





Learn

# Google™

**Michael Busby**



# **Learn Google™**

**Michael Busby**

**Wordware Publishing, Inc.**

**Library of Congress Cataloging-in-Publication Data**

Busby, Michael.

Learn Google / by Michael Busby.

p. cm.

ISBN 1-55622-038-3 (pbk.)

1. Google. 2. Web search engines. I. Title.

TK5105.885.G66B87 2003

025.04—dc22

2003019756

CIP

© 2004, Wordware Publishing, Inc.

All Rights Reserved

2320 Los Rios Boulevard  
Plano, Texas 75074

No part of this book may be reproduced in any form or by any means  
without permission in writing from Wordware Publishing, Inc.

Printed in the United States of America

ISBN 1-55622-038-3

10 9 8 7 6 5 4 3 2 1  
0310

Google is a trademark of Google Inc.

All screen shots used in this book copyright 2003 Google Inc.

All brand names and product names mentioned in this book are trademarks or service marks of their respective companies. Any omission or misuse (of any kind) of service marks or trademarks should not be regarded as intent to infringe on the property of others. The publisher recognizes and respects all marks used by companies, manufacturers, and developers as a means to distinguish their products. This book is sold as is, without warranty of any kind, either express or implied, respecting the contents of this book and any disks or programs that may accompany it, including but not limited to implied warranties for the book's quality, performance, merchantability, or fitness for any particular purpose. Neither Wordware Publishing, Inc. nor its dealers or distributors shall be liable to the purchaser or any other person or entity with respect to any liability, loss, or damage caused or alleged to have been caused directly or indirectly by this book.

All inquiries for volume purchases of this book should be addressed to Wordware Publishing, Inc., at the above address. Telephone inquiries may be made by calling:

(972) 423-0090

## **Dedication**

For Shane, Drew, and Stuart Busby.

This page intentionally left blank.

# Contents

Preface: Searching for and Finding the Golden Fleece . . . . .	xi
Acknowledgments . . . . .	xv
<b>Chapter 1 All about Search Engines . . . . .</b>	<b>1</b>
Introduction . . . . .	1
What Do We Search For? . . . . .	3
How Much Do We Search? . . . . .	6
Why Search? . . . . .	6
What Is the Web? . . . . .	8
What Is a Web Page? . . . . .	12
What Is a Search Engine? . . . . .	16
Search Engine History . . . . .	16
Robots, Spiders, and Metacrawlers . . . . .	20
Relevancy Ranking . . . . .	21
How Does a Search Engine Work? . . . . .	24
Issues Searching the Web . . . . .	27
Search Integrity . . . . .	28
The Race to Be Number One . . . . .	28
What Is Google? . . . . .	30
How Google Ranks Its Search Results . . . . .	31
How Is Google Getting Better? . . . . .	32
The Next Generation . . . . .	33
Summary . . . . .	33
<b>Chapter 2 Google Services . . . . .</b>	<b>35</b>
Introduction . . . . .	35
Google Services . . . . .	41
Froogle . . . . .	41
Google Answers . . . . .	47
Google Catalogs . . . . .	50
Google Groups . . . . .	58
Google Image Search . . . . .	58
The Hard (Google) Way . . . . .	68
The Wrong (Bright Spark) Way . . . . .	69
The Easy (My) Way ... That Works! . . . . .	71
Google Labs . . . . .	74
Google News Alerts . . . . .	74
Google Compute . . . . .	79

Google Viewer . . . . .	85
Google WebQuotes . . . . .	86
Google Glossary . . . . .	88
Google Sets . . . . .	90
Google Voice Search . . . . .	92
Keyboard Shortcuts . . . . .	92
Google News. . . . .	94
Google Special Searches . . . . .	95
Google University Search. . . . .	97
Google Web Directory. . . . .	99
Google Web Search . . . . .	100
Wireless . . . . .	104
Summary . . . . .	105
<b>Chapter 3 Google Preferences . . . . .</b>	<b>107</b>
Introduction . . . . .	107
Preferences . . . . .	107
Interface Language . . . . .	111
Search Language. . . . .	112
SafeSearch Filtering . . . . .	114
Number of Results . . . . .	118
Results Window . . . . .	118
Save Preferences. . . . .	119
Summary . . . . .	119
<b>Chapter 4 Basic Google Search Techniques. . . . .</b>	<b>121</b>
Introduction . . . . .	121
Keywords . . . . .	125
The Quote ("") Operator . . . . .	127
The Arithmetic (+, -) Operators. . . . .	130
Boolean Operators . . . . .	134
Complex Boolean Operations . . . . .	136
Search Issues . . . . .	140
Summary . . . . .	141
<b>Chapter 5 Advanced Google Search Techniques . . . . .</b>	<b>143</b>
Introduction . . . . .	143
Advanced Search Features . . . . .	146
Find Results . . . . .	147
File Format . . . . .	148
Date . . . . .	150
Occurrences . . . . .	150
Domain . . . . .	153
Language Tools and Services . . . . .	158

Google in Your Language . . . . .	165
Searching for Hilux: A Real-life Search Example . . . . .	169
SafeSearch Filtering . . . . .	175
Similar . . . . .	175
Links . . . . .	177
Topic-Specific Searches . . . . .	177
Summary . . . . .	178
<b>Chapter 6 Advanced Alternate Query Search Operators. . . . .</b>	<b>181</b>
Introduction . . . . .	181
Advanced Alternate Query Search Operators . . . . .	182
cache: . . . . .	182
link: . . . . .	190
related: . . . . .	191
info: . . . . .	193
stock: . . . . .	194
site: . . . . .	195
allintitle: . . . . .	198
intitle: . . . . .	199
allinurl: . . . . .	201
inurl: . . . . .	202
Summary . . . . .	206
<b>Chapter 7 Advanced Image Search. . . . .</b>	<b>207</b>
Introduction . . . . .	207
Advanced Image Search. . . . .	210
Find Results . . . . .	210
Related to All of the Words . . . . .	212
Related to the Exact Phrase . . . . .	216
Related to Any of the Words . . . . .	218
Not Related to the Words . . . . .	219
Size . . . . .	220
Filetypes . . . . .	222
Coloration. . . . .	223
Domain . . . . .	223
SafeSearch . . . . .	224
Summary . . . . .	224
<b>Chapter 8 The Google Toolbar . . . . .</b>	<b>225</b>
Introduction . . . . .	225
System Requirements . . . . .	226
Google Toolbar Privacy Policy . . . . .	226
Toolbar Installation . . . . .	227
Toolbar Help . . . . .	228

Drag-and-Drop Functionality . . . . .	229
Right-click Functionality . . . . .	229
Toolbar Features . . . . .	229
Google Toolbar Menu . . . . .	230
Google Search . . . . .	230
Search Site . . . . .	231
PageRank . . . . .	232
Page Info . . . . .	233
Up . . . . .	233
Highlight . . . . .	233
Toolbar Menu . . . . .	234
Google Home Page. . . . .	236
Google Images . . . . .	236
Google Groups . . . . .	236
Google Web Directory . . . . .	236
Google News . . . . .	236
Google Answers . . . . .	236
Advanced Search Page . . . . .	237
Search Preferences Page . . . . .	237
Language Tools Page. . . . .	237
Toolbar Options . . . . .	237
General Options . . . . .	238
Search Box . . . . .	239
Buttons . . . . .	241
Search Buttons . . . . .	241
Advanced Features Are ENABLED . . . . .	242
Page Information . . . . .	242
Navigation . . . . .	243
Finding Words within a Page . . . . .	243
Google Compute Is ENABLED . . . . .	244
Experimental Features . . . . .	244
Default Layout . . . . .	245
Clear Search History. . . . .	246
Help . . . . .	246
Privacy Information . . . . .	247
About Google Toolbar . . . . .	248
Contact Us . . . . .	248
Uninstall . . . . .	248
Toolbar Version 2.0 . . . . .	248
Summary . . . . .	254

---

<b>Chapter 9 Other Google Features . . . . .</b>	<b>255</b>
Google Help Central . . . . .	255
Google Site Map . . . . .	256
All About Google . . . . .	258
General FAQ . . . . .	259
Saving Preferences . . . . .	260
Accepting Cookies . . . . .	260
Accepting Cookies from Specific Sites . . . . .	262
Setting Google as Your Default Home Page . . . . .	263
Netscape 4.0 to 6.0 . . . . .	263
Internet Explorer . . . . .	263
America Online . . . . .	263
Setting Google as Your Default Search Engine . . . . .	264
Netscape Communicator . . . . .	264
Windows Operating System . . . . .	264
Macintosh Operating System . . . . .	265
UNIX Operating System . . . . .	265
Internet Explorer . . . . .	265
Windows Operating System . . . . .	265
Macintosh Operating System . . . . .	267
Mozilla/Netscape 6 Search Options . . . . .	267
IE QuickSearch . . . . .	267
Macintosh OS X Service . . . . .	268
Additional Google Web Search Features . . . . .	268
Calculator . . . . .	269
Dictionary Definitions . . . . .	271
I'm Feeling Lucky . . . . .	272
PhoneBook . . . . .	273
Spell Checker . . . . .	279
Street Maps . . . . .	280
Special Tricks and Treats . . . . .	280
Double Words . . . . .	281
Right-Click . . . . .	281
Miscellaneous Google Information . . . . .	282
Web APIs . . . . .	282
Webmaster's Information . . . . .	283
Submit Your URL . . . . .	283
Advertising Programs . . . . .	284
Organizing Your Google Bookmarks . . . . .	284
Your Security . . . . .	292
The Search Challenge . . . . .	294
Summary . . . . .	295

## Contents

---

<b>Chapter 10 Searching Newsgroups . . . . .</b>	<b>297</b>
Google Groups. . . . .	297
Posting to Google Groups . . . . .	303
Removing Your Post. . . . .	310
Summary. . . . .	317
Afterword. . . . .	319
Index . . . . .	321

## Preface: Searching for and Finding the Golden Fleece

In the mythological account Jason and the Argonauts, Jason and his heroic pals must sail the seas searching for the Golden Fleece.

Jason must return to his homeland with the fleece in order to reclaim his rightful position as king of Iolcus. In his quest, he encounters challenges of every description that he must overcome by shrewd thinking and more than a little help from the gods.

Finally, he finds the Golden Fleece, guarded by a horrific dragon, in the kingdom of King Aeetes, king of Colchis. With the help of King Aeetes' daughter, Medea, he makes off with the Golden Fleece and triumphantly returns to Iolcus to claim his birthright.

Sometimes, when I have a particularly difficult time trying to find some nugget of information on the Internet, I recall the story of Jason and how he did not yield to seemingly insurmountable obstacles. Always at the right moment, Jason achieved the specific objective required to pass on to the next step or phase of his quest because he did not quit. So it seems, too, that I ultimately prevail in my quest for information by perseverance and determination. Of course, Jason had help from the gods on occasion in the form of divine intervention. It is not likely that we will find help in the form of a (handsome) Greek god or (beautiful) goddess (shucky darn!) in our search for information on the Internet. But, like Jason, if we don't quit and we have a little help from a good book with useful information, then we will find our golden fleece (of information) too.

This book addresses two perspectives involving a web search. The first perspective is that of the user. There are an estimated 300 million web searches performed each day. While each user may perform two or three searches per day, that is still a lot of users. Most users do not know how to use the full array of search tools available to them. Because the tools are not fully utilized, searching

becomes a chore often filled with a sense of frustration. When the proper search tools are used, searching becomes a pleasure.

The second perspective is that of the web page designer. Every web designer wants the ranking of the web page in the search results to be as close to the coveted number one position as possible. How do you design a web page so the page's position in a search is in the top 30 or 40 results returned by the web search engine? You must understand why and how people search and how search engines rank pages. Simply put, the two perspectives are:

- How do you make your “golden fleece” visible to the world?
- How do you find your “golden fleece”?

Just like the mythical tale of Jason and the Argonauts, searching for your golden fleece (or nugget, as I refer to a search result throughout the book) of information on the web can be a long, tedious, and sometimes hazardous journey, fraught with unknown pitfalls. This book provides a clearly defined path for your quest that will lead you to the “golden fleece.” Before we begin that quest, let’s consider where we will be questing and why go on the quest at all. Our quest will take us through the electronic wilderness known as the Internet.

