



The Time Has Come for Visual Search

Visual search, the ability to review results graphically, has been on the horizon for quite some time. Industry prophets, those who stay awake until 2 or 3 a.m. exploring the most esoteric Web sites, have been pointing to graphical results as a likely method for improving users' ability to review search results in a more meaningful way. What initially caught the attention of the working searcher community, however, was the joint venture between Groxis and Yahoo! offering visual search, for free, to anyone who wished to try it.

Judy Luther, Maureen Kelly, and Donald Beagle pointed out in "Visualize This" (*Library Journal*, March 1, 2005) that visualization tools combine concept clustering with the ability to get "a 'picture' of the meaning behind the words." Users can quickly discover aspects or emphasis or content types that lead to more meaningful searches.

Chris Sherman, in "Visualizing Yahoo Search Results" (SearchEngineWatch, June 23, 2005), said that with "results ... grouped in topics, rather than presented in a linear list, you often see results that you might otherwise miss." Sherman also liked the preview features, filtering tools and the ability to "pinpoint specific types of content ... clustered together" within search results.

Pundits may applaud visual search, but the overwhelming majority of information professionals searchers stayed true to the linear, relevancy-ranked list of results, leaving visualization to the early adopters and techies. Change, however, is imminent.

Recently, both EBSCO and Factiva introduced visual search as part of their mainstream products. Is it possible that visual search is finally becoming a tool that will add value, and therefore attract more usage, now that it is part of key vendor offerings?

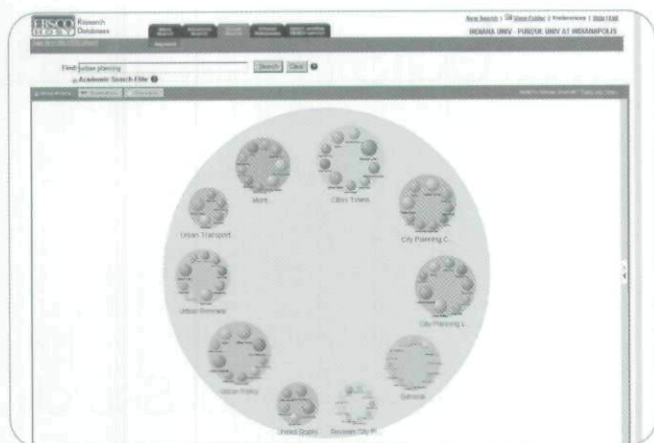
VIEW FROM EBSCO PRODUCT MANAGEMENT

In March 2006, Groxis and EBSCO Publishing announced a partnership to add visual search to EBSCOhost. According to the announcement, "The Grokker visual search feature on EBSCOhost organizes hundreds of search results from a single query into an intuitive, easy to navigate content map. The user can dynamically filter results by keyword, publication date, and full text to pinpoint specific articles and records without having to reframe the query. Through this partnership, EBSCO becomes the first major content aggregator to implement visualization technology into the aggregated database search experience."

Always curious to know how product initiatives happen, I asked Mike Gorrell, CIO at EBSCO, what triggered his interest in visualization. Interestingly, Gorrell said that Grokker and EBSCO had a mutual customer—Sun Microsystems.

The SunLibrary elected to use Grokker to provide federated access to IEEE, NetLibrary eBooks, EBSCO journals, the Web, and MailFinder, Sun's repository of research-related e-mail. Sun researchers could more efficiently view and evaluate

“Is it possible that visual search is finally becoming a tool that will add value, and therefore attract more usage, now that it is part of key vendor offerings?”



Visual Search results on Academic Search Elite for urban planning

relationships from the integrated results. The topically organized visual map, preview, filtering, and collaborative features proved popular with Sun staffers. SunLibrary surveys noted increased use of premium content subscriptions and improved user satisfaction.

