

Research Strategies (abridged) - 5

CONTENTS:

▶ Home	▶ 5 - Internet Research
▶ Preface	▶ 6 - Other Resources
▶ 1 - Taking Charge	▶ 7 - Case Studies
▶ 2 - Databases	▶ 8 - Learning to Read
▶ 3 - Information Fog	▶ 9 - Organizing Notes
▶ 4 - Periodical Maze	▶ 10 - Research Writing

[PDF of this Page](#) 

[How to Get the Paper Edition](#)

RESEARCH STRATEGIES - WILLIAM BADKE, COPYRIGHT 2008

Note that chapter order in the print edition will differ from the above. The online version will retain the former chapter divisions.

5

Internet Research

[The material below is now is Chapter 6 of the print edition. It has been condensed for the online version, and several sections have been omitted].

The early 1990s saw the beginning of an information revolution as dramatic as the invention of the printing press in 1440 (though, typical of our narrow Western mindset, we usually ignore the fact that the Chinese people had a printing press 400 years earlier). The printing press made it possible to replace the normal copying of documents by hand with a process that produced multiple copies in much more rapid succession. All of a sudden, the availability of written information increased dramatically.

The World Wide Web in the 1990s created the same kind of revolution, but on a larger scale. An information delivery system primarily concerned with paper and print (libraries, bookstores, hard copy journal subscriptions and so on) became an electronic universe available to anyone with a computer and a service provider. Today, the average student gets well over two-thirds of his/her information from the Net.

Please supply your own answer to the following question: **The Internet is a world where:**

Stumped? Let me supply some possibilities.

The Internet is a world where:

- you can truly interact with the rest of humanity.
- you can find information on virtually any topic.
- almost everyone is trying to sell you something.
- you can surf until your eyes fall out.
- anything that's worth anything costs *money* to retrieve (or else you're told that access is "forbidden!").
- you can meet new friends or pick up predators.
- you can't trust anyone or any piece of information.

- you can get your questions answered.
- you can waste a lot of time.

Maybe your response should be “all of the above.” The Internet is wonderful and frustrating, helpful and potentially dangerous, beneficial and a waste of time. Why bother with it then? Because it has become the common denominator of our daily lives.

What do I mean? All of us know that most transactions today are becoming digital—we use bank machines and online banking, we use computer library catalogs, we find the Net to be a venue for social interaction, and many of us are shopping from home over the electronic highway. As the phenomenon of doing things electronically grows, the only common medium that can meet its needs is the Internet. The Net has become the basis, the common denominator, for the things we need to get done. It is the number one vehicle for all informational research, whether through a search engine or a specialized subscription database.

True, there are now calls to start over with this thing, using what we have learned to create a new World Wide Web that works more efficiently. But no one is advocating that we abandon the whole idea. Thus we'd better try to understand it.

A Brief Introduction to the Net

The Internet has been around for a quite awhile, but most people weren't aware of it until the early 90s. Back in the late 1960s, the US military developed a worldwide computer network in order to remain in communication with everybody involved in the space program and defense research. This network (really a network of networks) eventually came into the hands of non-military people—scientists, computer buffs, and so on. It lacked a common communication language that was easy to use, so only specialists could profit from it. In the 1990s, a common communication language and a common communication protocol were established so that anyone who had access to this network could move around it with ease. This resulted in the formation of the World Wide Web.

The following sites on the Internet will provide you with good information about using it for research purposes:

<http://www.learnthenet.com/english/index.html>

This one, called *Learn the Net*, is very informative and user friendly. It also has a lot of links to other information about the Net.

<http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/FindInfo.html>

This is a tutorial with links from the University of California, Berkeley.

Disclaimer: In the pages that follow, several Internet URLs will be given. Addresses change rapidly on the Net, so I make no guarantee that you will find what you're seeking by inputting the URLs given.