In recent years, there have been numerous comparisons of the Internet and its associated World Wide Web with a library. It seemed to make sense, as both places are repositories of information. But two authors of a recent book about Google make the astounding claim that the Internet and its associated web are not a library! They advance the claim based upon their belief that the Internet lacks:

- A central source for resource information
- A paid staff dutifully indexing new material as it comes in
- A well-understood and rigorously adhered-to ontology

They go on to say, “Thinking of the Internet as a library can be misleading.” For sure, the Internet does not represent a brick and mortar building located in a downtown metropolitan area where no one in their right mind would be caught after dark. But the general

meaning of a library does not necessarily fit within the confines of someone's rigid definition. In a broad sense, a library is just a source of information. While we associated brick and mortar buildings with the term "library" in the twentieth century, we will come more and more to associate the term with the Internet (or at least certain portions of the Internet) in the twenty-first century. Why? Why not?

There really is no compelling reason to continue producing paper-based books other than it is in the vested interest of several industries to chop down trees, make pulp paper, and have printed matter. The publishing industry is beginning to recognize the power and impact of electronic publishing, and they will move, probably slowly, toward an all-electronic product. Eventually, paper-based books will become collector's items of a bygone era. But there is much more information in the world than what is offered by the (book) publishing industry.

There will probably be brick and mortar libraries around for some time to come, as the older members of the current generation are accustomed to visiting a library and checking out books. But, as the younger generation today moves toward retirement, fewer people will be comforted with a visit to a library. Instead, they will access all of their information needs via the computer and Internet. I am a research-intensive individual, yet I find little need to visit a library anymore. The Internet will become, if it is not already, the world's leading source of information, making it a twenty-first century library in my book (pun intended).

Information is power. Power is survival in a competitive environment that is increasingly hostile, as the world's burgeoning population struggles for fewer and fewer resources. The easy access to information that a computer and the Internet provide to a knowledgeable person in the privacy of his or her own home or office is awesome. But the access is "easy" only if you know how to find your golden fleece. This book is essentially a road map that leads you to your search objectives, making access to information easy for you in the truest sense of the word.

I know that everything you read or see on the Internet is not true. The same can be said for newspapers (recall the *New York Times* scandal of June 2003), books, magazines, television, radio, and, in general, any source of information. With such a vast source of information available at our fingertips, the challenge now for each of us is to wisely use that information to accomplish our objectives. The scope of this book is not how and what information you use but how to find any information you need. You must decide how valuable the information is to you.

## **Acknowledgments**

Thanks to Jim Hill of Wordware Publishing for the opportunity to write this book. I owe him an especially grateful acknowledgment for his patience and professionalism. Thanks to Judy West of CyberRecruiters.com for tech-checking the manuscript.

This page intentionally left blank.

# Chapter 1

## All about Search Engines

### Introduction

---

Many people today use the Internet to send and receive e-mail, shop, pay bills, search for lost relatives and ancestors, chat with kindred spirits in obscure chat rooms, play games, view movies, listen to music, find partners, search for jobs, search for employees, research medical issues, find inexpensive hotels and vacation packages, hunt for collectibles, and do countless other things. One estimate places the number of daily Internet users at 425 million people worldwide.

Adjunct to our use of the Internet is our need to find a particular nugget of information that we require to fulfill some need we have. Perhaps we want to find out when the War of 1812 was fought or the length of the Seven Year's War. Maybe we are looking for that special person. Or perhaps we search for a long-lost relative. How about an inexpensive vacation to Orlando? Whatever the reason, millions of people each day get on their computer and look for information.



According to SearchEngineWatch.com, the three most popular Internet activities are sending/receiving e-mail (81 percent of users), searching the web (57 percent of users), and looking for product info (46 percent of users). (*February 17, 2000, "How People Use the Internet," SRI — SearchEngineWatch.com*)

People are people, and searching for content on the web is a varied experience for us. The experience can be very gratifying ("Hey look, I found this thingy that I wanted for half price!"). Or, the experience can be very frustrating ("I have looked through 10,000 search results/sites and cannot find the thingy that I am looking for."). SearchEngineWatch.com reports that we are more frustrated than gratified. According to SearchEngineWatch.com, 77 percent of us find searching the Internet frustrating to some extent.

How long do people search for their thingy? Do they really spend hours or days looking through 10,000 search results to find their nugget of information? Again, people are people, and it doesn't seem likely that they would spend very long looking for their "acorn." According to the WebTop Search Rage Study, Americans typically experience "search rage" if they don't find what they are looking for within 12 minutes. (*WebTop Search Rage Study, August 2000 — SearchEngineWatch.com*)

But how frustrating does everyone else find web searching? Table 1.1 details the results of a SearchEngineWatch.com study.

**Table 1.1: Frustration on the web, from SearchEngineWatch.com**

5 — Very frustrated	29%
4	17%
3 — Frustrated	31%
2	11%
1 — Not at all frustrated	9%

An amazing 77 percent ( $31\% + 17\% + 29\% = 77\%$ ) from Table 1.1) of us find searching the web an unpleasant experience. Yet, it doesn't have to be that way. I never get frustrated searching the



web, and I always find the information I am looking for within moments of initiating a search. There is only one instance in ten years of web use that I could not find what I was looking for — the names of the three Rockwell International employees assassinated in Iran in 1976.

Here is a challenge to test your search skills, a kind of before-and-after test: Can you find the names of the Rockwell employees with just the scant information I have given? Try searching now and then search again after you have read this book. Compare your search results. The comparison will be a good benchmark that will illustrate how your search skills have improved.

Armed with a little patience and a few tools in your search kit, you can find anything that you seek within a minute or so and change your experience from frustrating to exciting and pleasurable. You must provide the patience, and I will provide the tools. Before we develop the tools that we need to reduce our search to a pleasant journey in cyberspace, it is useful to discuss the web and its associated web pages. Understanding what the web is, how web pages constitute the web, and what constitutes a web page will lay a foundation upon which we can erect our workshop before stocking it with the tools necessary to yield that promised pleasant experience.

## What Do We Search For?

Do you want to know what people are searching for? Check out <http://www.wordtracker.com/>. Click on the Keyword Report icon and then fill out the short form for a free weekly trial subscription, and you will receive the top 500 search words for each week.

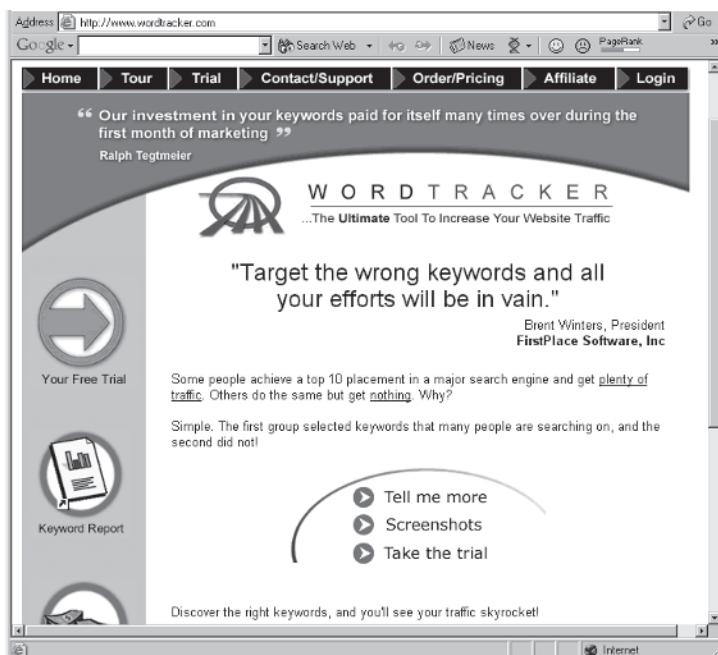


Figure 1.1: WordTracker home page

The top ten search topics for December 2002/January 2003 are:

**Table 1.2: Top ten search topics**

No.	Count (#of hits)	Keyword
1	180,863	travel
2	167,170	people search
3	159,908	autos
4	134,805	Google
5	123,209	Yahoo
6	118,467	jokes
7	118,129	eBay
8	113,537	hotels
9	109,661	health
10	87,108	Red Cardinal

**Note:**

Check out "Red Cardinal." No, it is not a bird or a priest! Red Cardinal is present in the top ten because of the season. In July 2003, it did not even make the top 500 search terms.

The list in Table 1.2 does not agree with the top ten items identified in a survey conducted by SearchEngineWatch.com, the results of which are shown in Table 1.3:

**Table 1.3**

<b>What kind of information do you look for on the Internet?</b>	
News	54%
Entertainment	53%
Health	50%
Business	49%
Academic	47%
Shopping	46%
Financial	44%
Career	37%
Sports	32%
Games	26%
Don't Know/None/No Answer	3%

News, listed as number one in the SearchEngineWatch.com survey, does not even figure in the top ten in the actual keyword searches, as reported by WordTracker. The difference in the two lists can be explained by the nature of people. Surveys tend to be static (people only remember what they last searched on) and misleading, whereas keyword search reports tend to be dynamic (recording all keyword searches) and accurate.



## How Much Do We Search?

---

“On average Americans spend 1.5 hours per week searching for information.” — *WebTop Search Rage Study, August 2000* ([SearchEngineWatch.com](http://SearchEngineWatch.com))

“Americans search the web practically every other day. Nearly 1/3 search once or more per day.” — *WebTop Search Rage Study, August 2000* ([SearchEngineWatch.com](http://SearchEngineWatch.com))

In my professions as author, engineer, and businessman, my colleagues and I spend much more time searching the web than the above WebTop Search study suggests. I spend at least one hour per day searching, and many of my colleagues spend even more time. On weekends, people spend enormous amounts of time searching the web, far exceeding any other form of recreation. Cruising the web has become the nation’s favorite pastime. My apologies to baseball and sex, but you were both superceded in this category around 1991 (baseball) and 1995 (sex).

## Why Search?

---

A revolution of global proportions and of great importance occurred in the past ten years. Perhaps few people noticed at first because the casualties were at a minimum and the body count was not projected every night on network TV. Only in the last couple of years did people start to take notice; by then it was too late. The world was wired and routed, and we, the consumers, were primed for the Internet. The consumer battle cry in this revolution was multimedia, all of those graphics-rich features available today at the click of a mouse button. But businesses and governments were also quick to jump on the Internet bandwagon. The Internet (or rather, access to the Internet) became in an amazingly short time the most sought-after commodity in the commercial marketplace.



Consumers wanted to overdose on all that graphics-rich multimedia, businesses wanted to capitalize on new markets and new revenue streams, and governments wanted to spy on their citizens. The varied interests of these three groups converged at the intersection of “global” and “unification.”

Global unification is the standard under which the current economic and, to some extent, political revolution is being fought. Prior to 2000, global unification was a practical impossibility. For global unification to be a realistic possibility, certain things are required: a global language, a global communications system, and a global library. English is the global language, modern telecommunications fulfills the need for a global communications system, and the Internet/World Wide Web is the global library. The most significant issue with the global library is that there are no web librarians to help us locate our nugget of information reposing among the billions of web pages. While you may walk into a public library and receive every consideration from the helpful staff, there are few, if any, helpful staff members in the web library.

Successful people know the usefulness of a public library and how important it is to be able to find information when they need it. Now the web has opened up a whole new library “frontier.” Knowing how to successfully manage (that is, find and retrieve) information in the web library is important for several reasons. It is important to the success of careers, it is important to the success of self-fulfillment, it is important to the success of economies (especially the global economy), and that makes it important to the survival of governments.

There is a metaphor that states “Even a blind squirrel finds a nut, occasionally.” The metaphor’s obvious meaning is that a squirrel without sight will stumble upon a nut if given sufficient time to search, assuming it does not give up searching. So it is with users who search the web. Given sufficient time and determination, we will eventually stumble upon our “nut” of information.

