

Information Today

The Newspaper for Users and Producers of Electronic Information Services | www.infotoday.com

January 2009
Vol. 26 | Issue 1

News Briefs

Reed Elsevier Takes Its B2B Arm Off the Market

3

Gannett Acquires Social Media Services Provider

3

OCLC Develops Search Service and More

3

Jordan, Edmiston Group Announces Two Acquisitions

8

IDEX Acquisition Adds 1 Million Expert Witness Records to LexisNexis

8

NewsMakers

23

Applause, Applause

At the start of 2009, we're taking a quick look back to the triumphs and victories of 2008. That's right, *Information Today* is announcing the winners of the 2008 People's Choice Awards.

Despite the ebbs and flows of the economy and the many mergers and acquisitions last year, some enterprises kept the innovations coming and the information technology industry moving ahead. And the votes in this year's People's Choice Awards can attest to this fact.

We kicked off the two-part voting process online last year at ITI's website (www.infotoday.com) for all interested parties. One round of nominations from members of the information technology industry created a sizable list of contenders—some past People's Choice Awards winners and some newcomers. After

the nominations were tallied, ballots were added online for the final votes. At the end of December, we disarmed the online voting process and moved ahead to recognizing this year's winners in each of our 10 categories.

On page 50, you'll find the winners and a profile of each company, person, or product, along with one or two comments from the voters themselves. We also included honorable mentions for those that came close to winning the top award.

The people have spoken, they have cast their votes, and the winners can enjoy some well-earned recognition for their contributions to the IT industry in 2008. The awards in each of the 10 categories pay tribute to the excellence that the winner has achieved, whether through personal innovation or through quality products or services.



Congratulations to our winners. And thanks to those who took the time to cast their votes online. Stay tuned for more information on the 2009 People's Choice Awards in the

next issue. We'll be shaking up the competition with a new category or two. Take a few moments to check out the winners on page 50.

—Barbara Brynko

Where Have All the Lawsuits Gone?

by GEORGE H. PIKE

Have you ever wondered what it takes to settle a massive copyright infringement lawsuit? Well, in the case of The Authors Guild lawsuit against Google over Google Book Search, the answer is 3 years, 323 pages of paperwork, and \$125 million.

In September 2005, The Authors Guild, the Association of American Publishers, and others sued Google over the Google Book Search proj-

ect. Under this project, Google entered into agreements with a number of libraries and other institutions to scan books and other materials into Google databases. Some of these materials were in the public domain, and Google negotiated agreements with some publishers. However, Google had neither permission nor a license to copy the lion's share of the materials.

(continued on page 52)

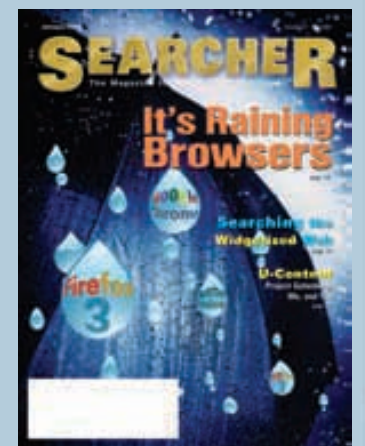
hakia Helps Librarians With Credible Search Results

by VANESSA FOX

hakia, a natural language search engine that uses semantic analysis of the web, recently launched a new initiative with librarians to identify "credible" content on the web. The program is designed to provide searchers with a quick way to access lists of sites that they know professionals have vetted as an alternative to the list of all-web results that search engines typically provide.

"A popular Web source may not always be credible, and a credible Web source may not always be

(continued on page 52)



Inside

Check out Lauree Padgett's new column called Meet the Editor on page 47.

PRESORTED
STANDARD
US POSTAGE PAID
SO. JERSEY, NJ 08031
PERMIT NO. 450

hakia Helps Librarians With Credible Search Results

(continued from page 1)

popular,” according to the site. “hakia is the first search engine to integrate librarians’ collective knowledge of credible Web sites into search results to guide searchers.”

hakia’s credible sites program initially covers health and environmental websites, although it hopes to expand to all topics. hakia isn’t the only search engine that realizes that the web contains a wealth of inaccurate as well as accurate information and that searchers can benefit from vetted results. For instance, Microsoft Live Search provides health results that have been preapproved from sources such as the Mayo Clinic in conjunction with Health Vault. Google also recently launched Knol, which is intended to provide authoritative information.

However, hakia is the only search engine that has specifically reached out to librarians to help with this effort. Melek Pulatkonak, president of hakia, says, “[W]e want to place the librarians and information professionals in the center of the information retrieval conversation. ... Search engines are the single most powerful distribution points of knowledge, wealth, and yes, misinformation. Librarians hold the key to quality Web sources and are the gatekeepers of quality information. Only with their active contribution to Web search and to building resources of Web search engines, can we stop misinformation propagation.” For example, the Medical Library Association recommends credible health sites.

According to Pulatkonak, “Today, almost all search engines rely heavily on ‘popularity’ measurements and rankings.” However, many searchers never look past the first page of results and take those results as “truth,” even though they might be simply popular or, worst case, the result of gaming.

How Do Librarians Suggest Credible Results?

Librarians can join “h CLUB” at club.hakia.com and suggest “credible” sites for available categories. The credibility criteria (<http://club.hakia.com/lib>) include the following details:

- Peer review
- No commercial bias
- Currency
- Authenticity

hakia then QDEXes these sites; QDEXing (query detection and extraction) is hakia’s semantic analysis-and-storage process that determines what questions each page can answer rather than simply cataloging the words on the page (<http://searchengineland.com/social-networking-through-search-hakia-helps-you-meet-others-12586.php>). Pulatkonak

believes “semantic technologies will enable companies like hakia to deploy filters like credibility without compromising on relevance. We think it is time for search engines to assume accountability of search results.”

