



In Search of Intergovernmental Organizations: The Development of an IGO Search Engine

David Oldenkamp and Amalia Monroe
Indiana University Bloomington Libraries
Bloomington, USA

Meeting: 119 Government Information and Official Publications
Simultaneous Interpretation: Yes

WORLD LIBRARY AND INFORMATION CONGRESS: 73RD IFLA GENERAL CONFERENCE AND COUNCIL
19-23 August 2007, Durban, South Africa
<http://www.ifla.org/iv/ifla73/index.htm>

Abstract:

For years, librarians have attempted to cope with the increasing flood of online governmental and intergovernmental organization (IGO) documents and publications by creating and maintaining lengthy online lists of documents, agencies, and organizations. Some libraries have attempted to catalog these massive numbers of online publications. Meanwhile, regular Internet searches using popular search engines do a good job of finding specific, known item, government documents. However, users who are not searching for known items find that high quality IGO publications are scattered throughout the average search engine results list.

Relief can be found in a new generation of search tools that librarians can create to search across hundreds governmental and IGO Websites. Search engines such as Yahoo!, Google, Rollyo, PSS!, and Swiki are offering tools that allow users to create their own search engines. Seizing this opportunity, librarians are now taking their lists of links and webiographies and turning them into new search tools for the world to use.

David Oldenkamp has created an Intergovernmental Organization Search Engine using several of these tools. The most promising search tool is the Google Co-op tool, which has been used to create a search of over 350 IGO Websites. This tool has also been used to create a search engine for Canadian Government Documents (<http://www.library.carleton.ca/madgic/>) by David Sharp (Carleton University Library).

This paper looks at the current state of these customized search engine tools and evaluates the future uses of such search tools for general reference work and specifically government information libraries. Since these tools allow users to search within a selected set of Websites, the possibilities for government researchers are immense. This paper also looks at how current government information users access information, and how customized government information search engines may be used by these patrons.

Biographical Information

David Oldenkamp is the International Studies Librarian at Indiana University Bloomington's Herman B Wells Library where he is responsible for developing international document collections. David is a member of the International Federation of Library Associations, American Library Association's International Documents Taskforce and the President-elect of the Indiana Networking for Documents and Information of Government Organizations (INDIGO).

Amalia Monroe is a Master of Library of Science candidate at Indiana University's School of Library and Information Science, and she is the 2006-07 Graduate Assistant for International Studies at Indiana University Bloomington Libraries. Amalia holds a MA in Political Science from Illinois State University.

Throughout the last two decades, intergovernmental organizations (IGOs) have used the Internet as one of their primary platforms for publishing and distributing their materials to their worldwide clientele. As publications from all sectors have transitioned from print to online, people with access to an Internet connection have been able to access a phenomenal amount of information. For people interested in IGO documents and publications, this has meant that they can access massive amounts of information that they would previously have only found in major research libraries. The development of Internet search technologies, search engines in particular, has allowed users to search for their information across billions of Webpages.

Libraries have responded to the explosion of online resources through a host of methods. Librarians have created webiographies, online pathfinders, cataloged important Websites, and taught workshops on Internet searching, among other initiatives.¹ All of these strategies have been employed by librarians when it comes to connecting users to IGO information.

The Problem

In a general sense, IGO information is not fundamentally different than any other type of information. If it is online, it is searchable through the average search engine. However, as an academic librarian whose job it is to assist and instruct hundreds of students, researchers, and faculty every semester, I have found that the average user has difficulty in discerning which of their thousands of search engine-produced results come from an IGO, which come from an NGO, which come from a governmental agency, and which results are simply random Webpages. No matter how many workshops and instructional sessions that are offered, there are simply too many places where users are getting stuck, frustrated, or misled as they conduct online research. It does not help our users that the world of IGOs is so large, with hundreds of IGOs that have spawned thousands of sub-agencies and related organizations.

While there are many librarians who have created online lists of IGOs in hopes of guiding users to the appropriate sources of information, another problem arises as users attempt to use these lists. When confronted with a list of IGOs, the user must often guess which IGO has

¹ Examples of major IGO webiographies include Northwestern University Library's "International Governmental Organizations" list (<http://www.library.northwestern.edu/govinfo/resource/internat/igo.html>), and the University of Michigan Library Documents Center's "International Organizations and Related Information" list (<http://www.lib.umich.edu/govdocs/intl.html>).

the information that they are seeking. In most instances, there are many organizations working in the same area, and selecting one of the Websites on a list will not yield a complete picture of how the IGO community is addressing a particular problem. For example, if a user is seeking information on international development activities in Ecuador, relevant information could be found from dozens of organizations such as: the World Bank, the International Monetary Fund, the United Nations Development Programme, the Organization of American States, the International Fund for Agricultural Development, the Inter-American Development Bank, and the United Nations Economic Commission for Latin America and the Caribbean, etc... Assuming that a user takes the time to locate all of these organizations in a list, it would be very tedious to have to search within each organization's Website for the desired information. A user doing a general search of Google for the terms Ecuador International Development will retrieve some 2,900,000 results, with IGO resources making up a small portion of the overall total of results.