But few, if any, of us are determined enough to search through millions, or billions, of pages of information to find our “nut.” So, to



reduce the problem to a, more or less, manageable solution, web “search engines” were introduced a few years ago. The search engines use search terms called query words, or keywords, entered by the user to find all the existing web pages containing the search terms. Voilà — instead of looking through millions of pages for our nut, we only had to look through, say, 100,000 or so. But humans, being the intelligent creatures that we are, soon realized that if we did not find what we were looking for in the first 100 or so search results, then maybe we really didn’t need to know the answer. But, be honest. After looking through the first 100 search results, haven’t you ever wondered what that 100,000th web page really was and if it contained something pertinent to your search?

To understand search engines and web searching, perhaps we first need to explain the “web,” the “Internet,” and “web pages.” The *web* and the *Internet* are terms loosely used to describe the vast interconnecting “web” of computers sitting in numerous offices, classrooms, homes, closets, etc., in every corner of the globe. *Web pages* are the electronic manifestations of books, magazines, catalogs, newspapers, commercial enterprises, government entities, educational institutions, home pages, and any other topic, subject, or creation that human invention can devise to electronically represent the interests of some entity. Enough of the bull, already! So, what, specifically, is the “web”?

## What Is the Web?

A company that warehouses or hosts web pages, usually for profit, on one or more computers is called an *Internet service provider* (ISP). Companies and individuals pay service providers to “host” their web pages. Two or more Internet service providers connected electronically form a “web.” Today, we often speak of the World Wide Web (WWW, or www, or just web). Figure 1.2 illustrates a simple “web.” To visualize the real World Wide Web, just imagine thousands of cities and hundreds of millions of users connected via, primarily, the worldwide telephone system.

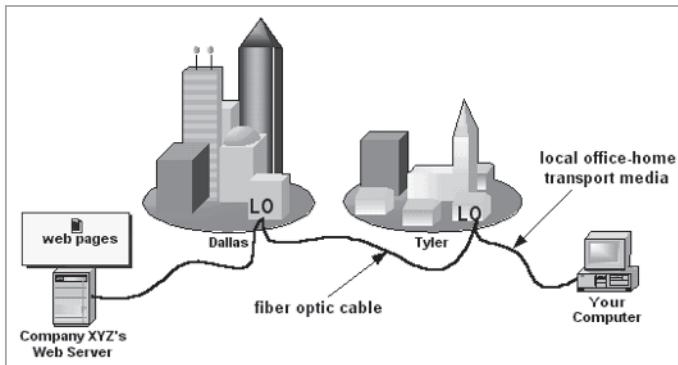


Figure 1.2: A World Wide Web

“Host” means the electronic representation of a particular web page residing upon the hard disk that constitutes the Internet service provider’s Internet/web computer assets. In the diagram, XYZ in Dallas is our ISP. Company XYZ owns and maintains one or more computers that contain the web pages of other companies and individuals who have paid XYZ to host the pages. The ISP can be a small, medium, or large company. One well-known ISP is America Online. Besides warehousing third-party web pages, ISPs may also host their own web site where they offer a variety of services.

A key element that is required for XYZ to be a successful business is XYZ’s ability to transfer huge amounts of digital data between its computer(s) and the people who want access to the web pages stored on XYZ’s computers. This high-speed data connection is shown as a dark line going from XYZ to the LO (local office) in Dallas. “Local office” is the term used by the telephone industry to describe the equipment sites scattered around communities where small bundles of telephone lines are grouped into larger bundles. A local office generally handles about 5,000 subscribers (telephone users). This connection between XYZ and the local office is typically one or more fiber optic cables. The number of fibers depends upon the amount of data, or traffic, passing between XYZ and the outside world. The telephone company connects XYZ’s computers to the outside world via the local office physically located near the service provider. The telephone company also connects your



computer to XYZ's computers through the local office nearest you in your community.

"But how does one distinguish between the millions/billions of web pages and where they are physically located?" an inquiring mind may ask. The <http://www.yourname.com/gov/org/etc.>, called a web address, or Uniform Resource Locator (URL), is the information the telephone company uses to identify and locate the particular computer hosting the web page that you are seeking. Fast electronic circuitry encased in boxes called routers convert the URL into addressing values that are used by electronic switches to connect the user (you) to the service provider (them).

The web pages hosted by the service provider are grouped by individual ownership into web sites. The topic or focus of a web site can be virtually anything imaginable. A few examples are Yahoo (<http://www.yahoo.com>), Google (<http://www.google.com>), the White House (<http://www.whitehouse.gov>), the FBI (<http://www.fbi.gov>), the Marine Corps (<http://www.usmc.mil>), Plano Independent School District (<http://www.pisd.edu>), Boy Scouts of America (<http://www.scouting.org>), and your home page (<http://www.yourhomepage.com>). Each of these web sites are composed of web pages describing the products, services, or interests of their owners. A web site may be composed of any number of individual "pages." The content of these individual pages may vary, such as an individual's "home" page, detailing the life and interests of the individual, or they may be logically related, such as a mail-order catalog's offerings. If a human can imagine it, it is in a web page, somewhere.

Today, the web spans the globe. In an effort to logically divide the web into manageable pieces, the early Internet pioneers assigned two-letter country codes to countries. Table 1.4 lists a few country codes. Note that the United States' country code is "us," but in this country you rarely see it appended to a web address. That is because in this country, it is the default country code; therefore, there is no need to append it to the address.

**Table 1.4: Top-level domains**

<b>Country Code (Top-level Domain)</b>	<b>Country</b>	<b>Example</b>
uk	United Kingdom	tvr.co.uk
us	United States	city.plano.tx.us
my	Malaysia	parlmin.gov.my
de	Germany	sgi.de
jp	Japan	hitachi.co.jp
to	Tonga	netsurf.co.to
tv	Tuvalu	internet.tv

You can view the complete list of top-level domain (TLD) country codes at <http://www.iana.org/cctld/cctld-whois.htm>. In addition to the TLD country codes, three-letter domain codes are assigned to various entities. See Table 1.5 for a sample.

**Table 1.5: Internet domain names**

<b>Top-level Domain</b>	<b>Meaning</b>	<b>Examples</b>
.com	Commercial, personal	eds.com, ibm.com
.net	Internet service provider	sbcglobal.net
.gov	U.S. government agency	whitehouse.gov
.edu	U.S. educational institution	stanford.edu
.org	Not-for-profit institution	redcross.org
.mil	U.S. military	usmc.mil
.int	International	itu.int

When you consider that the web is a “World Wide Web” (WWW), also called just the “web,” it is not difficult to imagine billions, if not trillions, of web pages with content as varied as the individuals inhabiting the planet, existing as magnetic media (little ones and zeros imprinted into the magnetic coating of hard drives) in the innards of thousands of service providers’ computers. That is what the web is — tens of thousands of physically interconnected (via the telephone company and its local offices (LO)) service providers



hosting web pages. And the Internet? Just add Usenet, Gopher-space, and private “webs,” also called networks, owned, operated, and maintained by businesses, institutions, and governments, and you have the much ballyhooed Internet.

You can think of the web as an “electronic curtain” that hides the stage before the actors are prepared to begin the show. You may be certain that the producers and directors are working feverishly behind the scenes to set the stage before the curtain opens. When you load the web page on your desktop, laptop, palmtop, thumbtop, etc., computer, you are opening the curtain upon the World Wide Web stage.

Generally, a web site will contain a *home page* that may have links to one or more subsidiary pages. The home page is identified as the [www.yoururl.com/index.html](http://www.yoururl.com/index.html) page. The “index.html,” sometimes just “index.htm,” in the URL is the default identifier for the initial, or doorway, page to a web site. When you download a web page from an ISP, you are “accessing” the page for viewing. Okay, we have managed to get the home page of some web site from XYZ’s host computers to your desktop. Now, what is a web “page”?

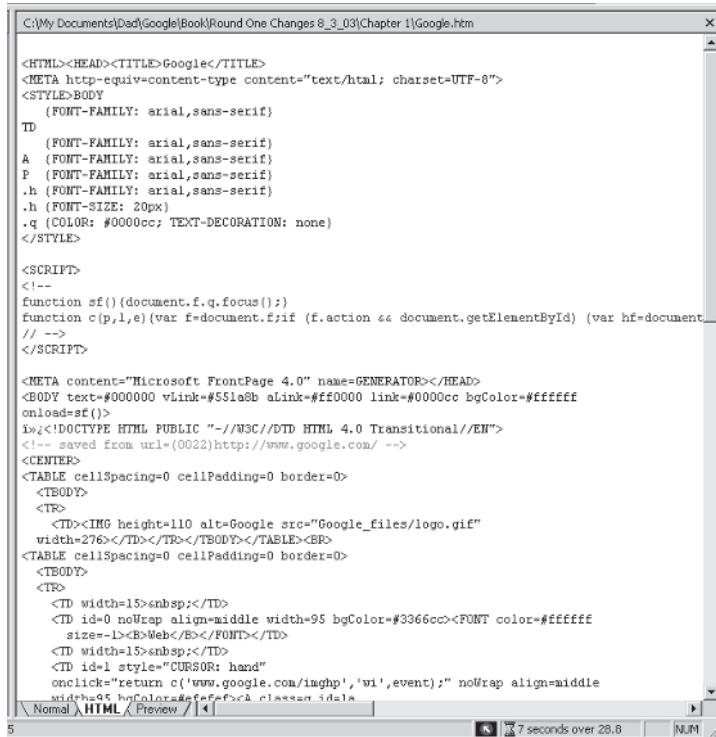
## What Is a Web Page?

A *web page* is a document in Hypertext Markup Language (HTML) format that provides some, hopefully useful, information to anyone accessing the page from his or her computer via a browser, such as Internet Explorer or Netscape. Whew! Glad we cleared that up.

HTML documents (called web pages from now on) are created using a special computer program, such as Microsoft’s FrontPage, and they are viewed on the computer using a web browser, such as Netscape, Internet Explorer, etc. A web page is much like any other type of “page,” such as the ubiquitous Microsoft Word document. The difference between a document created with a word processor and a web page is that the web page is organized in a manner that is designed to best present commodity data to the viewer, whereas



the word processor document is designed to best present textual information. Figure 1.3 illustrates the HTML text for Google's home page. Figure 1.4 shows what the HTML text looks like when a browser, such as Internet Explorer, opens the document.



The screenshot shows a Microsoft Word document window with the title bar "C:\My Documents\ Dad\Google\Book\Round One Changes 8\_3\_03\Chapter 1\Google.htm". The document contains the raw HTML code for Google's homepage. The code includes the DOCTYPE declaration, head and body sections with various tags like **HTML**, **HEAD**, **TITLE**, **META**, **STYLE**, **SCRIPT**, and **TABLE**. It also includes JavaScript code for handling form submissions and CSS styles for fonts and colors. The code is presented in a standard monospaced font used by most code editors.

```
<HTML><HEAD><TITLE>Google</TITLE>
<META http-equiv=content-type content="text/html; charset=UTF-8">
<STYLE>BODY
    {FONT-FAMILY: arial,sans-serif}
TD
    {FONT-FAMILY: arial,sans-serif}
A {FONT-FAMILY: arial,sans-serif}
P {FONT-FAMILY: arial,sans-serif}
.h {FONT-FAMILY: arial,sans-serif}
.h {FONT-SIZE: 20px}
.q {COLOR: #0000cc; TEXT-DECORATION: none}
</STYLE>

<SCRIPT>
<!--
function sf() {document.f.q.focus();}
function c(p,l,e){var f=document.f;if (f.action && document.getElementById) (var hf=document_
// -->
</SCRIPT>

<META content="Microsoft FrontPage 4.0" name=GENERATOR></HEAD>
<BODY text="#000000 vlink="#551a8balink="#ff0000 link="#0000cc bgColor="#ffffff
onload=sf()>
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<!-- saved from url=(0022)http://www.google.com/ -->
<CENTER>
<TABLE cellSpacing=0 cellPadding=0 border=0>
    <TBODY>
        <TR>
            <TD><IMG height=110 alt=Google src="Google_files/logo.gif"
width=276></TD></TR></TBODY></TABLE><BR>
<TABLE cellSpacing=0 cellPadding=0 border=0>
    <TBODY>
        <TR>
            <TD width=15>&ampnbsp</TD>
            <TD id=0 noWrap align=middle width=95 bgColor="#3366cc"><FONT color="#ffffff
size=-1><B>Web</B></FONT></TD>
            <TD width=15>&ampnbsp</TD>
            <TD id=1 style="CURSOR: hand"
onClick="return ( ('www.google.com/imghp','w1',event);" noWrap align=middle
width=95 bgColor="#ffffcc"><A href="http://www.google.com/imghp" id=1>
```