The Language of the Net

Let's consider some terminology:

Browser—a program that gives you access to the Internet and its searching tools as well as allowing you to manipulate and download data. Two common examples are *Internet Explorer* and *Firefox* (or *Safari* for you Mac users).

Home Page—the first screen you come to when you open your browser. It is also the front door or first room in any Internet site.

URL—Uniform Resource Locator, a specific “address” for a database or information site available through the Internet. It's analogous to a telephone number that helps to link your phone with someone else's phone, except that, for an URL, you are linking computers. The typical URL begins with **http://** then has a distinctive alpha-numeric description in which there are no spaces. Thus, you can find the ERIC database if you go to the URL: <http://www.eric.ed.gov>

Search Engine—the World Wide Web is a network of thousands (millions?) of computers linked, essentially, through telephone lines or other cable systems. Each database held by one of these computers has an “address”. But if you don't know a specific address yet you want information on a topic, a “search engine” is what you need. The average search engine asks you to input one or more keywords. The search engine then goes through thousands of networked computer databases (actually through the search engine's own indexed snapshot of the Internet), looking for your keyword(s). You are presented with a list of relevant database sites, with some summary information for each.

World Wide Web—often described as WWW. It is an organizational system for sites on Internet, allowing a user with a search engine to find information from many databases using a common language (HTML) and a common method of communication (HTTP).

HTML—the common language of the World Wide Web. It stands for *HyperText Markup Language*. This is the language in which data in the World Wide Web is written.

HTTP—*HyperText Transfer Protocol*, a formatting method to determine content and means of transfer of information on the Net. If we were to use the phone system as an analogy, http would be like the method that phone companies worldwide have agreed upon to make sure a phone in Nigeria can communicate with a phone in England. Http is the protocol that is used within the World Wide Web to allow communication among Web sites. Most Internet URLs begin with "http."

Document—on the Net, a document is any collection of information that you locate in your search (home page plus related pages on one Internet site).

Link—a feature of the World Wide Web that allows you to move from one document to another by way of a *hyperlink* or just plain "link." A link in a document is a word or phrase with colored letters, sometimes underlined. With it, you can move from document to document without typing in a lot of addresses.

Bookmark—your browser has a bookmark or "favorites" function that allows you to record the addresses of sites on the Internet that you want to return to. When you want to return, you can click your mouse on that bookmark, and you will be taken to that site without needing to type an address. You can use <http://del.icio.us/> to organize and share bookmarks, if that is your desire.

Google Scholar and other Free Academic Search Engines on the Net [more print edition]

Free Journals—An Academic Myth

While it is true that many databases provide full text, most of them require a subscription by an institution as well as password protection. As far as genuinely free journal articles are concerned, there are still limitations on the titles and scope of journals available. [more in print edition]

Google Scholar (<http://scholar.google.com>)

Google Scholar is not really a subset of the Google search engine, but a resource devoted to scholarly literature—books, journal articles, conference proceedings and academic Web sites found on the WWW. [more in print edition]

Microsoft Academic Search (<http://academic.research.microsoft.com/>)

Following upon the failure of Windows Live Academic, Microsoft has a new academic search tool. Its database is still smaller than Google's but it does have a few unique features. For more information on features, see <http://academic.research.microsoft.com/About/Help.htm>.

CiteSeer (<http://citeseer.ist.psu.edu/>)

CiteSeer is a search engine for academic scientific information that is freely

available on the WWW. [more in print edition]

Scirus (<http://www.scirus.com/srsapp/>)

This site bills itself as "the most comprehensive science-specific search engine on the Internet."

getCITED (<http://www.getcited.org/>)

Academic users of this search engine submit their own material or citations to it and have the ability (if properly signed in) to revise anything in the database. Think of it as the Wikipedia of academic search engines.

Search Engines for the Rest of Humanity—Google and Friends

Searching by Search Engine, Using Keywords

In the most common Internet research situation, you want information on a topic, but you don't have a specific address. This is where a search engine can help you by taking the keyword(s) you input and searching the WWW for data that is relevant.