Until now, there has not been a place for a user to search across all of these IGOs without having to deal with the previously mentioned issues of doing general search engine searches. For years, Google and other search engine companies have created specialized search engines that search subsets of the much larger Internet. Examples of previously existing special Google include a special search for books, blogs, catalogs, images, Linux, Apple Macintosh, and the U.S. Government.² The Google U.S. Government Search (previously known as Google UncleSam), is a wonderful template for what a specialized search engine tool can do for searching across governmental agencies and across branches of government.

Problem Solved: Customized Search Engines

The solution to the problem of IGO searching is obvious, even if it took a while for the tools to solve the problem to become available. Ideally, a person who is interested in any topic should be able to create a list of their favorite Websites, place these Website URLs in their own search engine, which would enable them to search across their selection of Websites. For IGOs, this would mean taking existing lists of IGO URLs and creating a search engine to only search within the selected Websites.

Thankfully, several custom search engine (CSE) products have become available to the general public, and these CSEs are designed to allow anyone with an Internet connection and some free time to create search engines anyone can use. The first few CSEs that came out are Rollyo, PSS! (Personal Search Syndication), Yahoo Search Builder, and Eurekster's Swiki.³ These four CSE products have been joined by several other products, including Google's Custom Search Engine.⁴ A common feature of all of the CSE products is that all that a person needs to do create their own CSE is to create a username and password at the CSE Website, select the options on how the search engine should operate, and populate the CSE with the "seed" or "core" URLs or terms (depending on the CSE being used). The CSE product will usually provide the user with a Webpage that hosts their search engine, but the

² Links to the various Google search engines can be found at <http://www.google.com/intl/en/options/> and at <http://www.google.com/options/specialsearches.html>.

³ Phil Bradley, "Your Search, Your Way, Part 1," *SearchEngineWatch*, September 19, 2006. <http://searchenginewatch.com> (accessed on Oct. 18, 2006); and Phil Bradley, "Your Search, Your Way, Part 2," *SearchEngineWatch*, Sept. 20, 2006. <http://searchenginewatch.com> (accessed on October 18, 2006).

⁴ Resource Shelf, "Briefs: Google Joins Rollyo, Yahoo, and Microsoft Live Search in Allowing Users to Create Customized, Domain Specific Search Tools," October 25, 2006, <http://resourceshelf.com> (accessed October 25, 2006).

CSEs also provide users with the HTML coding that, when inserted into a Website, will create a search box for the CSE on any Webpages that the CSE creator chooses. Many of the CSEs allow multiple contributors for the same CSE, and the creator can determine the level of changes that contributors can make in their CSE.

Each CSE product has its own quirks.⁵ In creating an IGO Search Engine, we hoped to cultivate an exhaustive list of IGOs and then limit our CSE to search across our carefully-tailored list of organizations. Initially, we set up three experiments with the following CSE products: Google Custom Search, Swiki, and PSS!. We attempted to use Rollyo, but Rollyo has a limit of twenty five Websites, and this number is just too limiting for the large number of IGOs that we wanted to include in our search. Yahoo has similar limits on the number of sites that can be included in a search engine.

Three Experiments

When starting to create our IGO Search Engine, it was difficult to determine which CSE product would do the best job in searching across all of our IGO URLs. We decided that it would be best to experiment with several of the CSE products, and we created a Webpage where we placed our three IGO Search Engine experiments.