Figure 1.3: HTML document

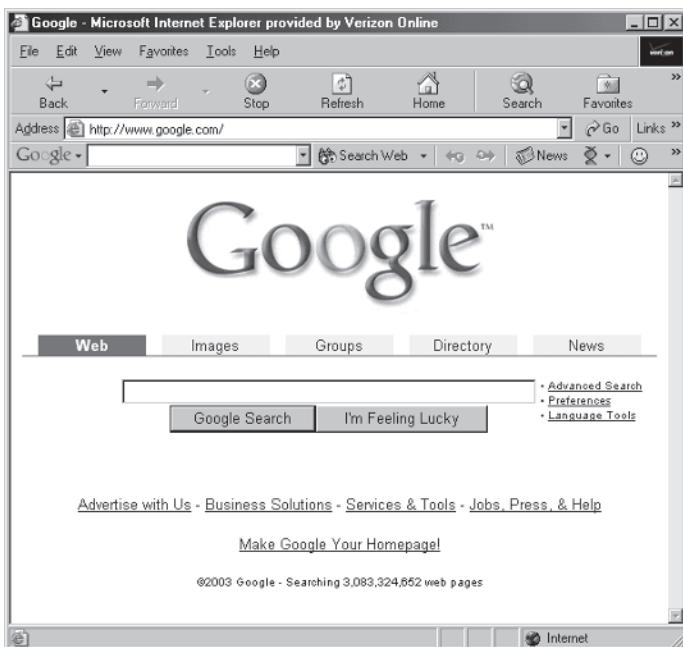


Figure 1.4: HTML document as seen in a browser

A well-designed web page entices you to purchase something, whereas the word processor document is designed to inform. Another difference between the two types of pages is the use of hyperlinks in web pages to move the viewer to another web page, to another location within the same page, or to another web site. But what is a hyperlink, you ask?

A *hyperlink* is a group of characters that include the URL of another web page or site. The purpose of a hyperlink is to move the user from the current web page to the web page designated by the hyperlink. The hyperlink (from now on called a link) is interpreted by browsers as a path to the specified URL. A left-click on a link whisks the user to the new page.

The web page has many hidden features that operate behind the scenes to achieve the objectives of the page designer. Meta tags, particularly meta description and meta keyword, are probably the most important behind-the-scenes web page features with respect



to finding a web page. Search engines typically use the contents of these two tags to identify the particular information contained within the web page. We examine meta tags in more detail in a later chapter.

Other document formats may be downloaded and viewed from a web page. One common format is Adobe's PDF document files. Other material of interest to us is accessed via web pages. Audio, video, and images are brought into our homes via the humble web page. It is the starting point from which our journey into cyber (or, more appropriately, with 77 percent of us frustrated when searching the web, hyper) space is launched.

The web contains millions of documents placed there by businesses, educational institutions, governments, and individuals. All of these documents comprise a "virtual library" of information accessible to everyone with a computer and an Internet connection.

But how do we find the particular piece of information that we are searching for in the virtual library? Let us assume for a moment that we live in a world without Internet search engines. In this world without search engines, we only have access to web pages through the Uniform Resource Locator (URL) address. A sample URL looks like this: [www.myhomepage.com](http://www.myhomepage.com). There are no printed "yellow pages" of URL addresses. If there were "yellow pages" of URLs, the URL list would be practically useless, unless we knew what service or commodity was provided by the URL owner.

Unless we have acquired a URL of interest through some means, such as an exchange between co-workers, we have no idea how to access pages of interest to us. Well, we could guess (like a blind squirrel?), but that approach is very time consuming and most likely fruitless. I have just described the Internet in its infancy, circa 1993.

The issue for you and me becomes how do we find, as expediently as possible, the particular piece of information, residing within the body of a particular web page, we seek in the vast repository of the World Wide Web? Just when the web was about to outgrow its



diaper and start its run to the millennium, search engines were created.

## What Is a Search Engine?

*Search engine* is a loose term for several types of software, typically owned and operated by companies other than ISPs, that “crawl” or “spider” the web space, identify the web pages’ URL and the web page content via the meta tag and actual page content, and store the information on the search engine owner’s computer(s). When a user enters the words of interest into the search engine’s search text box, the search engine searches its own database for web pages with content matching the search words and returns the URLs for those pages.

Today, some search engines also include “yellow page”-like directories that a user can browse to find web pages offering content of interest. Google is a search engine combined with a directory (Open Directory Project). Yahoo is a directory combined with a search engine (Google).

A search engine is sometimes called a portal. A *portal* is a web site that is considered an entry point to other web sites, each identified with a unique URL. Without portals, cruising the Internet quickly slows to a crawl. Portals maintain huge lists of URLs with instructions, embedded in the innards of the servers, routers, and gateways, on how to find those URLs.

## Search Engine History

The Internet is composed of the original FTP, Gopherspace, e-mail, Usenet, and Telenet services in addition to the World Wide Web. What distinguishes the WWW from, say, Gopherspace is primarily the format of the documentation. FTP sites include any digital data that can be stored on a hard drive. FTP sites store such things as computer programs, images, documents, audio, support files, utilities, FAQs, games, more technical stuff than any ten people can



read in a lifetime, plus other kinds of binary files. FTP sites were the original inhabitants of planet Internet. Then along came Gopherspace. Gopherspace is limited to text-based documentation, including technical and scientific treatises, publications, documents, and ancillary nerdy stuff. Most of the software is shareware/free-ware. If you have an adventurous spirit and a thirst for knowledge, check out some of these sites. Go to Google Search ([www.google.com](http://www.google.com)), enter “FTP,” and then click on Google Search. Enjoy the adventure of exploration.

The first search engine was Archie, released in 1990. There was no World Wide Web at the time. Data resided on defense contractor, university, and government computers, and techies were the only people accessing the data. The computers were interconnected by Telenet, a precursor to the web connection of today. File Transfer Protocol (FTP) was the methodology used for transferring files from computer to computer. There was no such thing as a browser. Files were transferred in their native format and viewed using the associated file type software. Archie searched FTP servers and indexed their files into a searchable directory. The Internet was small in comparison to today’s web, requiring a very unsophisticated Archie. Gopherspace came into existence with the advent of Gopher in 1991. Gopher cataloged FTP sites, and the resulting catalog became known as Gopherspace.

In March 1994, the burgeoning World Wide Web was beginning its geometrical growth toward today’s entity. Two watershed events occurred during that month. The amount of information passing among Internet computers via Telenet/FTP was exceeded by the information passed among the new web browsers accessing not FTP sites but WWW sites. Also in that month, WebCrawler, a new type of search engine that indexed the entire content of a web page, was introduced.

**Table 1.6: Search appliances**

<b>Search Appliance</b>	<b>Date Started</b>	<b>Designer</b>	<b>Profession</b>	<b>Accomplishments</b>
Archie	1990	Alan Emtage	McGill University student	Archived anonymous FTP sites/searchable directory
Gopher	1991	Mark McCahill	University of Minnesota student	Alternative to Archie
Veronica	1992		University of Nevada students	Archived Gopherspace
Jughead	1993			Added keyword and Boolean operator search features to Gopherspace search
JumpStation	1993			Spider
RBSE	1993			First keyword/spider
Architext (Excite)	February 1993		Stanford University students	First concept-based spider
WWW Wanderer	June 1993	Matthew Gray	MIT student	First fully automated robot
Aliweb	October 1993	Martin Koster		First web directory
WWW Worm	1994	Oliver McBryan		Spider
Galaxy (Tradewave Galaxy)	January 1994	MCC Research, University of Texas		First browsable web directory
WebCrawler	March 1994	Brian Pinkerton	University of Washington student	Indexed entire text of web pages in increasing relevancy
Yahoo!	April 1994	David Filo and Jerry Yang	Stanford University students	Directory
Lycos	August 1994	Michael Maldin	Carnegie Mellon student	Directory
Infoseek	February 1995			
MetaCrawler	June 1995	Eric Selburg and Oren Etzioni	University of Washington students	First metacrawler



Search Appliance	Date Started	Designer	Profession	Accomplishments
AltaVista	December 1995	Digital Equipment Corp.		First natural language and Boolean search techniques/first searchable, full-text database on the World Wide Web
Excite	December 1995	Mark Van Haren, Ryan McIntyre, Ben Lutch, Joe Kraus, Graham Spencer, and Martin Reinfried	Stanford University students	
HotBot (Inktomi)	May 1996	Paul Gauthier and Eric Brewer	UC Berkeley students	Powered by Inktomi search engine
LookSmart (Go2net.com)	October 1996			Directory
Argos	October 1996			First limited topic (ancient and medieval) search engine
Direct Hit	1998	Gary Culliss		Analyzes past Internet searches
MSN Search	September 1998	Microsoft		Search and directory
Google	September 1998	Larry Page and Sergey Brin	Stanford University students	Directory/link relevancy
FAST Search	1999	Norwegian owned		First to index 200 million web pages
Snap	November 1999	NBC		Search and directory

In 1994, webmasters and web site owners begin submitting sites for inclusion in the growing number of web directories. In 1995, automatic web directory search submission software was introduced. This type of software allowed webmasters and web site owners to automatically submit their web site to the major web directories with a click of a mouse. Also, meta tags in the web page were first utilized by some search engines to determine relevancy. In 1997, search engine rank-checking software was introduced.



This valuable software feature provided web site owners and web masters with an automated tool to determine their web site's position and ranking within the major search engines. In 1998, search engine algorithms begin incorporating esoterical information in their ranking algorithms. One such ranking approach was the inclusion of the number of links to a web site to determine its "link popularity." Another ranking approach was to determine the number of clicks (visitors) to a web site based upon keyword and phrase relevancy. By the year 2000, marketers determined that pay-per-click campaigns were an easy yet expensive approach to gaining top search rankings. Today, web site owners and webmasters understand that the best approach to elevate their sites in the search engine rankings is to build web sites that have useful and relevant content while optimizing their web pages for each specific search engine.

## **Robots, Spiders, and Metacrawlers**

Robots, spiders, and crawlers are the search engine machines that visit web pages and catalog the content of the pages. Collectively, they reside on their search engine owner's computer(s). Based upon a routine called Query, they search the Internet for servers and then collect the appropriate information when they find a server. Most search engines will use the links that they find on a web page to find other servers/sites. The information that the search engine collects is organized into a database. This database also contains those web sites submitted for inclusion by the web site owners/webmasters as part of the search engine's directory submittal process. The information viewed by the user as a result of a web search is actually data residing in that search engine's database. So, if the web site is not in the search engine's database, the user will never know about it (at least from the search results).

Now let's discuss the difference between robots, spiders, and crawlers. *Robots* only read a web site's URL and any embedded links in the web page. Since robots read the links, they (actually a separate software routine residing on the search engine's computer) can decipher every link to another site on the web page,



thereby forming “link trees.” A *spider* reads the URL/links, reading the title and the keyword section of a web page. Besides URLs, links, and title, a spider also reads the complete document, including meta tags. Therefore, spiders are indispensable for search engines incorporating “link popularity” and “relevance ranking” in their search rank algorithms. *Metacrawlers*, on the other hand, visit the various search engines’ servers and compile their search results from the sites cataloged on the engine’s server. They then list the results according to the collective relevancy of the individual search engine results. While this may seem “unfair,” it certainly speeds up the search process.

## Relevancy Ranking

Search engines are not all the same. Of course, each search engine uses some method to find its results. The search method that a search engine uses determines the quality and quantity of results that a user will get for any particular search. Generally, search engines can be categorized as author-controlled (keyword relevancy), editor-controlled (directories), user-controlled, or pecuniary-controlled. Each of these different methods that search engines use to find the user’s topic of interest skews the relevancy ranking differently.

Author-controlled search engines such as Google and AltaVista operate on keywords that the user supplies to the search engine interface. The search engine then searches its cache of web documents for matches to the keywords, returning those web page links in some predefined order as the search results.

