

The Internet research component of the American Foreign Policy course: Beyond random gleaning for bits of “information”

Professor Wayne A. Selcher, Professor of International Studies *Emeritus*
Department of Political Science, Elizabethtown College
Editor, WWW Virtual Library: International Affairs Resources <http://www2.etown.edu/vl>
September 2010

You are accustomed to using the Internet heavily and sometimes preferentially or exclusively for academic assignments, but are you familiar with effective Internet search methodology? The Internet provides cost-free access to valuable and practical foreign and domestic news, information, and analysis sources in many languages. Really **effective** and **efficient** research on the Internet, however, is definitely much more difficult and complicated and takes far more patience and efforts to stay current than the traditional and relatively static paper-based library research. Drifting off-point through interesting but unrelated hyperlinks is a constant temptation. Misinformation, out-of-date information, and deception are all too easy to come by.

The Internet is huge-- in September 2010, the search engine Cuil <http://www.cuil.com> claimed to have indexed 127 billion webpages. In June 2009, a Bing staffer, in Bing’s community blogs, estimated that the Internet had over one trillion pages of content. So the searcher who wishes to go beyond random or haphazard gleaning of chance bits of information or analysis must learn the basics of serious Internet research just as thoroughly as one must learn library research for printed materials. Only an informed, careful, disciplined, and patient strategy with discerning techniques can overcome the overwhelming information overload in Internet use and allow focused, thoughtful consideration, context, and analysis from the most valuable sources for the topic being researched.

Internet indexing and searching has become a highly specialized major industry in rapid change, presently trending toward natural language, visualized, clustered, more relevant, contextual, deep web, and personalized search capabilities. Web searcher behavior is being thoroughly researched and the results very much affect what you see online. Search engine optimization (SEO), or coding webpages to rank higher on search results, is an established and widely used technical and marketing skill and definitely affects the order of the retrieved results that you see. There are now thousands of local, regional, national (country-specific), global, and limited topic or file-type search engines. See <http://www.search-engines-2.com>.

The currently dominant Google search engine <http://www.google.com> is constantly being refined and augmented. It is so complicated in its features, possibilities, and changes that there are numerous websites, weblogs, and printed users’ manuals that one can consult to increase its research functionality for a given purpose. Definitely look at the advice at <http://www.google.com/help> and <http://www.google.com/help/cheatsheet.html> for elaboration on how to use Google more thoroughly. Google’s specialized sites include two fine ones for academic use: Google Scholar <http://scholar.google.com> that is limited in its searching to scholarly sources, and Google News <http://news.google.com> for searching through thousands of online news sources around the world.

Find some search engines that match your purposes and master at least one, but use several in each search, because they each yield somewhat different results. To be preferred now for general use are Google <http://www.google.com>, MSN <http://www.msn.com>, Yahoo! <http://www.yahoo.com>, and Ask <http://www.ask.com> because they all have their own (and different) indexing systems. Bing is also useful <http://www.bing.com/>. Metasearch engines such as Dogpile <http://www.dogpile.com> and Mamma <http://www.mamma.com> compile responses from several major search engines into one set of results.

Yippy <http://search.yippy.com>, Quintura <http://www.quintura.com>, and Gigablast <http://www.gigablast.com> helpfully cluster results by category. Glearn <http://www.glearn.com> allows you to search for language and country-specific content in “top results from Google, Yahoo and Bing as well as the most popular search engines for the selected country.”

About's “Web Search” <http://websearch.about.com> explains search engines and techniques, and offers a weekly newsletter. Information specialist Phil Bradley's website <http://www.philb.com> has lots of tips to help you select the proper search engine or technique for your task at hand. Marcus Zillman has posted a large series of fine cost-free online Internet guides in PDF at <http://www.whitepapers.us> that include academic and subject matter topics.