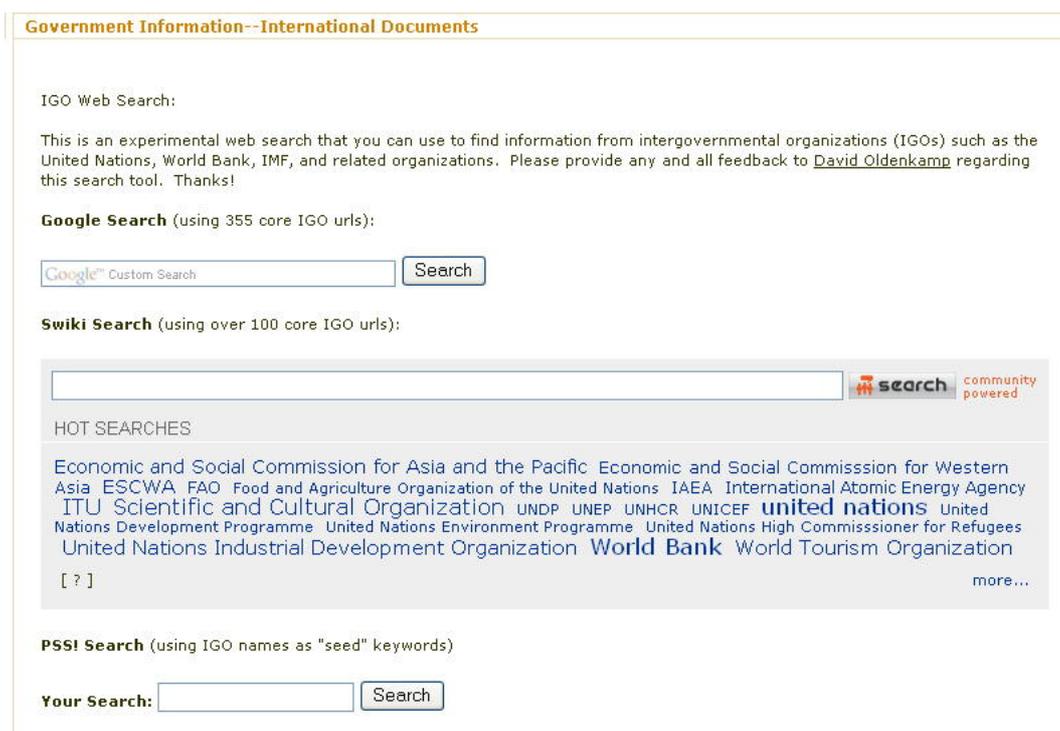


Figure 1: The Intergovernmental Search Engine Experiment Webpage:
<http://www.libraries.iub.edu/index.php?pageId=4173>

Google Custom Search Engine (Google Co-op: <http://google.com/coop/>)

Google's CSE product permits users to create a CSE with up to 5,000 URLs, so Google immediately became one of our favorite CSE options. Google's dominance in the search

⁵ A good discussion of the major CSE products can be found in the following article: Stephen Abram, "Growing Your Own Search Engine," *Information Outlook* 10, no. 12 (December 2006), 42-43. http://www.sirsidynix.com/Resources/Pdfs/Company/Abram/IOCcolumn_54.pdf.

engine market means that our users are already familiar with Google's interface. Once we created our initial Google CSE, we were able to add and subtract URLs to the list with relative ease. Google provided us with the HTML coding that can be inserted into any Webpage, which results in an IGO Search Engine box appearing on that webpage. Google also provided our own Google IGO Search Engine Webpage.⁶

The HTML coding for the IGO Search Engine has been inserted into several International Studies Class Webpages at Indiana University Bloomington, and this coding has been shared widely through the American Library Association's Government Documents Round Table (GODORT) and the International Documents Task Force (IDTF) of GODORT. When the IGO Search Engine was shared with the IDTF at the 2007 ALA Midwinter Meeting, the group endorsed the search engine and the search engine was opened to collaborators from the international documents community.⁷ Several librarians have taken the HTML coding and integrated the IGO Search Engine into their own international documents Webpages.

The Google CSE has proven to be the best CSE that we have tested. Whereas the general search of Google.com for our Ecuador search had yielded 2,900,000 results, the same search conducted in the Google IGO Search Engine yields about 388 results, all of which come from major and minor IGOs. A user doing this search will discover IGO documents and publications from a host of organizations, and the user can then follow-up with further explorations on individual organization Websites.



The screenshot shows the GODORT search engine interface. At the top left is the GODORT logo, which is a red circular emblem with an open book in the center and the text 'INTERNATIONAL • U.S. FEDERAL • STATE • LOCAL' around the perimeter. Below the logo is the search input field containing the text 'ecuador international development' and a 'Search' button. Below the search field are two radio buttons: 'Intergovernmental Organization Search Engine' (which is selected) and 'Web Search'. Below the search area, the results are displayed. The first result is 'SAR Directory - Medical Capacities' with a release authority of 'U.S. Agency for International Development ... National Directorate of Civil Defence of Ecuador ...' and a URL 'www.reliefweb.int/tocha_ol/programs/response/sarnet/medical.html - 23k - Cached'. The second result is 'Statement by Ecuador at the International Conference on Financing ...' with a URL 'www.un.org/ffd/statements/ecuadorE.htm - 9k - Cached'. The third result is 'FAO - PC 87th Session' with a URL 'www.fao.org/docrep/MEETING/004/Y6172e/Y6172e00.HTM - 45k - Cached'. The fourth result is 'Contributions: Ecuador: Galapagos Islands Oil Spill - Jan 2001 ...' with a URL 'www.reliefweb.int/rw/RWB.NSF/db900SID/SKAR-63JHQQ?OpenDocument - 19k - Cached'. At the top right of the results area, it says 'Results 1 - 10 for ecuador international development (0.08 seconds)'. At the bottom right, there is a 'Google Custom Search' logo.