Gorrell's idea was to get the user past looking at the first 10 results and was "looking for ways to improve search and the user experience." He wanted to, "allow the user to use their results" and offer a "compelling visual experience." He was attracted to Grokker's ability to enable users to "navigate the visuals, zoom, click, and explore results and tangents."

Visual Search uses EBSCO's proprietary natural language processing and relevance ranking, with weighting on EBSCO subject terms, to retrieve a results set of 250 items. Next, this results grouping is handed off to Grokker to organize into a content map. As Grokker has its own internal dictionary, work was done to map the terms from EBSCO search to the Grokker dictionary.

Gorrell told me that subject term clustering will soon be added that will show the top subject terms. Clustering will operate in visual, basic, and advanced search modes. Databases within the EBSCOhost interface that have their own controlled vocabularies, such as MEDLINE and INSPEC, will also offer clustering. The Grokker visual search is offered to EBSCOhost users with no additional fees because EBSCO seeks to "increase loyalty and product differentiation," according to Gorrell.

SEARCHING EBSCO VISUAL SEARCH

The EBSCO interface now offers four main search tabs: Basic Search, Advanced Search, Visual Search, and Choose Databases. By clicking on "Visual Search," the user moves to this new option, which is a stand-alone offering. Typical EBSCO search options, such as Keyword, Publications, Subject Terms, Indexes, and Images, are not presently available.

The Visual Search option is a single "find" box and the results are presented in a visual map organized by topic. As with other Grokker applications, circles represent categories and

squares symbolize individual articles. My search on Academic Search Elite for **urban planning**, a typical student research paper topic, returned 10 perfectly reasonable circles corresponding to a variety of related categories, including: Urban Transportation, Urban Renewal, Urban Policy, Reviews City Planning Non-Fiction, NY New York State, General, City Planning Social, City Planning Land, Cities & Towns. A "More" category offered further exploration. You can drill down further for more clarification. Clicking on "Cities & Towns," for example, revealed nine subcategories.

The dynamic, "zoomable" Grokker map includes Show/Hide filters, enabling users to refine their search. Searchers can do a date restriction with a slider bar that, based on the new retrieval, changes the Grokker map. This functionality displays the relationship between the topical category and the selected time frame. Users can immediately determine that the most recent articles deal with particular aspects of the overall topic. Additional filter options include a limit to full text and the ability to refine by adding a keyword.

At the article level, a mouseover reveals a minitization that includes article title, author, publication, date, and a full-text indicator. Complete article displays include the standard EBSCO options of print, e-mail, save, and add to folder. (Folder options are dependent on the user setting up personalized folders using the My EBSCOhost feature.) Once the user is viewing articles, a split-screen is offered, showing the individual article and the visual map simultaneously. The user can click to see just the map or article as well.

VIEW FROM FACTIVA PRODUCT MANAGEMENT

Factiva's Dennis Cahill, newly promoted vice president-product from associate VP of technology, was charged with examining the emerging world of text mining to determine if this technology offered a method for providing added value to Factiva products.

Initially, Cahill and his associates focused on using text mining for the Factiva Insight product line. According to the Factiva Web site, "The Factiva Insight suite of products combines research methodologies, trusted content and advanced text-mining and visualization tools into powerful media and reputation intelligence solutions enabling you to demonstrate the effectiveness of your communications strategies." Insight, targeted at marketing communications, media, and competitive intelligence professionals, provided a "role-specific" context for the text-mining application.

Next, the Factiva product team looked to leverage the Yahoo!/Google search popularity by creating a tool to generate "powerful search results from a simple search experience." That tool, called Factiva Search 2.0, is available within Factiva's core product, Factiva.com. Specifically, Cahill's team tried to construct "intelligence within a search result" by taking advantage of the "underlying structure of the retrieved data, such as taxonomies." This approach allows the searcher to enter potentially ambiguous search terms, such as acronyms, in which the retrieval would allow for refinement and offer greater implicit relevance.

