

Chapter 9

More Google

Google features much more than a Web search engine. This chapter discusses some of Google's other tools.

9.1 Google Directory

A search engine index is a computer constructed and ranked database of pages automatically crawled from the Web. A directory is a human hand-selected and hand-categorized collection of Web pages. In this computer versus human comparison, the computer easily outmatches the humans for quantity, while the humans know how to pick out and organize some really valuable sites. As a result, you have two different ways to find content on the Web. The Google search engine has indexed a few billion pages, while the volunteers at the Open Directory Project (the source of Google's Directory pages) have selected a few million pages.

A powerful advantage of the Google Directory is that Google uses its PageRank™ technology to order the pages, producing a results list that is both hand selected and ordered by perceived importance (or at least popularity). While the regular Google Web search includes results from the Google Directory, a Directory search is restricted only to Directory entries.

Tip 9.1 Use Your Head

The Google Directory can be searched two ways. First, you can enter your query in the search box just as you would with a regular Google Web search. Or, you can drill down through the directories. The challenge of drill-down is knowing how a particular topic has been categorized. Look at the Jack the Ripper example below. Not everyone would have known to start with "Society," the first category, to find out who the suspects are. The best advice, then, is to use the search term approach. Start your search by clicking on "directory" and then create your own queries for the directory search box. Enter your term, and Google will provide the directory and subcategories, from most general to most specific, in the second line of each result.

Go to the first result and click on the first category (the broadest) listed. That will bring up the "Directory" page (see below). You can move along through the subcategories by clicking on the next subcategory in your original list, the subtopic listed in the boldest type. Other large subcategories of interest will also be in slightly lighter boldface type.

Below is a screen capture of part of the Google Directory results page, showing the first result of the query.



In this example, your choices include the following:

- If you click on the page title “Casebook: Jack the Ripper—Main,” you will be taken to that Web page.
- If you click on “Similar pages,” you’ll get other pages in the Directory from the same and other categories in the Directory. Consider this a “wide sweep” approach that broadens your search by looking for more general but thematically related pages.
- If you click on the “Society” hyperlink under “Category,” you’ll be taken to the Directory itself, to a page that includes the “Casebook” page, together with other pages chosen by the directory editors. A Directory page looks like the screen capture below.



This Directory page includes the following information:

- Each selected Web page includes a green bar to the left of it, indicating its relative popularity based on the PageRank™ system. To access a page, simply click on the hyperlink as you would with a regular results page.

- Under the hyperlinked page title and URL is a description of the page, written by the directory editors. Instead of snippets of text from the page, this description is a summary of the page's content.
- Under the search box are two radio buttons (providing alternate choices). If the Directory category contains a large number of entries, you can search within your results by adding keywords in the Directory search box and leaving the default button selected (in this example the default is "Search only in Jack the Ripper").
- Beneath the category title for the one you are currently viewing may be one or more "Related Category" listings. These categories contain pages with content similar to your chosen category, but located in a different part of the Directory.

The Google Directory is an excellent place to begin a search if you have a general topic and want to learn more about it before narrowing your search to a specific problem or issue. Even though the Directory has only a few million entries, they are almost always well chosen and will get you started.

9.2 Google Images

In addition to Web pages, the Google database includes several hundred million images, searchable by keyword. After clicking on "Images," you can use the same kind of search query you would for Web pages, and the Image Search box supports the same advanced operators (such as site: or inurl:) as the regular search box. Your search results will be displayed as a set of thumbnail images, together with a description of each image, including its file name, dimensions in pixels, size in kilobytes, and the URL of the originating page. Pay attention to the file extension (jpg, gif, or png) if you are interested only in certain file types.

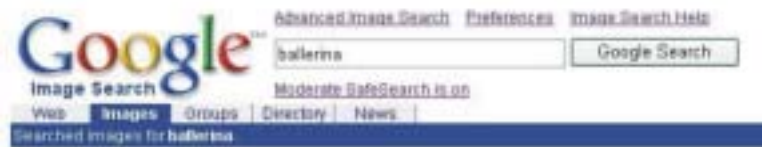
Scroll through the results. When you find an image of interest, click on it and Google will display the image (still scaled down) on a new page, with the original context of the image in a window below. You can scroll through the lower window to examine how the image has been used on the originating page. If you want to save an image, click on the image itself or on the hyperlink "See full-size image" under the scaled down image. The image will be displayed in its full size. To save the image, right click and choose "Save image as."

Tip 9.2 Mousing the Images

Many pictures and other graphics you will find do not include titles with them on their Web pages. However, there is a way to find a title for each image to use in your bibliography. Many Web pages include

a text description of each image for the benefit of users who have graphics turned off in their browsers. These descriptions are included in the alt (for alternate) tag. To read the alt tag of an image, use a mouseover. That is, just move your mouse pointer over the image and wait a moment. The description that appears can be used as the image's title. If there is no alt tag, no description will appear. In that event, right click, choose Properties, and you will see the actual file name of the image. For your bibliography, use the alt tag title if available, or the file name if that is all you can get. (Note that while you are using Google's Image Search, you can mouseover the image on the page in the context window.)

Below is a screen capture of the top of an Image Search results page.