Figure 2: Results from a Search of the Google Intergovernmental Organization Search Engine

Eurekster's Swiki (<http://swiki.eurekster.com/>)

⁶ The direct URL for the Google Intergovernmental Organization Search Engine is:
<http://google.com/coop/cse?cx=006748068166572874491%3A55ez0c3j3ey>

⁷ The IDTF has a webpage for the Google IGO Search Engine:
http://wikis.ala.org/godort/index.php/IGO_search. Those interested in collaborating can contact the author.

Whereas experiments with Google's CSE have been quite successful, our experiment with Eurekster's Swiki CSE product has been less successful.⁸ Swiki only allows about 100 URLs to be included in the search engine, which limits us in that we cannot achieve as comprehensive IGO coverage as we would like. The Swiki IGO Search cannot be controlled in the same way as the Google CSE. Swiki allows creators to exclude results from Websites that have not been preselected. In other words, users of the Swiki CSE will retrieve results that are not from IGOs. Librarians who have spent hours warning students away from citing Wikipedia in their research papers will find Wikipedia results in the mix.

Swiki is unique in that any user can "vote" for the results that they found to be most relevant and "vote" against results that were not as relevant. This community involvement is meant to strengthen the Swiki search results over time. However, it seems doubtful that users will spend much time rating and ranking each search result that they receive. Swiki does manage to provide some interesting search results, but it ultimately fails the job of providing users with a well-defined pool of information to be searched.

PSS! (Personal Search Syndication: <http://www.pssdir.com/>)

PSS! is an unusual search engine in that a person creates their search engine by providing up to twenty four search terms that are joined by Boolean operators.⁹ Instead of specifying URLs, PSS! CSE creators can guide the search tool through keyword selection. The result is that searches in the PSS! IGO Search Engine do not retrieve a comprehensive list of relevant IGO results. Instead, the results are a mix of some IGO documents and publications along with a variety of news and other sources. PSS! does not meet our needs very well, and we are looking to discontinue its use in the future.

The Future of Customized Search Engines in Libraries

With library users expecting to be able to access information in an online format, and with the continued momentum of IGOs publishing the bulk of their materials online, the future of customized search engines seems to be bright. These experiments with using CSEs to improve the IGO search experience have been fruitful, and our initial experiences have led us to think about how we can improve our existing IGO CSEs and brainstorm additional CSE projects that we may want to undertake.

The broad category of government information can be an excellent starting point for CSE projects, but there is nothing to stop creative librarians from creating and implementing CSE technology to meet a whole host of reference needs. Librarians are creating search engines on topics ranging from Canadian Government Documents to Library and Information Science blogs.¹⁰ CSE technology is the solution to the underused collections of lists of online resources that populate many library Websites. Many of the problems endemic to searching for IGO documents and publications can be addressed through the implementation and use of the Intergovernmental Organization Search Engine. Major and minor organizations should

⁸ The direct URL for the Swiki IGO Search Engine is: <http://igo-search-swiki.eurekster.com/>.

⁹ The direct URL for the PSS! Search Engine is:

http://www.pssdir.com/dir/category_8608/category_8613/category_8966/.

¹⁰ David Sharp at the Carleton University Library has created a Canadian Government Document Search Engine using Google's CSE (<http://www.library.carleton.ca/madgic/>), and Kevin Yezbick has created LISZEN (<http://liszen.com/>), which searches over 600 library and librarian blogs.

be able to be found, and IGO resources from across spectrum can be seen in a whole new context when IGO search results are listed side-by-side.

As the CSE technology develops and improves, librarians need to be at the forefront as we bring our significant subject, area, and technical expertise to the table. We are intimately aware of the all of the good features and all of the deficits in generic search engines. While not every reference problem will be solved through the creation of a CSE, the opportunities definitely exist for the creation of CSEs that will serve users at all levels and areas of interest. Instead of waiting for vendors to provide reference tools to meet our user's needs, librarians are capable of using CSEs to create the kinds of reference tools that we have always wanted to be able to offer to our users.