[more in print edition]

A Basic Introduction to the Best Search Engines [more in [3rd print edition](#)]

Google (<http://www.google.com/>)

Google also offers separate search functions for books (Google Book—<http://books.google.com>), scholarly material (Google Scholar—<http://scholar.google.com>), blogs (Google Blog Search—<http://blogsearch.google.com>), things to buy (Google Product Search—<http://www.google.com/prdhp?tab=wf&ie=UTF-8>), and so on. All of these are found under the “more” link on the Google home page, or you can put icons to any of them into your Google toolbar.

Yahoo (<http://search.yahoo.com/>)

Windows Live (<http://www.live.com/>) - Now replaced by Bing

Ask.com (<http://www.ask.com/>)

The New Semantic Search Engines [more in print edition]

Searching by Subject Tree

All information exists within hierarchies. For example, cell phones are a subclass of telephones which are a subclass of electronic communication devices, which are a subclass of all communication devices. Information hierarchies form a tree-like structure. There are certain sites on the Internet where you can search down various hierarchies or subject trees from more general categories to specific ones.

For example, Google, while majoring on keyword searches, also has a Directory (<http://www.google.com/dirhp>) that leads you to a list of categories like Art, Business, News, Science, and so on. Click on any one, and you can follow a hierarchy down to more specific information.

Portals [more in print edition]

A growing area of significant development on the Internet is the “portal,” a site that serves as an introduction to important Internet sites on a subject. Typically, its main feature is a collection of links to sites that have been checked out for quality. What you end up with is a hierarchical way of searching the Net without a search engine. Its advantage is that someone has evaluated the sites, so you have a better chance of finding material that you can actually use.

Portals are organized by subject.

But how do you locate such sites in the first place? Here’s where a more general portal can help. A general portal is usually much larger, and often serves, at least in part, as a doorway to more specific subject portals. For example:

- Infomine: Scholarly Internet Academic Collections (<http://infomine.ucr.edu/>) boasts over 100,000 academically valuable sites, organized by broad subject categories or searchable by its internal search engine. Many of the sites it lists are actually subject portals.
- Multimedia Educational Resource for Learning and Online Teaching (<http://www.merlot.org/>). This is a project of several universities and academic organizations to organize and peer review Web sites valuable for higher education. You can search it directly or hierarchically by subject categories.
- Academic Info (<http://www.academicinfo.net/>) is an educational subject directory to programs of study and test preparation sites. It includes a section of subject gateways to most disciplines.
- IPL2 (<http://www.ipl.org/>) is searchable directly or through hierarchies.

The Hidden Internet [more in print edition]

Evaluating Information from the Internet

Let's be realistic for a moment (aren't we always?) and ask the question: *Why would people with data want to put it up for free on the Internet?* The answers are varied:

- They want to sell you something, and their web page is simply advertising or a doorway to a credit card purchase.
- They have something they want to say, and this a cheap and easy way to do it. Here you can have anything from "Hi, I'm Tim, and here are some pictures of my iguana Frank," to "I was abducted by Martians, and I want to warn the world before they destroy us all." Here too are the blogs and social networking sites like MySpace and Facebook.
- A government or public agency that would normally not charge for its information wants to make it available. Here you can find everything from the ERIC Database to census data, and so on.
- An educational body sees providing information as part of its mandate. Here you may get open access journal articles, electronic editions of out-of-print books, guides to this and that, occasionally even electronic dictionaries or encyclopedias.
- Sincere scholars and other individuals who have valuable information to share and want to make that information freely available for the edification of everyone.

But take careful note of one foundational rule of life: **Few people, except those related to all but the first reason above, provide information for free on the Net unless they can't find anyone to pay for it or they have the financial resources to give it away freely.**

What does that mean for you, the anxious Internet searcher? It means several things:

- A lot of what you hoped might be on the Net is not there or you have to pay for it—the full texts of most journal articles, the full texts of recent books, a vast array of reference tools, etc.
- Most keyword searches through search engines also bring you a ton of less than wonderful Web sites that you have to weed through to find a few gems.
- It's difficult to evaluate the quality of the material you find on the Net.

