integrated Technology Group; Oleg Boyarsky, Library Automation Technologies; Charles Boyer, Tagsys; Emmet Erwin, Bibliotheca; Dan Denault, VTLS; Doug Karp, Checkpoint; and Frank Mussche, Libramation. Vendors that responded to the privacy question have differing opinions concerning whether privacy concerns are reasonable or overblown. Ayre, who is also the author of the March/April 2004 issue of *Library Technology Reports* ("Filtering and Filtering Software," 40, no. 2), provides a concise wrap-up of the RFID panel discussion in her November 30, 2004, entry at the blog URL listed under "Contact."—TP

Contact: EPCglobal: [www.epcglobalinc.org](http://www.epcglobalinc.org)  
Lori Bowen Ayre's blog entries:  
[www.galecia.com/weblog/mt/archives/  
cat\\_rfidandwireless.php#000154](http://www.galecia.com/weblog/mt/archives/cat_rfidandwireless.php#000154)

static or even fall, libraries are exploring webinar software as a productive, cost-efficient alternative. Added ergonomic bonuses include: no cauliflower ear and crooked neck from an interminable conference call, and no "quality" windshield time.—TP

Contact: Macromedia: [www.macromedia.com/  
software/breeze](http://www.macromedia.com/software/breeze)  
Talking Communities: [www.talkingcommunities.com](http://www.talkingcommunities.com)  
Microsoft: [http://office.microsoft.com/en-us/  
FX010909711033.aspx](http://office.microsoft.com/en-us/FX010909711033.aspx)  
WebEx: [www.webex.com](http://www.webex.com)  
Donna Cohen's iVocalize presentation about webinar software and libraries:  
[www.dcoheninfo.com/WebCollab/ISDec04.htm](http://www.dcoheninfo.com/WebCollab/ISDec04.htm)



**ALA TechSource**  
www.techsource.ala.org

Smart Libraries Newsletter  
American Library Association  
50 East Huron Street  
Chicago, IL 60611-2795 USA

---

NON PROFIT  
US POSTAGE  
PAID  
PERMIT 1479  
ROCHESTER, NY

---

## February 2005 Patent Prods Metasearch Market Melee

### Smart Libraries Newsletter

Smart Libraries Newsletter delivers hard data and innovative insights about the world of library technology, every month.

---

#### Contributing Editors

Marshall Breeding  
615-343-6094  
marshall@breeding.com

Tom Peters  
816-228-6406  
tapinformation@yahoo.com

#### Editor

Teresa Koltzenburg  
800-545-2433, ext. 3252  
312-280-3252  
tkoltzenburg@ala.org

#### Administrative Assistant

Judy Foley  
800-545-2433, ext. 4272  
312-280-4272  
jfoley@ala.org

### TO SUBSCRIBE

To reserve your subscription, contact the Customer Service Center at **800-545-2433, press 5 for assistance**, or visit [www.techsource.ala.org](http://www.techsource.ala.org).

The 2005 subscription price is just \$85 US.

Production and design by Christine Velez,  
American Library Association Production Services.

*Smart Libraries Newsletter* is published monthly by ALA TechSource, a unit of the publishing division of the American Library Association.

Copyright American Library Association 2005. All rights reserved.