The “Advanced Image Search” link takes you to a page that allows you to create a very specific search. You can restrict your search in these ways:

- Find an image related to all of the words, an exact phrase, any of the words, or not related to certain words.
- Locate images of any size, icon-size, small, medium, large, very large, or wallpaper size.
- Locate images of any file type, or restrict results to only jpg, only gif, or only png.
- Locate images in any colors, black and white only, gray scale only, or full color only.
- Locate images in a specific site or domain.
- Apply no filtering, moderate filtering, or strict filtering. This choice allows you to screen out adult content images. The default is moderate filtering, which works quite well. If a particular query returns too many adult-content images, you can increase the filtering to strict. (Your choice will remain in effect during the remainder of your search session.)

The “Moderate SafeSearch is on” indicator shows that moderate filtering is in effect. You can change this setting either by going to the Advanced Search page as described above or by clicking on the “Moderate SafeSearch is on” hyperlink.

Tip 9.3 Get the Image Information

When you save or print a Web page, the URL automatically accompanies the document. Not so when you save an image. Then all you get is the file name of the image itself. But you'll need the page title and URL of the image for your citation and bibliography. Remember that images need to be cited, too, just like words. Therefore, it's a good idea to keep a word processing file for your images, where you can not only copy and paste each image, but also its URL and the other page information needed for your citations.

9.3 Google News

When you perform a regular Google search of the Web, if your topic words have been in the news recently, Google will display three links to news stories about that topic. You can also perform a search looking only for news articles. To see the difference, run a Google search in the regular Google Web search box on *supreme court*. You'll see three links labeled "News" followed by the results pages from the Web. Now click the News tab and your query words will be used to search only the news. All the results will be news stories.

Google's computers crawl several thousand news sources continuously and update the news results regularly, so that each time you perform a search on a breaking event, you will get updated reports.

Whenever you want to follow current events, you can go either to the News home page and see if the story is listed, or to the News search and enter your query words there. News stories on the results page are arranged by relevance, but they can also be sorted by date, with the most recent listed first. (Click on the "Sort by date" hyperlink near the top right of the results page.)

Tip 9.4 Save That News

Many news sites (newspapers and magazines especially) allow access to their articles free, but for a limited time, often seven days. Therefore, if you perform a news search related to your topic and locate one or more articles of interest, be sure to save or print that article. If you wait and come back more than a week later, you may be asked to pay to visit the archive. Google saves its news links for 30 days, so that may give you some extra time, but if you return after that, you may not find the link. (Try a regular Google Web search in that case, to see if the content is still available on the original site and hope it is still free.)

The News is offered in several regional versions, including the United States (the default), United Kingdom, Canada, Australia, India, and New Zealand. If you are following a story from one of these other countries, you may want to select that version before performing your search, so that your results will include more news sources from that country.

9.4 Additional Features

Here is a brief look at several other information sources available through Google.

Translation. Web page results in foreign languages include a link that allows you to “Translate this page.” In addition, the Google Web search home page and results pages have a “Language Tools” link that takes you to a page where you can translate text from one language to another. You can type in, cut and paste, or name a URL and translate to or from English, German, Spanish, French, Italian, or Portuguese. (Yes, your French professor knows about this.) The translations are what you would expect from a computer, often awkward and sometimes even a challenge to make fully sensible, but they are usually good enough for you to get the gist of the information and sometimes surprisingly good.

The “Language Tools” page also allows you to:

- Search for pages in a specific language.
- Search for pages from a specific country.
- Set the Google interface language to any of about 35 choices, including not just Malay and Uzbek, but Elmer Fudd and Klingon.
- Visit the Google site in a foreign country.

Maps. You can get a map of a city or location by typing in the appropriate information in the regular Google Web search box. For example, if you type in Albuquerque, one of the first links will be “Show map of Albuquerque NM on Yahoo! Maps—MapQuest,” giving you a choice of two map sites. (Note that MapQuest has the option to see an aerial photograph also, on which you can zoom in or out.) If you type in a specific address, such as the address of the White House, the map link will be to that address, as in “Show map of 1600 Pennsylvania Ave Washington DC on Yahoo! Maps—MapQuest.”

Groups. Google has indexed the archives of Usenet Newsgroups, an interactive discussion forum with more than 700 million postings over the past 20 years. You can access Groups by going to a particular group (there are thousands) or by performing a keyword search. The “Advanced Groups Search” page allows searching by subject, author, language, and date.

Tip 9.5 Use the Groups with Caution

Newsgroups suffer from a bad reputation for reliability, much of it deserved, so be careful about using information from them in your researched writing. Some of the groups involve conversations among experts, and some groups are moderated (edited) to keep out as

much junk and irrelevance as possible, so there is good content available. But beware:

- Threaded discussions like Usenet Newsgroups are sometimes difficult to follow. With various snippets of previous postings included in each one, it may be difficult to know to whom to attribute an idea, unless you painstakingly locate the original post.
- Many postings include copy and paste pieces of articles from other people. Many of these may be taken out of context, and there is no guarantee that they are correct copies.
- Newsgroups are traditionally a free-for-all, where you will find wild ranting and invented facts. Credibility is always a question.

Special searches. Searches restricted to specific universities, computer platforms, or the U. S. Government are available through the “Services and Tools” link on the Google Home Page. Access to several other tools (shopping, catalogs, language, and so forth) is also available from this page.

© 2003 by McGraw-Hill/Dushkin, Guilford, CT 06437, A Division of The McGraw-Hill Companies.

Copyright law prohibits the reproduction, storage, or transmission in any form by any means of any portion of this publication without the express written permission of McGraw-Hill/Dushkin, and of the copyright holder (if different) of the part of the publication to be reproduced. The Guidelines for Classroom Copying endorsed by Congress explicitly state that unauthorized copying may not be used to create, to replace, or to substitute for anthologies, compilations, or collective works.