Let's consider this latter point a bit more closely. In normal publishing, there are gatekeepers to make sure that material that is inferior doesn't get published (at least we hope there are gatekeepers). On the Internet, anybody who wants to say anything has the chance to say it. Unless it is criminally obscene, violently racist or a clear and present danger to society, no one challenges it. Thus people can tell lies on the Net, and they probably won't suffer any nasty consequences (at least not in this life).

So what happens when you download a document that has no author clearly named but seems to be reliable information about B.F. Skinner's behaviorism? How do you determine whether it's good or bad information? Here are some clues:

- Look for the name of an author and/or organization responsible for the information. One way to do this is to recognize that Internet URLs are hierarchical and that the slashes in an address (/) define levels of the hierarchy. For example, the Web site advertising my book, *Beyond the Answer Sheet: Academic Success for International Students* (an example showing my penchant for blatant self promotion) has the URL http://www.acts.twu.ca/lbr/answer_sheet.htm. If I chop off the portion [/lbr/answer_sheet.htm](http://www.acts.twu.ca/lbr/answer_sheet.htm), I have the root left: <http://www.acts.twu.ca>, which is the educational institution that has allowed me access to their Web site because I am one of them. Chopping back on a URL will often lead you to the source behind the document you're interested in, though sometimes it just reveals the name of a person's Internet service provider, which, if it's some generic vendor of web space, will give you no help at all.
- Look for signs of scholarship—good language level, analytical thinking, bibliography and/or footnotes, logical organization.
- Look for signs of a *lack* of scholarship—lots of opinion without the support of evidence, indications of paranoia (as in *somebody's out to get us*, or *we're victims of a conspiracy*), poor spelling and grammar, lack of references to other sources, poor organization.
- Ask yourself—does this person have a vested interest in promoting a viewpoint or is he/she simply sharing information? Vested interests may be OK as long as you are well aware of what they are. A site selling Toyotas is going to be different in its very nature from a site offering independent reviews of Toyotas.

Ultimately, you will have to evaluate the information itself. Does it make sense? Does it ring true? Is there sufficient backing for viewpoints presented?

Have you or your professor ever heard of the people involved? Remember that Internet data may lack all the proper signposts of good scholarly work and yet still be valuable. On the other hand, it may have footnotes and a bibliography but be a racist rant. For proper evaluation, the buck stops with you actually reading the material and making sense of it.

If you want some help learning how to evaluate Internet resources, take The Internet Detective, a free Web site tutorial with quizzes. It's very cool. (<http://www.vts.intute.ac.uk/detective/>).

Some Internet Addresses Valuable for Research Purposes

In presenting the following addresses, I must warn you again that addresses, like phone numbers, go out of date fairly quickly. On the Net, something you found today might not be there tomorrow. So, some of these addresses may not work. If you have a problem, try searching for the title of the source, using a search engine.

Reference Sources

Refdesk, an amazingly detailed site that bills itself as "The single best source for facts" (<http://www.refdesk.com/>)

A biography encyclopedia: <http://www.biography.com/>

6.9.2 Searchable Library Catalogs

Libdex: <http://www.libdex.com/>

. (For this one, it's best to click on the "Country" link).

Library of Congress searchable catalog: <http://catalog.loc.gov/>

WorldCat: <http://www.worldcat.org/>

6.9.3 Directories

Switchboard (addresses & phone numbers—US plus Yellow Pages):

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

Canada 411 (Canadian phone numbers & addresses): <http://www.canada411.ca/>



[For a study guide to this chapter along with practice exercises (and key) and assignment, see the print edition of this book]

[TOP OF PAGE](#)

[ON TO CHAPTER SIX](#)

Updated June 22, 2008