*Keywords*, also called *query words* and *search words*, are those words describing the subject of interest. For example, a person searching for information on a Florida vacation might enter the words “Florida” and “hotel.” Of course, the search engine cannot distinguish between the hotels in Miami and Orlando, so the results will include all web pages containing those two words without regard to any specific Florida city. To narrow the search focus to the city of interest (say, Orlando), the searcher must include “Orlando” as another keyword in addition to “Florida” and “hotel.” So, the more



specific the keywords utilized, the fewer but more targeted search results the user will receive. Then he or she is on the path to search nirvana.

An important consideration of author-controlled search engines is the methods used to determine the relevancy of web pages to the search criteria or the keywords. Author-controlled searching relevancy ranking will depend upon the search algorithm that the search engine uses. Some may skew the results based upon size of company, link popularity, quality of web page, and/or whether the page is commercial (.com), government (.gov), or educational (.edu).

Editor-controlled search engines, such as Yahoo in its original form and LookSmart, place web page links into structured directories based upon subject matter. This type of search engine may be totally software driven or may involve some human intervention. The user may travel the top-down directory structure until arriving at the topic of interest, or he or she may search the directory for the topic of interest. The web pages are usually placed in alphabetical order within the directories.

User-controlled search engines depend upon not link popularity but rather visitor popularity. That is, the more visitors a web page receives, the higher it will rank in a search result. An example of a user-controlled search engine is Direct Hit. Direct Hit rankings are dependent upon the number of visitors to each site. The greater the number of visitors, the greater the relevancy ranking. Relevancy ranking in this case seems to be democratic in a way. However, behind every search engine is a group of individuals who can manipulate the results as they deem appropriate. Additionally, an enterprising web site operator can easily design a software program that can automatically visit the site repeatedly, garnering a top position in a very short period of time.

Pecuniary-controlled relevancy ranking is simply paid placement. It is also controversial, as it ensures a certain spot in the search results. Of course, the highest spot in the search results costs the most money. It's not very democratic, as one usually will not find a



small, enterprising company trying to establish a market on the Internet using this method of gaining web traffic to its site.

Relevancy ranking is an important concept for search engine technology. *Relevancy ranking* means how relevant a web site is to the search words entered by the user. When we search for web sites, we enter certain keywords in the search engine's search box. We expect the search engine to return the URLs of web sites that contain relevant information. If we search on "flowers," we do not care to discover that the search engine returns web sites whose content is "furniture." So, when the search engine's database is searched, an algorithm that ranks each site containing one or more of the keywords is employed. What are the mechanics of the ranking algorithms? Well, knowing that is a substantial part of the cat and mouse game that search engine owners and web site developers have played since about 1993. Suffice it to say that useful and relevant web site content is very important.

Relevancy ranking is a very important issue for both the user and the web page developer. For the user, relevancy ranking determines whether or not you find the particular piece of information that you are looking for in a reasonable amount of time. Studies indicate that most users do not scroll beyond the first 100 search results, while the top 30 results are considered the crème de la crème by the industry. So, relevancy ranking should place the most likely web sites of interest in the top 30, with the next 70 sites listed in order of relevancy to the topic. For web developers, their objective is to achieve a top 30 relevancy ranking for obvious reasons — more site traffic. So, the magical top 30 is everyone's goal. But should it be?

A statistic from one study says that 77 percent of Americans are frustrated with web searching. This statistic reveals that either relevancy ranking is not working or we do not know how to search the web. My own experience suggests that it is some combination of both. However, I believe the greater fault for our frustration lies with us, the users, due to our inexperience.



When a user only utilizes the basic services of a search engine, the search results identify every web site that contains the search keywords without regard to textual relationships and meanings displayed for the user's viewing in a rank ordering that realistically may not represent the user's interests. With three billion web pages categorized, it should be no surprise when a search based upon a couple of keywords returns 3,400,000 results. After reviewing the first 100 or 200 search results, we give up in frustration. Been there, done that! However, by utilizing a search engine's advanced search capabilities, a user can easily narrow the engine's focus to the most relevant sites. Then the power of relevancy ranking is projected from the screen. The next nine chapters investigate the basic and advanced search capabilities of the Google search engine.

## How Does a Search Engine Work?

---

First, let's make sure we know what a search engine is and then we can explain how it works. The use of the word "engine" to describe the Internet search function is usually misleading. We typically think of any type of engine as an internal combustion machine useful for providing the motive power of innumerable automobiles, airplanes, ships, etc. Internal combustion machines are certainly one form of "engine." In the modern, general sense, "engine" is "any machine that converts forms of energy into mechanical force and motion." This does not describe a "search engine"!

Another definition of engine exists that is archaic, but the meaning is closer to what a web "search engine" is. This definition describes an "engine" as "something used to effect a purpose." In our case, the "something" is software code running on a computer and the "purpose" is finding and indexing documents residing on other computers.

Search engines attempt to locate and index as many web sites (URLs) as possible. Search engine features vary greatly, as does the



actual scope, size, and accuracy of the databases from engine to engine. Okay. But what the heck is a search engine? A search engine is a loose term for several types of software, typically owned and operated by companies other than Internet service providers, that “crawl” or “spider” the web space, identify web page content via the meta tag and actual web page content, and store the information on the search engine owner’s computer(s). Then, when a user enters the words of interest into the search engine’s search text box, the search engine searches its own database for web pages with content matching the search words and returns the URLs for those pages.

When we search the web, we are looking for content embedded within HTML-formatted documents. HTML is the default web page format, and search engines look specifically for these types of documents residing on the service providers’ computers. Other data formats can only be found through a reference in an HTML document.

Let’s say that we want to find a certain MP3 file (a music file format) and download (move from the Internet service provider’s computer to ours) it so we can enjoy listening to our favorite music CD. We must first find a web page formatted as an HTML file that includes a reference/link to this music file. From this web page, we can download the music file. We cannot go directly to the music file and download it from the search engine results.

Search engines attempt to locate and index as many web pages as possible, given the search engine’s physical resource limitations. Search features vary greatly from engine to engine, as do the scope, size, and accuracy of their databases.

Search engines find the information that the user is looking for by comparing the keywords the user enters in the search text box to the search engine’s indices and returning the URL when a match is found. Search engines operate on:

- Unique keywords
- Combinations of unique keywords



- Special search characters
- Boolean operators
- Field searching and limiting parameters
- Advanced search commands (i.e., advanced modifiers)

While search engines may include millions, even billions, of pages in their databases, none of them index the entire web, much less the entire Internet. The following lists some of the content not found by search engines. Note that I said content. Search engines will find references in HTML documents to these items:

- Content of Adobe PDF and other non-HTML formatted files:
  - CGI (i.e., form data)
  - MP3 (music files)
  - MPG/MPEG/AVI (movie files)
  - JPG/JPEG (image files)
  - TXT (text files)
  - DOC (document files)
  - Any number of other less familiar file types
- Content protected from external access, usually requiring a login (intranets):
  - Corporate
  - Government
  - Educational
  - Institutional
  - Private (i.e., my home LAN)
- Web sites/pages not associated with an InterNIC domain name
- Commercial resources with InterNIC domain name limitations
- Web pages utilizing a robots.txt file (keeps files out of search engines' reach)



- Limited access sites (private domains)
- Other non-web resources

Google has the capability, which is unique among search engines, to search the content of certain file types listed above. We discuss the file types that Google can search in a later chapter.

## Issues Searching the Web

You probably have some experience searching the web for information. You enter your keywords into the search box and breathlessly hit Enter and, voilà, there are your 3,467,972 search results. Some search engines display search results in a sequential manner. That is, you must view these items 10, 20, 30, 50, or 100 at a time, starting from the first and going sequentially to the last. When you are at the last item, you must sequentially go back through the previous items if you decide the page of interest is somewhere in the previously displayed results. Such an approach is very time consuming.

Let's say I have the patience to view 250 results, 20 at a time:

1 - 20 of 3,400,000 | [Next 20](#)

21 - 40 of 3,400,000 [First Page](#) | [Previous 20](#) | [Next 20](#)

Now, I want to go back to the 60th result. I must return to the first page by clicking on the [First Page](#) choice (if available — not all search engines provide this option) and then select [Next 20](#) until I get to the 60th result. Or I must select [Previous 20](#) from the spot where I decided my best choice was the 60th page. Neither approach is very efficient for moving around search results. Ideally, the search engine interface would allow me to jump to any search result I wanted. When searching for a nugget of information, who really didn't wonder what was on the last page of those 1,237,401 search results?

In any case, why do Google and other search engines tell me there are 1,237,401 search results? Am I ever going to see more than



about the first 250 results? Not likely. So, if they do not provide a means for me to easily see those results, why don't they just limit the results to the first 1,000 results? Then our curiosity would not be aroused. Which search engine would you use — one that said it found 1,000 matches or one that said it found 1,000,000? Me too! But there is a ray of hope shining at the end of our search tunnel. Before we are finished with this book, we are going to demystify the net and find our nugget of information in just a few moments of searching without worrying about that last page.

## Search Integrity

---

Google claims its automated search method using PageRank does not easily allow for human tampering. Other search engines make similar claims. But who knows for sure?

Most search engine user interfaces do include paid ads in their search results. However, only the pecuniary-controlled search engines accept money (above the table) for paid placement. Who knows what goes on under the table? After all, humans will be humans, most of the time, and every search engine has at least one human operating behind the scenes.

## The Race to Be Number One

---

The race to be number one in the search engine industry has been a closely watched race for a number of years. Marketing dollars are precious, and companies want to invest them where they think they will get the greatest return on their investment. The search engine race has been quite an interesting affair to watch. Longtime front-runners can be overtaken by upstart new companies in a matter of months. The search engine industry is still a very dynamic



environment, with new and untried approaches to web searching being introduced quite frequently.

The current search engine leader is Google. Its rise to the premier industry position has been amazing. In five years, despite spending little or nothing on marketing itself, Google's home page has become the 15th-most-visited web site in the U.S., according to Jupiter Media Metrix, Inc. With the company handling 114 million web searches every day (Search Engine Watch, April 2003), Google traffic leaped over 75 percent last year, making it the top search site on the web. Google itself claims to receive 150 million hits per day and index over three billion web pages. These pages must be visited periodically by Google's search application in order to keep its database up to date.

Last summer, a survey by market researcher NPD Group ranked Google the most effective search engine. Of the users surveyed, 97 percent reported locating what they were looking for "every time" or "most of the time." That number is amazing considering the number of searches performed in the vast ocean of web content. The following table shows Google leads its nearest contender by a nearly 2:1 margin.

**Table 1.7: The most effective search engines**

Google	Approximately 150 million searches per day (per Google's web site)
Inktomi	Approximately 80 million searches per day (per SearchEngineWatch.com)
AltaVista	Approximately 50 million searches per day (per SearchEngineWatch.com)
Direct Hit	Approximately 20 million searches per day (per SearchEngineWatch.com)
FAST	Approximately 12 million searches per day (per SearchEngineWatch.com)
Galaxy	Approximately 100,000 visitors per day (per Galaxy's web site)



Google claims to run its search application on a unique combination of advanced hardware and software. The speed that a user experiences is attributed, by Google, partly to the efficiency of its search algorithm and partly to the thousands of low-cost PCs it has networked together to create a super-fast search engine.

## What Is Google?

---

Google is primarily a web search engine combined with a commercial catalog. Many Internet pundits claim that Google dominates the Internet search engines. At 114 million hits per day (Search Engine Watch, April 2003), it exceeds Yahoo's 42 million daily hits (Search Engine Watch, April 2003) easily. Web sites are seeing a tremendous growth in the use of Google. One web site claims that it has seen a 185 percent growth in Google search referrals in the past 12 months. By another pundit's estimate, Google now accounts for 66 percent of all search referrals on the Internet. According to Google, its web search catalogs over three billion web pages.

Google is a feature-rich search engine that incorporates numerous tools and options, giving web searchers many different approaches to solving their search issues. Google is not a solution to searching the web; it is an instrument, a method, or a tool itself, and it is a very good instrument. You are the solution to your search issues. It is up to you to learn Google and apply what you learn, using Google as a craftsman would use an expensive and well-built device, to solve your search issues.



