

## The Internet research component of the European Union Simulation course: Beyond random gleaning of bits of “information”

Professor Wayne A. Selcher [selchewa@etown.edu](mailto:selchewa@etown.edu)

Editor, WWW Virtual Library: International Affairs Resources <http://www2.etown.edu/vl>

September 2006

You are probably 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 Misinformation, out-of-date information, and deception are all too easy to come by.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 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 on the Internet 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. 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-- <http://www.search-engines-2.com>. The currently dominant Google search engine <http://www.google.com> indexed over 8.1 billion webpages in mid-2005 and 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 only 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 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 Jeeves <http://www.ask.com> because they all have their own (and different) website indexing systems. Meta-search engines such as Dogpile <http://www.dogpile.com> and ProFusion <http://www.profusion.com> compile responses from several major search engines into one set of results. Clusty <http://clusty.com> and Gigablast <http://www.gigablast.com> helpfully cluster results by category. Be sure to use the advanced search page on each engine if available, not just the simple initial interface, and go well beyond the first two pages of results.

About's “Web Search” <http://websearch.about.com> explains search engines and techniques, and offers a weekly newsletter. “Recommended Web Finding Tools” <http://infodome.sdsu.edu/research/guides/recommend.shtml>, from the San Diego State University library, thoroughly reviews search engines, subject directories, website reviews, and other finding tools.

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 produced many fine cost-free online Internet guides in PDF at <http://www.whitepapers.us> that include subject matter topics and how to find sources in the huge “invisible,” “deep,” or uncatalogued and largest portion of the Internet that search engine robots do not penetrate, index, and integrate into their retrieved results. Also see Complete Planet at <http://www.completeplanet.com> for further explanation about searching the deep web.

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.

Not everything necessary to do your assignments well is easily available online somewhere, 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 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.

For the European Union Simulation 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. Metasearch sites such as Dogpile <http://www.dogpile.com> and Mamma <http://www.mamma.com> draw results simultaneously from several different major search engines;

- 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. See <http://www.google.com/help/operators.html> for explanations.
  - F. go well beyond the first two or three pages of results (a common error);
  - G. distinguish between “vertical” versus “horizontal” search methods and their best uses; i.e., delving more deeply into a topic (say, specifics of EU human rights policies) as contrasted with moving outward into related topics (concepts or theory about human rights in general);
  - H. avoid wandering away from the main topic “horizontally” 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 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.
  3. Knowing how to find something of **real value** is far 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 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. Check the Human Rights and Humanitarian Affairs page at <http://www2.eto.edu/vl/humrts.html> to find such dependable sources as Amnesty International, the Human Rights Library, Human Rights Watch, the Human Rights Network, and Human Rights Internet. 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, but few persons have ever heard of them. There are no limited area search engines specifically for E.U. matters at present, but those in other areas, such as in the subject matters of the resolution to be debated, may be of help. Two excellent ones that may prove very useful for this year’s human rights topic are HURISEARCH <http://www.hurisearch.org> that searches the content on the sites of almost 1500 human rights NGOs in 58 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 European Union news search page <http://news.google.com/news?q=European%20Union&hl=en&lr=&ie=UTF-8&sa=N&tab=wn> is an excellent and focused news search engine to retrieve news about the European Union and offers an e-mail news alert service.

6. **Subject matter directories, databases, or gateways** such as the **WWW Virtual Library system** <http://vlib.org>, the **Social Science Information Gateway** <http://www.sosig.ac.uk>, and the Economic and Social Research Council's **Society Today** <http://www.esrcsocietytoday.ac.uk> (all searchable) are mediated by subject matter experts, virtual information specialists, or "cyberlibrarians." 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 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. The European Union page <http://www2.etown.edu/vl/eurunion.html> of the WWW Virtual Library: International Affairs Resources <http://www2.etown.edu/vl> will be especially useful to you, but other pages of that Internet directory will point you to good sources with relevant analytical information.
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 persons 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 College Life.

**Note:** Vesey, Ken. "Eliminate 'Wobbly' Research with the Information Resource Tripod." Teacher Librarian. Vol. 32, No. 3 (February 2005): 35-37.