Kim Megginson, a librarian who is involved with the program, explains that many librarians have been keeping lists of credible websites for their clients for some time. While it’s not a new idea for librarians, it is a new application for a search engine.

Librarian Beverly Butula says, “[T]his is an excellent program. With the wealth of information available online, a semantic search engine is a great alternative. The benefits are obvious. Producing results that are on point and from quality resources not only ensures relevant information, but can make the researcher more efficient.”

How Do Searchers Access Credible Results?

Searchers simply go to hakia.com, enter a query, and view lists of regular search results, images, news, and credible results.

Pulatkonak says hakia launched this program “to offer a quality search experience to Web searchers where they can find relevant results from clearly organized sources such as credible sites, news and more. The current search experience makes us feel that we are searching for a needle in a haystack: the searcher gets millions of search results and has to assess relevancy, credibility and freshness manually by scanning the page. hakia’s analogy would be searching for a needle in various buckets of hay where each bucket is clearly marked and the searcher saves time.”

How’s the Program Working?

hakia reports a great response from librarians so far; hakia has already started to incorporate librarian suggestions into results with a “trusted results” release in October (<http://searchengineland.com/hakia-relaunches-site-with-trusted-results-14949.php>). The feedback received is being incorporated into the program. Pulatkonak says hakia is committed to its relationship with librarians, and it highly values their contributions in helping searchers find truthful information online. A University of Wisconsin law library blog post noted that hakia is interested in the legal vertical next and encouraged librarians there to participate (www.law.wisc.edu/blogs/wisblawg/2008/09/haika_semantic_search_engine_1.html).

In Today’s Searching Culture, Does Google Equal Truth?

The last 10 years have brought an amazing change in our culture. Although we suddenly appreciate the value of searching,



finding answers to our searches in some cases is useless if we don’t find accurate answers. Of course, librarians provide a service: ensuring a searcher receives an accurate answer. Google doesn’t promise accuracy, but some searchers seem to assume it does anyway. Could hakia’s outreach with librarians be a way to bridge that gap and combine the instant searching needs of our culture with the reliability of accurate results that librarians can provide?

I talked about this with some of the librarians involved in the program. They noted that many searchers are in a hurry and may not realize that they’re not getting the best results available. Often, the results they get are sufficient. Not every search requires peer-reviewed journals. And the librarians had faith that the searchers realize that not everything online is reliable.

Megginson notes, “[C]asual searchers will probably be content with the top 10 results with Google. Searchers wanting specific or more in-depth information can be frustrated by the results of just one search engine. To Google’s credit, there are now other Google search engines for such topics as scholarly literature and patents. However, there is so much more out there and hakia can certainly fill a gap with the potential to limit to credible websites.” She says searchers know that some websites are more trustworthy than others, and they do critically evaluate sources of information.

“It is very likely that Internet users will be the ones to leverage between the searching culture and trusted results,” she says. “Of course, librarians have a role in educating Internet users about searching for information on the world wide web.”

Butula notes that the trend to search online will continue to evolve and that librarians can evolve with that trend. “An engine that focuses on producing quality, reviewed results will have an obvious benefit for not only researchers, but also the ‘average’ web user. If the trend is to rely more on online materials, as a librarian I am in favor [of] building search engines that produce quality, evaluated sites. Many searchers do not evaluate sites for authority to the degree they should—that is why relevant results are imperative.”

Vanessa Fox, an entrepreneur-in-residence with Ignition Partners and features editor at Search Engine Land, is co-chair for O’Reilly Found, a conference for web developers about SEO. She previously created Google’s Webmaster Central, which provides both tools and community to help website owners improve their sites to gain more customers from search. She was recently named one of Seattle’s 2008 top 25 innovators and entrepreneurs. Send your comments about this article to itletters@infotoday.com.

Where Have All the Lawsuits Gone?

(continued from page 1)

Of course, Google claimed that copying was a fair use. Its use of the books was in an index. Searches would obtain only limited text of a particular book, along with a link to a library or seller of the book for more information. In fair use parlance, Google’s use was “transformative.” The Authors Guild and publishers disagreed just as obviously, claiming that the copying was too extensive to be fair, and that Google’s copying of entire books for commercial use was not a fair use.

Fair Use Unresolved

The settlement resolves the dispute between Google and The Authors Guild, but the legal question of whether Google’s copying was fair will remain unresolved, as is often the case. We will never know how a court may have balanced Google’s transformative use (an essential element of the fair use test that favors the copyright user) with its copying of entire books and its profiting from the program, elements of the fair use test that favor the copyright owner. It is only through court decisions, not settlements, that legal precedents are created that can help to resolve future fair use disputes.

Arguably, Google’s willingness to pay \$125 million to settle the lawsuit suggests that it knew it might lose its fair use claim. However, when examined against the risk of hundreds of millions—even billions—of dollars in potential damages (as well as the benefit of using the settlement as the foundation of a new business model), Google’s settlement makes sense.

Book Rights Registry

Under the settlement, Google will pay at least \$45 million to authors whose books have already been scanned. Google also agreed to fund a Book Rights Registry, a nonprofit organization that will coordinate payments to authors for online uses of their books.

More importantly for Google and copyright owners, the settlement allows a significant expansion of the Google Book Search program, both in the amount of content available and through new marketing opportunities such as institutional subscriptions, library access, and additional purchasing platforms. Copyright owners will get two-thirds of any revenue generated by the Book Search project, including Google’s advertising revenue. Google will receive the remaining third.

Viacom Versus YouTube

Copyright infringement remains at the center of other major lawsuits still

(continued on page 54)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.