

## **InfoTech Column**

Due Nov. 20, 2006

*Information Outlook*, Dec. 2006 Issue

### **Growing Your Own Search Engine**

By Stephen Abram

There have been a few new and exciting custom search engines out there lately. And they have some applicability to the special librarian's and information professional's daily work.

Here's the short list:

Google Co-op (beta) Custom Search Engine

<http://www.google.com/coop/cse/overview>

Rollyo (beta) Roll Your Own Search Engine

<http://www.rollyo.com/>

Yahoo! Search Builder

<http://builder.search.yahoo.com/m/promo>

PSS: Personal Search Syndication

<http://www.pssdir.com/>

Eurekster's Swickis

<http://www.eurekster.com/>

OK, now you learners who just want to go play and see what they do, go play. It's a fast way to learn. For the rest of us who need to read a bit first, here's a quick overview of this trend. Then go play. For some of us, it's like washing our hands before we eat, brushing our teeth before you leave the house, or stretching our arms before we keyboard. Reading first is a necessary habit for many of us.

Arriving at desktops near you is the next step in search – personalized searches that go beyond just simple alerts but let you tune your search by topic, domain, and presentation style or down to specific root URL's.

Workaday, many of you find yourself searching, repeatedly, small groups of websites that matter to you on a regular basis. It is particularly common to check competitors' sites, special sites like retail pricing sites or country, military and government domains. Sometimes you have used web watch services to track

changes and sometimes you just search normally hoping to pick up new information. Frustratingly you might be using URL searches in Google, in the old-fashioned reiterative mode. Rarely, but it still happens, information professionals will visit a list of important sites daily to check them out. All of this is a Hoover™ level personal productivity suck even though it can deliver high value information and intelligence.

Recently, a load new search services are providing the ability to create your own search engine, and search the way you want on a small scale. Many are in beta but still useful and lie on top of standard web harvests like Google and Yahoo! There are ways with the right budget and ROI to create specialized search engines on any scale such as the FAST service from SirsiDynix in Rooms. However, many of our smaller research needs can be addressed by these mini-custom search tools. Besides, it's worthwhile to experiment with these, since it will make us a better evaluator and purchaser of enterprise-wide, value-added search environments and tools.

## **Google Co-op**

Google Co-op is a platform that enables you to customize the web search experience for users of both Google and your own websites. It's really a simple development tool that can be applied to your internal or external work environment. You can create your own search engine and use the ubiquitous Google search technology to create a free that reflects your needs. As an aside, but an important consideration, you need to assess how comfortable you might be in earning advertising revenue from the resulting traffic (and serving up your organization's search traffic to another entity).

With Google Co-op you can deliver specialized search results by encouraging users to integrate your information or services into their Google search results. You will also help users refine their searches and use your information professional expertise to help improve Google web search for specific subjects by labeling the best sites.

One recent example of a Google Co-op application is LISZEN, (<http://www.liszen.com/>), which has created a single search engine to search over 500 library and information science specific blogs. It's a useful, targeted search aimed at a specific market of librarians. I have written in the past about the need for a search engine for library websites by type of library and subject. For instance, why is it so difficult to search for public library websites aimed at teens or academic library websites that offer e-reserves, virtual reference or e-books? How about associations or charities? Google Co-op offers the opportunity to create a tuned search engine like this. The holes in the Swiss cheese could be filled collaboratively.

I can easily imagine applications where you search all of your competition using this tool. Also, pick a topic, any topic, and your bookmark file of websites on that topic area and create a search engine that searches them all at once. You could even make a useful public presence by creating a specialized search pages for every good site you know about Alzheimer's, MS, MLK or Nanotechnology. Your primary limitation is your own imagination, time and energy.