It is vital when doing research on the Internet to think in terms of a coherent research strategy while online. A common mistake is to prefer “bursts” of information (“infoclips”?) and to scan webpages far too rapidly, which will frustrate your effectiveness. Haphazard and hasty approaches are common, but produce mediocre results at best. Be sure to consider carefully the trustworthiness, bias, or reputation of the source of the information or the perspective that you include and cite. As a general (but not absolute) rule, for higher degrees of credibility, prefer sites that are educational (.edu), governmental (.gov), military (.mil), organizational (.org), and international organizational (.int) in origin. You can limit Google searches to include only any one of these types of sites—see the Google helps above.

A growing number of key periodicals commonly used in academe now post their searchable archives online without cost. These presently include the New York Times, Washington Post, Time magazine, and the Christian Science Monitor.

Not everything necessary to do your assignments well is easily available online somewhere, either free or by institutional subscription. Vesey (2005) notes that a wise academic research strategy is like a **tripod** and will always incorporate both print and electronic sources from 1) copyrighted books in paper copy, 2) copyrighted peer-reviewed journal articles in fiche and paper copy, and 3) copyrighted full-text online databases that the college library subscribes to and cost-free reputable Internet sources. It is also advisable to use longer and more in-depth analytical online sources instead of the usual shorter and merely descriptive ones, because shorter ones tend to be very focused on details or a certain point in time and are often superficial.

The major issue for most academic users of the Internet is not really a scarcity of quality web sources, but rather learning how to find the good ones out there. For those needing a broader orientation on search techniques, excellent free online tutorials on effective Internet use are available. An annotated list of quality tutorials is available from Academic Info at <http://www.academicinfo.net/reffind.html>. Several of the best free tutorials and tips websites on the Internet are linked and annotated at <http://www2.etown.edu/vl/starter.html>. Do take some time to try these out, because the skills that you learn there will help your research in all college subject matters and in your personal searches.

The U.S. government posts a huge quantity of information on American foreign policy and related international matters, including statistics. See page <http://www2.etown.edu/vl/usgovt.html> for the best sources. Of special interest to American Foreign Policy are the excellent and objective **Congressional Research Service Reports**, prepared by a special office of the Library of Congress at the request of members of Congress or their staff. Information on these reports and where to find the foreign policy ones online in PDF format is explained at the top of the above page.

For American Foreign Policy course needs, we have found that the key beginning principles and skills to observe in your Internet search and usage are the following.

1. One of the most basic skills is more effective use of search engines, with which you are already basically familiar. It is important to:
 - A. be skeptically aware of the engines' algorithmic and mechanistic methods in their inclusion and ranking of results and therefore of their weaknesses relative to human reasoning;
 - B. recognize the limitations of essentially advertisement-driven search engine companies in producing the most relevant academic results. The top results returned for your search are not automatically the best or most authoritative ones for your specific purposes;
 - C. identify top-of-page sponsored results (paid inclusion, usually advertisements) in contrast to generated ("organic") results;
 - D. master one search engine well but use several search engines for best results, plus "national" versions for results from specific countries or in specific languages. Results definitely vary by search engine. Be sure to use the advanced search page on each engine, not just the simple initial interface;
 - E. frame queries properly, vary wording of queries, and use advanced features including Boolean and appropriate "operator" terms to refine results by varying the syntax and the wording of search terms. Prefixes such as *near:*, *inurl:*, *site:*, *intitle:*, *daterange:* and many others allow considerable search refinement in Google, for example. Use of quotation marks around string of words in Google will treat the string as a phrase instead of as separate words. See <http://www.google.com/help/operators.html> for more explanations;
 - F. go well beyond the first two or three pages of results (a common error—many users do not even go beyond the first page!);
 - G. go beyond the default features of the search engine to use some of the advanced refined features that are constantly cropping up as customizing improvements;
 - H. distinguish between "vertical" versus "horizontal" search methods and their best uses; i.e., delving more deeply into a topic (say, specifics of U.S. human rights policies) as contrasted with moving "sideways" into related topics (concepts or theory about human rights in general);
 - I. avoid wandering away from the main topic through less relevant hyperlinks or distracting advertisements on a webpage, a constant temptation, especially for the easily bored.
2. There is a huge "invisible," "deep," or uncataloged portion of the Internet that search engine robots do not penetrate and integrate into their retrieved results, especially in the cases of databases and very large websites such as those of the United Nations, the European Union, the World Bank, or the International Monetary Fund. The **deep web** is believed to be far larger than the indexed portion of the Internet, so you should learn how to try to find items there, mainly through top-quality subject matter directories. See Complete Planet at <http://www.completeplanet.com> for further explanation about searching the deep web, as well as Jessica Hupp's "[99 Resources to Research & Mine the Invisible Web](#)" and Alisa Miller's "[100 Useful Tips and Tools to Research the Deep Web](#)."
3. Knowing how to find something of **real value** is more desirable than just "finding something." Sheer information or data (as disconnected bits of facts) is less useful than analysis, yet serious analysis is much harder to find than facts on the Internet. Use persistence in locating and evaluating quality and in-depth sources to avoid a two-screen scroll hit-and-run attention span.