## How Google Ranks Its Search Results

Google believes the web is a democracy. That is, Google believes every URL link embedded in a web page is a “vote” for another web page. Operating on this belief, the founders developed a search algorithm called PageRank. PageRank monitors how many web pages/sites link to a page. Factored into the algorithm is site importance. A higher search result occurs for sites that are linked by other important sites than for sites that are linked by “unimportant” sites. That is, New York University’s Department of Computer Science’s “Apple Pie Parser” web site will be weighted heavier (in importance) than Mom’s “Apple Pie Recipe Home Page,” regardless of the search criteria. The more important the sites linked to the page, the higher the ranking. It is rumored that Google gives a slight advantage to .gov and .edu sites.

Google claims it does not allow anyone to purchase a higher ranking in its search results. Google does, however, include what it labels “Sponsored Links” on some results pages that may be confused with actual search results. For example, if you search on the keyword “hotel,” you will see PageRank’s search results on the left side of the results page and sponsored links relating to hotels at the top of and to the right side of the results page. The items labeled “Sponsored Links” are actually ads.

By analyzing links using “sophisticated text-matching techniques” and showing favoritism to non-commercial sites, Google assigns a PageRank to sites that it deems “important.” According to Google, this “democratic” way of reaching your search goal will assure you of “high-quality” results. What you really see in the top rankings is what Google deems important. Not really very democratic, is it? Probably the greatest fallacy to this approach of ranking search results is the obvious — anyone can pay other sites to link to their site.



Why is this a problem? Well, early in its development, the web was seen as the Great Leveler. The web was going to level the playing field for all who were interested in competing in the commercial markets of the world. The lone guy hawking pencils, the solitary woman selling bagels, and the big corporations had an equal opportunity to ply their products on the Internet. You did not need a big marketing budget to sell your products, and your market was essentially worldwide. In short order, Google managed to reduce this small merchant advantage by raising the bar for everyone but Fortune 500-type corporations.

## How Is Google Getting Better?

---

Google plans to capture more of the web's ever-expanding sea of content. The company's software currently searches more than 3.308 billion web pages. According to Google, that is an increase of 60 percent from more than a year ago. There are an estimated five billion searchable web pages, including foreign language pages, according to Internet search company Cyveillance. At present, Google searches and indexes about three-fifths of the available web content.

To extend its ability to search and index those additional pages, Google has increased the number of foreign languages that it will index. Currently, Google indexes web pages in more than 60 languages. For a comparison of Google's reach, consider rival AltaVista Co. AltaVista performs searches in only 25 languages. Also, Google is developing technology to search a wider variety of documents, including images.

The Google search technology is now powering the search areas of Yahoo, AOL, CompuServe, AOL.com, and Netscape. Google also provides WAP (Wireless Application Protocol) to AT&T Digital PocketNet Service Premium Plan customers. Google's WAP application automatically converts HTML documents on demand to the WAP format.



## The Next Generation

Search engines have become the gateway to the World Wide Web. Virtually every user enters the web via a search engine. A search engine's profitability is directly related to the number of people who access the web via its portal. Search engines, therefore, have a great interest in keeping their following satisfied. We can expect the competition among the search engines to produce better search results than currently produced. Also, we can expect niche search engines to flourish as people discover the higher quality results they provide. Some of these will be fee-based, such as Northern Light. Others will be free. We can expect to see improved searching techniques and a much better user interface. Improved relevancy should be a natural consequence of improved spiders and larger databases. But until the Internet bosses decide to include more domain types, splitting the domain space into finer groupings such as separate divisions between commercial and personal web pages currently grouped together in the .com domain, advances in search engine technology will be incremental.

## Summary

God created the heavens with its bountiful stars and earth in six days. The United States Department of Defense, in association with universities and commercial companies, created the Internet and its associated bountiful web in 20 years. One seems to me about as complicated as the other. Just as the firmament with its innumerable galaxies and stars was unmapped upon its creation, so too was the Internet. Only by more or less stumbling around blindly could a user find what he or she was looking for on the web. Then, creative university graduate students gave us search engines to help us find our nuggets of information. Now we can navigate the Internet firmament with greater ease than ever before.



Clearly, Google provides the clearest and most preferential map into the Internet firmament. The remainder of this book examines the inner workings of Google to provide insight for users of the application searching for their nugget and for web developers who desire a higher relevancy ranking from the Google search engine.

# Chapter 2

## Google Services

### Introduction

---

Most people, when confronted with a need to search on the Internet, jump right into a search engine, enter their keywords in the search (or query) box, hit the Enter key, and then wonder why they are presented with millions of search results that may or may not have anything to do with their particular search need. Each of the millions of search results represents a choice that must be made. Is my nugget of information in this web page or in another? Do I click on this link or another? Do I have time to review 200 web pages to find the particular piece of information I need?

Time is all we are born with that is a negotiable commodity. In simple terms, time is money and money does make a difference in our lives. It makes no difference if the search need is one at work or one at home. The choices we make determine the physical and emotional environment that we live in and our destiny. If we are presented with a baffling array of choices, we will flounder around while experiencing an uncomfortable sense of bewilderment, not-so-strangely wondering why we must utilize such incomprehensible products and services as computers and web searching in our daily lives. *That is why 77 percent of us experience search frustration.*

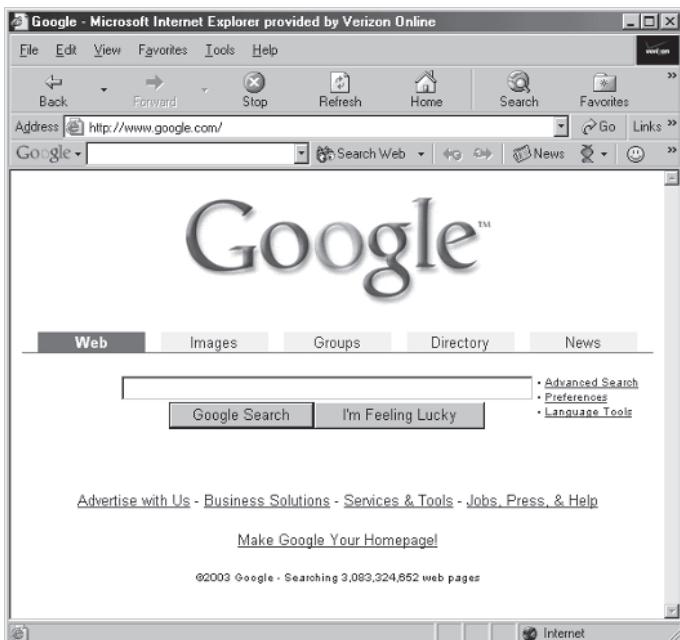


Figure 2.1: Google's home page

Thirty years ago computers were touted as the machines that would relieve us of the daily grinding, boring tasks that ground us into senseless robots after 25 years of working. Yet, it seems like the use of computers in the workplace has just added to that burden, not relieved it. There are many reasons for this, and you have no control over most of them. But one thing you do have control over is how you manage searching the Internet, and by taking control of that portion of your daily work routine, you empower yourself. Empowerment equals satisfaction, and satisfaction equals happiness. Voilà! No more search rage.

Google, by using PageRank technology to rank search results, attempts to minimize your need to sift through hundreds or thousands of search results to find your information nugget. However, PageRank does not understand your search need in the context of your life. Google displays search results according to the PageRank assigned to them. Keep in mind that clever web developers are constantly seeking to spoof PageRank into placing their web



page(s) into the magical top 30 search results. I have played around with web pages and moved pages from the 400th spot to the top 100. I know PageRank results can be manipulated, in a fashion. The point is, you may or may not find what you seek in the top 30 search results. If the top 30 is all you view, then you may be missing a real nugget in result 250.

But we do not have time to review 250 web pages to find that nugget. So we are right back where we started from, right? Well, we will be if we do not study this chapter (and the remainder of the book). In this chapter, we discover that Google offers more than just a basic search of the Internet. Google offers many services and tools that allow us to frame our search in the context particular to our search needs. That is, we can use Google services and tools to eliminate choices that do not add value to our search process. That makes us better managers of our search time, and empowers us.

Before we examine Google's search tools and techniques, we need to study the search services offered by Google. By reviewing the search services that Google offers, we will be able to make wiser choices about our approach to searching. Do we search Google's directory? Do we search the web? Or do we search a specific domain to find our nugget of information? After studying this chapter, we should understand that the service or method we choose to search the Internet is as important to our search results as the keywords we select to search on.



### Note:

We discuss keyword selection in Chapter 4.

The Google search application (from this point forward referred to as Google) is service and feature rich, including a wide array of tools to help you find exactly what you're looking for. Some of these services and features are mature and offer excellent results, while others are still being developed. We examine each of the services and features in detail in the remaining sections of this chapter. Google Services, shown in Figure 2.2, is found at <http://www.google.com/options/index.html>. (Bookmark it!) The web link



given is titled Google Services & Tools. Google Tools are discussed in Chapter 8.

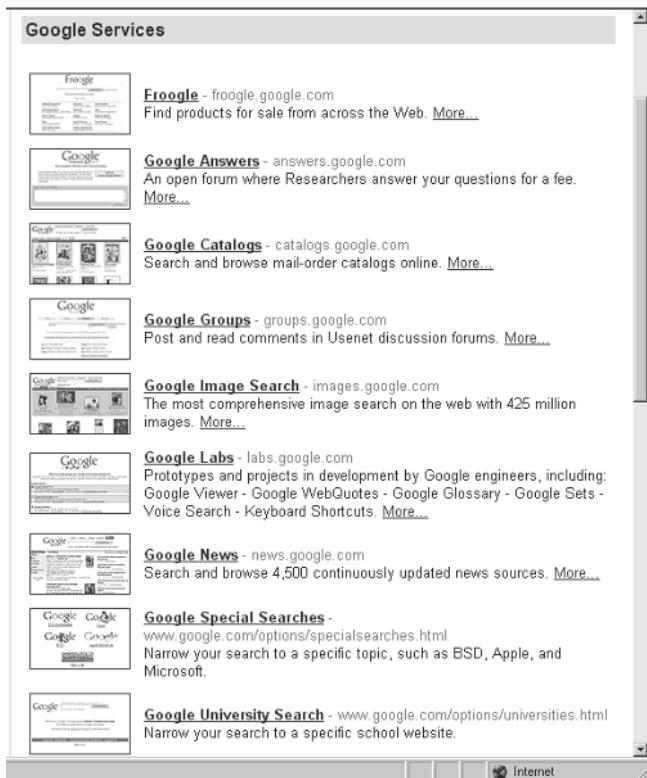


Figure 2.2: Google Services page

Google offers these specialized services:

- Froogle
- Google Answers
- Google Catalogs
- Google Groups
- Google Image Search
- Google Labs
- Google News



- Google Special Searches
- Google University Search
- Google Web Directory
- Google Web Search
- Google Wireless

Froogle is a service that lists commercial and retail web sites. Google Answers is a fee-based service offering answers to your questions. It is similar to Ask Jeeves, but real people answer your questions. Google Catalogs is a free directory of mail-order catalogs. Google Groups is the Usenet discussion forums. Google Image Search is an image search engine. Google Labs is a mixed bag of experimental features. Google News is a browsable and searchable database of over 4,000 news sources. Google Special Searches allows the user to narrow the search field and focus on a specific topic. Google University Search focuses the search topic on specific school web sites. Google Web Directory is Google's indexed web sites organized by categories. Google Web Search is Google's search engine. Google Wireless is Google's WAP portal for wireless and handheld devices.

We examine each of these services in greater detail in the remainder of this chapter. Before we do, let's look briefly at the features that we find in Google Labs, shown in Figure 2.3. Google Labs is a mixed bag of experimental features that includes:

- Google News Alerts
- Google Compute
- Google Viewer
- Google WebQuotes
- Google Glossary
- Google Sets
- Voice Search
- Keyboard Shortcuts

A screenshot of a web browser displaying the Google Labs page at <http://labs.google.com/>. The page lists several experimental services: Google News Alerts (launched 8/4/03), Google Compute (launched 3/26/03), Google Viewer (launched 12/10/02), Google Glossary (launched 5/20/02), Google Voice Search, Google Webquotes, Google Sets, and Keyboard Shortcuts. Each service entry includes a brief description, a 'Give us feedback' link, and a 'Discuss with others' link.