## **Rollyo**

Rollyo (its tagline shows this is a true child of the 70's - Roll Your Own search engine) is free. The limits are that you have to register and you can only select up to 25 websites to search. You can choose any set you want so it could be personal hobbies or music sites or professional, subject or topical sites. I find that I learn things better when it's a personal interest so I pick that as my test. I learn faster and have a tool that I can use for fun later. I can then transfer my learning to work. Besides, sometimes I know more about my personal interests and can then evaluate the tool better. Rollyo is built on top of the Yahoo! web harvest and uses their defaults. Rollyo calls these custom search engines 'searchrolls'. You can share your searchroll with anyone, keep it to yourself or paste it into your website, blog or intranet. There are even ways to add the searchroll to your Firefox browser as a bookmarklet or search bar button.

Rollyo is just simple and easy to use. It's free so there's no real downside other than you will learn something about creating simple federated search.

## **PSS! (Personal Search Syndication)**

PSS! or Personal Search Syndication, is pretty new. With the cool tip of the hat to RSS, it has a winning name. It's free like Rollyo. The differentiator here is that you can set up custom searches rather than just offer a search box on top of a collection of sites. You are limited to 24 search terms and Boolean "and" operators. You cannot use phrases or any of the more complex Boolean stuff in your kitbag. This isn't fatal but you might have to be creative and suffer a few more false drops. Since it follows the syndication model you can't expect on demand results. It updates daily and daily alerts can be set via e-mail.

I can envision that PSS has the potential to work for you in areas where you have a fairly unambiguous search (like a distinctive trademark) for website information and need regular info. However, I still find Technorati better for the blogosphere. PSS shows promise though.

## **Yahoo Search Builder**

This is another one of the new personalized search tools. It works like the others. You choose 5 websites and then some keywords. This builder also lets you build an unlimited news feed and filter by types of news like sports, politics

and entertainment. The Yahoo! product also lets you have some control over the display such as text size, width, color, and font. You can choose your own banner.

Yahoo! Search Builder says on its site that it can bring the power of Yahoo! Search to your site by easily add web search and site search to your site, tailoring the look and feel of the search experience to match your site and enhancing the search engine algorithm to focus on your site's topic. It does this and it does it well.

### **Eurekster Swickis**

Eurekster's Swicki has been around for a while and some think that this harkens to the future of search as the web harvest gets too big, too unmanageable and too disparate. The idea of community based search and relevance is not new but it's taking a while to emerge (as with all complex ideas). With a swicki you tap into the wisdom of crowds using social networking tools. You also can build a search environment that is informed by the team that collaborates or shares an interest in the quality of the results. This seems like a good place to look for search that is oriented to work teams, R&D groups, and special projects. The behavior of the searchers drives improvements to the swicki. The tag cloud on every swicki allows you to view the ongoing search preferences of the users and use those insights to improve the content. This is more complicated to understand than standard search but shows promise to make user-centric and team-centric search that gives a more targeted and expert driven service. It's worth keeping an eye on.

### **Conclusion**

If you read this column regularly, you already know that I value play as a key learning option for us to adapt our skills and competencies for the emerging world. As we arrive at the tipping point where human factors trump technology we must set up information professionals as guides and experts to success in this new world. We can go a long way to positioning ourselves as information fluent and technology adept professionals by playing and learning the new tools as they emerge. The ones outlined in this column are a fun place to start. Specialized search tools created to meet the needs of identified communities and special niche users could be a goldmine for you and your organization. And you will be better prepared to deal with the world of advanced federated search when your organization matures beyond simple search boxes in the middle of a white field.

*Stephen Abram, MLS is Vice President, Innovation, for SirsiDynix and the President-elect of SLA. He is an SLA Fellow and the past president of the Ontario Library Association and the immediate past president of the Canadian Library Association. In June 2003 he was awarded SLA's John Cotton Dana*

*Award. This column contains Stephen's personal perspectives and does not necessarily represent the opinions or positions of SirsiDynix. Stephen would love to hear from you at [stephen.abram@sirsidynix.com](mailto:stephen.abram@sirsidynix.com).*

,