The team also used the text-mining approach to essentially build a new indexing field called News Clusters. The articles are run through the text-mining application to extract the top 10–12 phrases that are weighted by location (title and lead paragraph are more significant) and frequency in the article. These News Clusters are included in the visual search Discovery Pane, which also includes Date, Companies, Industries, Subjects, and Keywords. These refining concepts are displayed as visuals with interactive mouseovers. Experienced Factiva searchers will recognize that Companies, Industries, and Subjects are based on the Factiva Intelligent Indexing.

USER INPUT

Cahill's team is making extensive use of user feedback and input during the product development process. Opinion is obtained via several mechanisms, including the following:

- Surveys within the application that asked users what they thought of the search results and which discovery tools worked best for them
- Focus groups with Factiva's largest customers that involved users of all types, from professional searcher to occasional user. The searchers were presented with "use cases" that took them through a variety of search scenarios, followed by discussion.

Overall, feedback has been positive. Users will also be pleased to hear that the visualization and discovery tools within Factiva 2.0 are provided at no additional charge. According to Cahill, in what emerged as a common theme, Factiva was seeking to "drive value and differentiation" rather than increase prices.

Factiva Search 2.0 is a true beta. According to Factiva, "After a larger amount of customer feedback is evaluated, the beta will conclude and some or all of Search 2.0 will be integrated into various Factiva products." Changes are being implemented as the beta progresses. For example, Factiva Search 2.0 has replaced the start page in Factiva.com, a session-based Clipboard has been added, additional search options have been provided, and a remove filters function has been added. Future enhancements will be based around driving

increased implicit relevance with more focus on job functions, taxonomy elements, term frequency, proximity, and publication date.

SEARCHING FACTIVA 2.0 SEARCH BETA

Users start with a simple single search box. Web search conventions, such as use of parenthesis to indicate immediate adjacency, work nicely. For example searching on **Experian market share** produced 315 hits, whereas searching **Experian "market share"** retrieved 30 hits. Interestingly, the first search also provided a Google-like suggestion: "Did you mean Experian market square?"

Retrieval is divided between Publications and Web News, with All Publications being the default display. Publications include newspapers, magazines, and newswires. Click to limit retrieval to that publication type. Sort options include relevance and date. There is an "Alert me" option. What's new about Factiva 2.0 is the Discovery Pane located on the right side of the display.

The first option in the Discovery Pane is the Date display, arrayed as a vertical bar chart that shows the article distribution by date range. Mouseover to indicate the start date and end date represented by the bar and the number of hits within that time frame. You can clickthrough to see the articles, as you can with all Discovery Pane options.

The second option is the News Clusters mentioned above. Essentially offering the ability to search on "aspects" of the search, the mouseover shows number of hits. Users can click to see those specific articles. The Factiva Intelligent Indexing, consisting of Companies, Industries, and Subjects, is collectively shown as horizontal bar charts, with the number of hits displayed within the bar. The final display is keywords. Clicking this option retrieves articles ranked by the most occurrences of that keyword within your initial search results. Companies, Industries, and Subjects are sequenced in descending order by number of hits.

Full article displays include highlighted search terms, options to clickthrough to "More Articles Like This," as well as to related Industries, Companies, and Subjects. Navigation features include Return to Headlines, Previous Article, Next Article, Print, and E-mail in addition to Manage Alerts, Feedback, and Help.

VISUALIZATION TOOLS—HERE TODAY ...

Since user search styles are as diverse as the Web itself, integrating knowledge of visual search into your storehouse of online tools is probably a good thing. Even if you prefer a list format for results, your users could be more visually inclined. Be assured that libraries of all types and their vendors are also exploring visual search.



Results searching Factiva 2.0 using the term **Experian "market share"**

George Plosker [gplosker@365media.com] is director, business development, 365 Media. He writes here in a personal capacity.

Comments? E-mail letters to the editor to marydee@xmission.com.

Copyright of Online is the property of Information Today Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.