Figure 2.3: Google Labs page

Google News Alerts are e-mail messages Google sends you to inform you of breaking news stories. You select the criteria for news topics of interest, and when a story appears with your criteria, Google sends you an e-mail informing you of the story.

Google Compute enables your computer to participate in worthwhile collaborative scientific computing projects. This is the easiest way for you to be a hero or heroine. Simply by loaning your computer for academic use when you are not using it, you have contributed to an important scientific project benefiting all of humankind.

Google Viewer allows the user to continuously scroll through search results, as opposed to viewing 20 search results per page. Want to know what other sites are saying about your site? Google WebQuotes is a unique service that displays other web site comments about the web sites returned as search results. Google Glossary is a dictionary of words, phrases, and acronyms. Google Sets is an interesting approach to searching the web that returns results based on sets of related keywords. Voice Search enables web searching using a telephone connection and the user's voice.



Keyboard Shortcuts provides a means of using keyboard keys to navigate through search results.

## Google Services

Now we examine Google Services in detail. To reiterate, Google Services are Froogle, Google Answers, Google Catalogs, Google Groups, Google Image Search, Google Labs, Google News, Google Special Searches, Google University Search, Google Web Directory, Google Web Search, and Google Wireless.

Each of these services offers a unique way of searching and discovering information on the web. You will find them entertaining and useful to use. We begin our examination of Google Services with a look at Froogle.

### Froogle

Froogle itself is not a store. There is no shopping cart, and you cannot purchase anything in Froogle. Froogle is merely a “yellow page”-style site that links the consumer to a retail store’s web page where the item of interest may then be purchased. In Froogle’s list of stores, shops, and catalogs, only retail establishments are listed. You will not find cousin Lizzie’s home page describing her summer vacation with her parents in Froogle. Froogle comes from a play on the words “frugal” and “Google.”

Froogle home page:

<http://froogle.google.com/> (Bookmark it.)

Froogle FAQ:

<http://froogle.google.com/froogle/about.html>

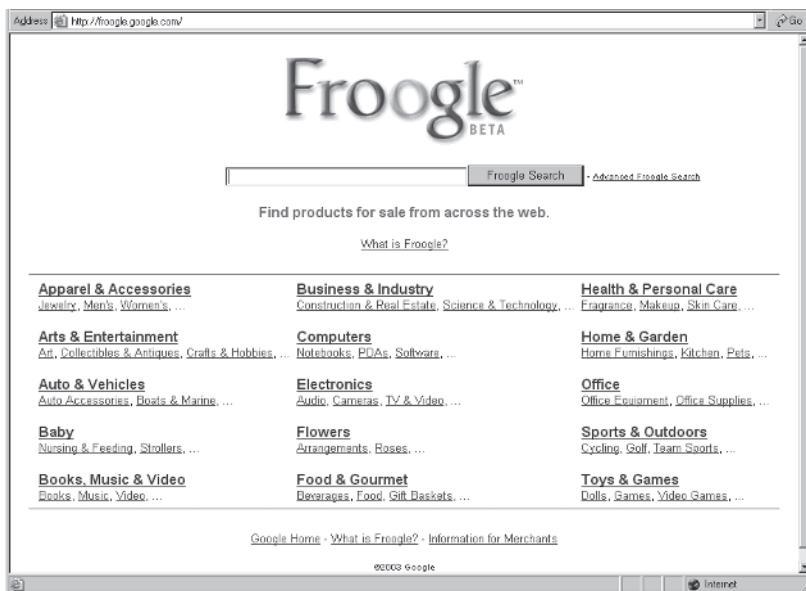


Figure 2.4: Froogle home page

Froogle is a new Google Service, so it may be excused for not working quite as planned on occasion. In the fall of 2002, a search for “computers” yielded results other than actual computers. The top 100 results were almost exclusively computer accessories, such as innumerable automobile cigarette lighter power cables for laptops. I did not have the patience to view more than the first 100 results. I searched on other subjects, and again the search results were equally disappointing. Finally, I decided to search on something simple, so I chose “socks.” I was amazed at the number of valid returns the application returned, and I was equally impressed with the (countless) number of variations on what I thought was a simple clothing item. I will not tell you how many I looked at, but I never found an invalid search result in the approximately 1,000 web pages Google displayed of the 1,000,000+ links found. So for retail shoppers, Froogle will ultimately be a shopping nirvana, but it appears that Google still has a few issues to address regarding search result relevancy.



I repeated this experiment in the spring of 2003. The search topic (“computers” this time) yielded results with actual computers in the first 14 search results, but the next five were:

- (#15) A service plan that provides printer installation, training, and driver installation for a maximum of two Phasor computers
- (#16) nFORCE-2 Special System Builder (a real computer result)
- (#17) A laptop light
- (#18) A computer switchbox kit
- (#19) A web page describing how to troubleshoot wired Ethernet problems

From this point forward, the quality of the remaining search results dropped below the threshold of acceptability. Almost all of the next 200 results were computer accessories and books.

Google states that its “spidering software” crawls the Internet identifying web pages that offer merchandise for sale. When a web page is so identified, the spidering software associates images found on the page with the search keywords. In this manner, Froogle builds its database seemingly apart and in a differing manner than the Google search engine that uses PageRank. Since the approach is totally software driven, mistakes, such as associating the wrong image with a particular keyword, may result. I can only imagine the possible combinations and resultant humorous associations. Besides utilizing Google’s spidering software, Froogle builds its database using product information submitted electronically by the seller.

Google does not specify how Froogle determines keyword relevancy. It makes sense if Google uses the pages’ meta tags to identify keywords, as this is the purpose that the meta tags were originally intended to serve. However, Google’s spidering software does not use the meta tags but instead operates upon the words found in the body of the web page. The results of my experiments indicate the approach used to match keywords to products is not



very accurate in terms of matching the principal products to the keywords.

Froogle is a great concept, but Google's primary implementation seems to be a software-driven engine that cannot discern search term meanings and product nuances. That is, the engine cannot distinguish between a noun and an adjective, ergo the difference in search results quality between "computers" and "socks." Socks do not seem to have many accessorized items that Froogle can mix with the desired search results, whereas computers may have thousands of accessorized items intermingling with actual computers. As an example of the disconnect, I think every Amazon.com computer and software book was included in my "computer" search results.

Searching for product-specific retail establishments seems at the outset to be a simple and straightforward endeavor. But the way to search results nirvana is always mined with the unexpected and unforgiving. Google wants to project the power of its search technology into the retail marketplace. But its link popularity algorithm ignores meta tags. Interestingly enough, the original purpose of meta tags was to allow web page developers to specify keywords that pinpointed the exact purpose of the web page's content.

Froogle will not be a useful feature of Google's stable of services until its software includes reading and understanding meta tags and the difference between a noun and an adjective. Conversely, retail establishments with a web presence must include valid and useful meta tags in their web pages to help identify the specific products that they are selling without resorting to games to spoof search engines into placing the page in places that it does not belong.

Froogle is easy to use. Just go to <http://froogle.google.com/>. Key in the keywords of interest, and the Froogle search will return nothing but businesses selling the item(s) and all of its accessories that you are seeking. Of course, you may still get more search results than any one person can view in a lifetime. Online shopping really does enable you to "shop until you drop." By focusing on product search, Froogle does weed out the inappropriate sites (at least that is the



ultimate goal). Until Google fine-tunes the word association portion of its ranking software, we will continue to get inappropriate search results. But, hey, it can be fun. A search on “pipe,” while I expected to receive thousands of bong results, yielded Native American peace pipes, books, golf visors (What does a cap have to do with a pipe? Do people really smoke them?), and two bona fide tobacco pipes in the first ten results. The next 100 results yielded a great variety of products that had at least a remote association with the word “pipe.” Alternatively, you may browse Froogle’s directory of merchandise categories.

Froogle’s search results are not the result of paid placement nor does the seller pay to be included in the Froogle directory. Froogle search results are automatically generated by Google’s “ranking software,” according to Google. Google does not state that Froogle uses PageRank — only its “ranking software.” It does not seem plausible for Froogle to use PageRank, as link popularity is not a useful concept in a directory model, and my limited experiments do not reveal that link popularity is the approach used. Google claims, “As with all other Google search results, Froogle ranks store sites based only on their relevance to the search terms you’ve entered.” That is a general statement that can mean just about anything.

It is not clear to me how Froogle associates a .jpg file with the keyword(s). Google says that it scans each web page and associates the .jpg files found embedded in the web page with the page’s keywords. Web page developers for retail product web sites would be well advised to determine through experimentation how Froogle associates .jpg files with page keywords. In the limited experimentation I have done, I have seen a few blunders in Froogle’s results.

Froogle does include paid “sponsored” advertising with Froogle’s search results. You will find the sponsored links on the right side in vertically stacked rectangular boxes. Placement of these paid ads is based upon the amount of money that the advertiser is willing to pay per click (cost per click, or CPC) and the click-through rate (CTR). The higher the CPC and CTR, the higher the placement. There is an inverse mathematical relationship between CPC and



CTR. As the CTR increases, the CPC may decrease and the advertisement will maintain its position. In other words, as the volume to a web site increases, Google will charge less to maintain the advertisement's current position in the sponsored links.

What is especially appealing about shopping with Froogle is the inclusion in the search results of a product image in about 90 percent of the search results. Clicking on the image generally takes you to the seller's web site, where a larger image of the product is displayed. Item purchase and checkout is generally just a click away. Shopping in this manner sure beats getting jostled and bumped in those awful Christmas crowds, and you cannot beat the parking space — your favorite easy chair.

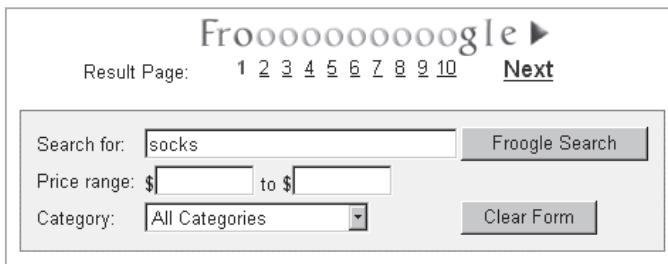


Figure 2.5: Froogle search box

Google has recently added a great feature to Froogle's search box that is displayed on the search results page. Now you can refine your search by placing the minimum and maximum price that you are interested in paying into the search box. See Figure 2.5. Note that the Price range text boxes are not present on the Froogle home page. You must search on some item before you can refine your search with a price range.

Google could greatly enhance Froogle by displaying some type of seal identifying those sellers with a record of honesty and integrity. Too many unscrupulous people use the Internet to bilk unsuspecting shoppers out of their money. Google could render invaluable assistance to countless consumers by establishing a seller ranking system based on consumer feedback. eBay has such a seller/buyer



ranking system that appears to offer some protection to both buyer and seller.

## Google Answers

Google Answers is an open forum where your questions are answered for a fee.

Google Answers URL:

<http://answers.google.com/answers/main> (Bookmark it? Don't bother.)

Google Answers FAQ:

<http://answers.google.com/answers/faq.html> (Bookmark it? No.)

The screenshot shows the Google Answers homepage. At the top, there's a navigation bar with an address bar containing 'http://answers.google.com/answers/main'. Below the address bar is the large 'Google Answers' logo with a trademark symbol. A sub-headline reads 'Ask a question. Set your price. Get your answer.' To the right of the logo are 'Log in or Create a Google Account' buttons. The main content area has a large input field for asking a question, with a placeholder 'Step 1 - Enter your Question. Tips for great results.' Below this is a search bar with fields for 'Search Google Answers for' and 'Google Search'. On the left, there's a sidebar titled 'Browse previously asked questions' with categories like Arts and Entertainment, Business and Money, Computers, Family and Home, and Health. On the right, there's a sidebar titled 'Recently answered questions' with links to 'Life span of a potted ivy plant', 'Dear Journalist, In response to my ...', 'Song Identification', and 'Pink -- why feminine?'. At the bottom, there are links for 'Google Home', 'Answers Help & Tips', 'Answers FAQ', 'Terms of Service', and copyright information '©2003 Google'.