4. There are many kinds of reliable and content-rich web sources of various sponsorships— intergovernmental organizations, governments, academic institutions, research foundations, nongovernmental advocacy groups, portals, gateways, academic databases, etc. Try to identify and favor such academically-sound sites and to search thoroughly within megasites such as those of the European Union, the United Nations, the U.S. and other governments, and sites of major think tanks to find relevant material. Almost all such sites have quality internal search facilities.
5. **Limited area search engines** search only high-quality sites in a specific subject rather than the whole Internet. Two that may prove very useful for, say, human rights topics are HuriSearch <http://www.hurisearch.org/> that searches the content on the sites of over 5000 human rights organizations in many different languages and the Meta Search Engine for Searching Multiple Human Rights Sites <http://www1.umn.edu/humanrts/lawform.html> from the University of Minnesota Human Rights Library. Google's news search page <http://news.google.com/> is an excellent and focused news search engine to retrieve world news from many sources and offers an e-mail news alert service and an archive. There is a very helpful limited area search engine at <http://zfacts.com/p/576.html> that searches (by keyword, phrase, or order code) only those sites that post Congressional Research Service Reports, high quality unclassified studies on a vast range of topics, done at the request of members of Congress. Many of these reports are on foreign affairs topics or U. S. foreign policy issues.
6. **Subject matter directories, databases, or gateways** such as the **WWW Virtual Library system** <http://vlib.org> and **Intute--Social Sciences** <http://www.intute.ac.uk/socialsciences/> are mediated by subject matter experts, virtual information specialists, or “cybrarians.” These sites index, annotate, and link key sites in a subject matter or provide a search facility that accomplishes that purpose from a database of the current content of high quality sites. Searchers thus have mediated access to optimum, refereed locations where they can seek more precisely, say, professional papers or reports that a major search engine would miss or would rank very low on the most likely search terms. Become aware of and use such directories and gateways in your field of study.
7. **Online Portable Document Format** (PDF) files are common as especially valuable “containers” for academic and research institution information such as scholarly papers and U.S. Congressional Research Service studies. Few students recognize this fact and tend instead to prefer shorter html-based information pages. Some attention should therefore be given to proper use of the Adobe Acrobat reader for PDF files. Search engines index both the titles and the contents of PDF files.
8. The Internet makes **cut-and-paste plagiarism** a strong temptation, so proper usage and citation style for online sources must be specifically learned and observed. You are expected to abide by the Elizabethtown College Pledge of Integrity in all of your work at the College, in and out of class. It is online at <http://www.etown.edu/web/policies/academicPolicies.html#pledge>. Also see the yearly booklet Academic Integrity at Elizabethtown College, used in the Freshman Seminar and available in the Office of the Dean of Student Life. **Plagiarism at Elizabethtown College is punishable by failure in the course.**

Note: Vesey, Ken. “Eliminate ‘Wobbly’ Research with the Information Resource Tripod.” Teacher Librarian. Vol. 32, No. 3 (February 2005): 35-37.