

Online Spotlight **Yahoo! Correlator**

ONLINE January/February 2011

I love Google as much as the next person. That being said, Google is starting to annoy me. Google Instant, which attempts to guess what I want and starts feeding me search results even before I finish typing, makes me feel like I'm watching hyper-caffeinated hamsters. Yes, I can turn it off, but this is one more reminder that Google is trying its best to anticipate what I want, and it's doing a seriously poor job of it. Of course, that is because most of my Google searches aren't like the 99% of typical searches. I'm not looking for sports scores, news, celebrity gossip, or porn, so Google's attempts to anticipate my query don't usually work that well.

In fact, one of the drawbacks of most search engines is that they don't automatically understand the context of a query. A search engine may look at the last queries I have made and extrapolate, but this is not a helpful feature for info pros, who are searching on a wide variety of topics over the course of a day. While I can add more search terms to narrow the query, what I often want to do is just say, "Focus on this aspect of it." What I want is a search engine that says to me, like a good reference librarian, "Tell me more about what you are looking for." I have seen startups trying to tackle this problem come and go. I was starting to wonder if this was one of those things, like sentiment analysis, that is always 5 years away from development. Then I ran into Yahoo! Correlator (<http://correlator.sandbox.yahoo.net>). This is part of Yahoo's Sandbox area where Yahoo! tries out new ideas. This particular demonstration is useful if you aren't quite sure what aspect of a topic you really need. (I don't know about you, but I often find that flailing wildly at the beginning unearths interesting leads.) Correlator gives you information - not just search results - that helps you get a better sense of the various aspects of an idea. At this point, Correlator only works with Wikipedia. For ambiguous research projects, I'm likely to start there anyway, and I find that I can get a much better idea of the information landscape with Correlator. Type your query in Correlator just as you would in any other search engine. Below the search box, though, is something different. There are icons for Wikipedia, Names (geographic, personal, company), Places, Events, Concepts, News, and Answers.

The default search result includes results grouped by Wikipedia categories. The search query bioreactor, for example, gives me information organized into areas such as biochemical engineering, biotech, chemical engineering, and environmental engineering. Right there, I get a sense of which direction I might take in my research.

The content is extracted from a number of Wikipedia articles, so I am getting a more three-dimensional view of the topic. When I click the Names icon, the result is a word cloud and a list of frequently occurring names related to bioreactor technology - Nanobac Pharmaceuticals, Shussel, and Neva Ciftcioglu (a company executive). If I click through to any of these names, I could see a new set of articles. Clicking the Places icon displays a map of the world, with flags in cities or

regions that are significant in the context of bioreactors. This alone gives me ideas on where to look next - local papers in San Francisco, for example, if I wanted to know what was in testing phases.

The Events icon displays a timeline of events, again culled from Wikipedia, pertaining to bioreactors. This gives me a sense of the time frame with which to limit my searches in the professional online services. The News icon produces snippets from 20 or so recent stories in Yahoo! News on the topic. Clicking the Answers icon runs my query in the Yahoo! Answers database. Since Yahoo! Answers tends to be consumer-oriented, the results help show some basic questions regarding bioreactors and are useful when you're looking for consumer reactions or concerns.

The Concepts icon produces another word cloud and list of keywords and phrases from the relevant Wikipedia articles on bioreactors. Of all the icons, this is the most variable in terms of usefulness; at the least, it serves as a reminder of alternative words or phrases that often appear in the context of my query. Interestingly, Correlator lets me "correct" results in the Names and Concepts field; if I think that its algorithm included a spurious entry, I can flag it.

While limited to Wikipedia results, I have found Yahoo! Correlator to be a quick way to see the wide range of possibilities for me to explore with a new concept, particularly when I am looking at a subject that spans a variety of possible topics.

Mary Ellen Bates (mbates@batesinfo.com; www.batesinfo.com) enjoys playing in (virtual) sandboxes.