Figure 2.6: Google Answers



Google Answers provides the user an alternative to online searching. Suppose you have searched repeatedly without success for the answer to a particular question. At some point, you may just give up. But for a question of some importance, many people will continue the search, perhaps asking colleagues at work, family, or friends or visiting the nearest public library. At some point, if they still do not find the answer they seek, they may decide that the amount of time spent searching is not worth the effort. Google Answers provides an alternative to just giving up. Remember my challenge in Chapter 1? What were the names of the three Rockwell employees assassinated in Iran in 1976? This question fits the criteria for giving up.

Google Answers is fee based. When you ask the question, you are given the opportunity to specify how much money you are willing to spend for the answer. Seventy-five percent of the fee is paid to the researcher, while the remaining twenty-five percent goes to Google.

Your question can be edited until a researcher “locks” the question. A locked question indicates a Google researcher has undertaken to research and answer the question, thereby preventing you from changing the question and preventing other Google researchers from attempting to answer it.

Google researchers scan questions and select those that pay the most for the least amount of effort expended. There do not seem to be any controls in place that guarantee you an answer. If you have not indicated that you will pay enough to warrant the amount of time the researcher must expend to get you the answer, then no one will bother to help you out. Your question will “expire” without resolution.

The time limit for Google “researchers” to answer your question is one month. One month seems unreasonably long, as most people’s questions have some sense of immediacy. Google researchers do not necessarily have any expertise in the subject matter, and Google does not provide resumes of the individuals that it contracts with to answer your questions. When a researcher decides to



undertake your question, he or she will look for the answer and then post it to you. Only when the researcher posts the answer are you charged for the service. Google provides a mechanism for feedback if you are dissatisfied with your answer. The researcher gets the chance to clarify the answer, and if you are still dissatisfied, you can then request a refund up to 30 days after receiving the original answer. All communications between you and the researcher/Google are posted on the Google Answers web site. Finally, you are given the opportunity to rate the researcher.

To ask a question, you must first establish an account. You establish an account by entering your e-mail address into a form. Then Google sends a confirmation e-mail to your e-mail address. You click on a confirmation link that verifies your e-mail address, and the sign-up continues with you entering a username. Then you are taken to a page where you can enter your question. After entering your question and the amount of money that you are willing to spend, you are taken to a form to complete your financial information. Google charges a nonrefundable administration fee of \$0.50 for each question.

Google posts every question and answer on its web site. The purpose is to allow other registered users, not researchers, to add their perspective/opinion and “share the research.” The individual asking the question is identified by a nickname that he selects when he registers to use the service. Folks who comment on these posts do not get paid (sorry). If a registered user posts a comment that “answers” your question prior to your question being locked or answered by a Google researcher, you may cancel your question and only pay the nonrefundable \$0.50 fee.

Google discourages questions that:

- Seek private information about individuals
- Seek assistance in conducting illegal activities
- Are designed to sell or advertise products
- Include material related to adult content
- Involve homework or exam questions (darn!)



### ■ Seek specific information about Google or Google Answers

Basically, Google Answers serves the needs of people who do not know how to do fundamental research, either in a library or on the Internet. So much information is moving from the brick and mortar library to the web that individuals are well advised to learn how to quickly and effectively mine the web for their nugget of information. Every individual can realize fantastic gains in productivity just by learning how to effectively search the web. As an example, while researching the material for another book, *Solving the 1897 Airship Mystery*, I spent hundreds of hours at the National Archives researching census records. After the census records were placed on the web by several different organizations, I could find the same information in a couple of hours. After studying the material in this book, you will be as well trained to find information, without undue frustration, as any Google researcher.

## Google Catalogs

Avid catalog shoppers will find Google Catalogs a boon. No more catalog clutter around the house. You can toss out your old catalogs and spend lazy afternoons shopping until your vision blurs and your head hits the monitor. Google Catalogs has something for everyone, including the person who has it all. There are enough catalogs in the Google Catalogs database to keep any well-respected shopper busy for the foreseeable future.



The screenshot shows the Google Catalogs beta homepage. At the top, there's a navigation bar with an address bar containing 'http://catalogs.google.com/' and a 'Go' button. Below the address bar is the 'Google Catalogs™ BETA' logo. To the right of the logo are 'Google Search', 'Advanced Catalog Search', and 'Catalog Help'. A main heading says 'Search and browse mail-order catalogs online'. Below it are links for 'Help Google Add More Catalogs' and 'Info For Catalog Vendors'. The page is divided into several sections of categories:

- Apparel & Accessories**: J. Crew, L.L. Bean, Lands' End, ...
- Computers**: Dell, PC Connection, ...
- Home & Garden**: Crate and Barrel, Gump's, Petsmart, Williams-Sonoma, ...
- Arts & Crafts**: Keepsake Quilting, Shar Music, ...
- Consumer Electronics**: B & H, Crutchfield, ...
- Lifestyle & Gift**: Crazy Crow Trading Post, RedEnvelope, ...
- Automotive**: Camping World, Stylin Concepts, ...
- Education**: Curriculum Associates, Discount School Supply, ...
- Sports & Outdoors**: Bart's Watersports, Eastbay, Golfsmith, West Marine, ...
- Books, Music & Film**: Acorn, Bas Bleu, ...
- Food & Gourmet**: Harry and David, Mackenzie Limited, ...
- Toys & Hobbies**: HearthSong, Leggo, ...
- Business to Business**: Life Uniform, OfficeMax, ...
- Health & Personal Care**: FootSmart, Garden Botanika, ...

At the bottom right, there's a link 'Browse complete list of catalogs...' and a standard Windows-style toolbar.

Figure 2.7: Google Catalogs

Google Catalogs is a directory of over 5,000 catalogs selling just about everything conceivable that is legal to sell. Most well-known catalog vendors are part of the database, including Neiman Marcus, L.L. Bean, Abercrombie and Fitch, and Frederick's of Hollywood.

Google Catalogs home page:

<http://catalogs.google.com> (Bookmark it!)

User FAQ:

<http://catalogs.google.com/googlecatalogs/help.html>

Google Catalogs vendor FAQ:

[http://catalogs.google.com/googlecatalogs/catalog\\_vendors.html](http://catalogs.google.com/googlecatalogs/catalog_vendors.html)

Left-clicking on any category displayed on the Google Catalogs home page will display a list of 28 catalogs in that category.

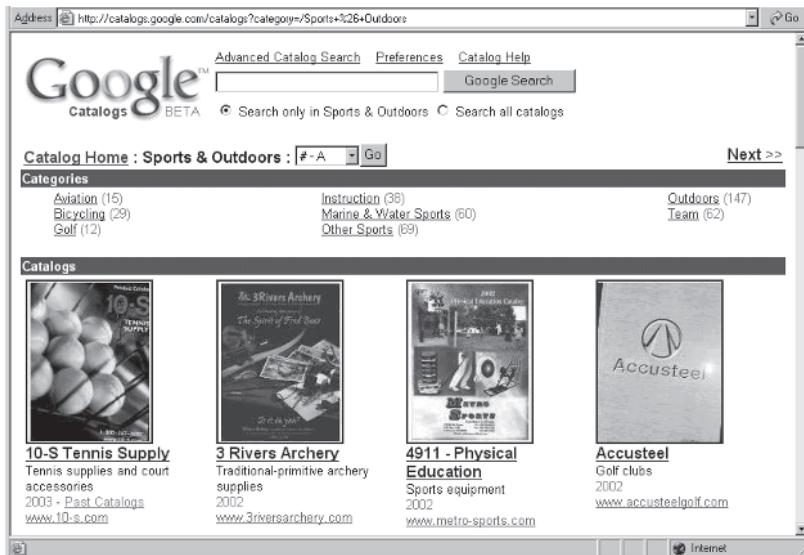


Figure 2.8: Sports & Outdoors catalogs

Let's check out a catalog. How about L.L. Bean? Under Apparel & Accessories (see left side of Figure 2.7), left-click on L.L. Bean.

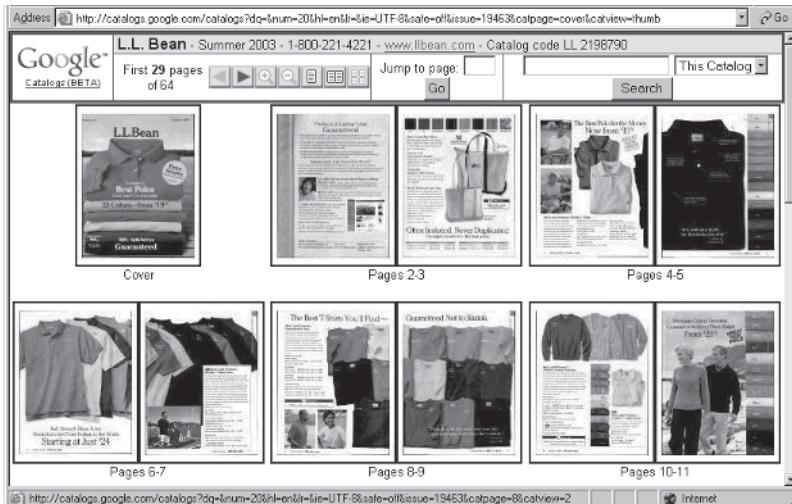


Figure 2.9: L.L. Bean catalog



When perusing the catalogs, notice the tools and vendor contact information at the top and bottom of the page. Figure 2.10 illustrates the information Google provides at the top of a vendor's page. You are provided both the telephone number and the web site of the catalog vendor. From either contact, you can place your order. From the vendor's catalog page, you can view a larger image by left-clicking on the page of interest.

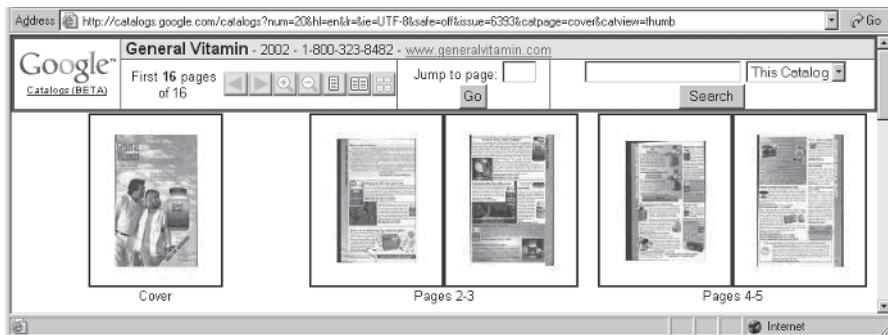


Figure 2.10: Vendor contact info

The toolbar at the top and bottom of the catalog's page has several items of interest. You can go to the previous page or the next page . You can magnify or reduce the image size. You can also modify how the catalog's pages are formatted and displayed. A left-click on any of the page formatting buttons will change the number of pages displayed per browser window. A left-click on will display a single page, while left-clicking on will display two pages, and left-clicking on will display four pages.

In addition to the toolbar tools, you can jump directly to any page of interest by left-clicking on the Jump to page box and entering a page number. Enter only a numeric value (1, 2, 3, 4...), then left-click on the Go button.

Jump to page:	<input type="text"/>	Go
---------------	----------------------	----

If you want to search the catalog you are currently viewing, left-click in the This Catalog search box, and key in your search word(s). Left-click on the Search button and you are taken to a search results page, if the item you searched on exists within

<input type="text"/>	This Catalog	Search
----------------------	--------------	--------



that catalog. If it does not, you are presented with a “did not match any pages” message. Return to the previous catalog page by left-clicking on the browser’s Back button. Here is a keyboard shortcut that works wherever you are, in any Internet Explorer window — press the Backspace key to return to the previous browser window.

Google scans the actual catalogs from catalog vendors, and then uses word recognition technology to return search results from the catalog database. The Google Catalogs search technology seems to have the same shortcoming that Froogle has — namely that it cannot distinguish between the primary item and accessories. Using the same term (“computer”) used to illustrate Froogle’s search technology returns both computers and computer accessories mixed in the results. This can be especially frustrating if you are looking for computers and have to shift through innumerable ancillary and/or accessorized items without finding the specific items that you are looking for. But do not give up yet! Google Catalogs has an advanced search page (see Figure 2.11), where you can modify your search query and narrow the field. However, the advanced features did not seem to be working at the time of publication. Let’s illustrate with an example.

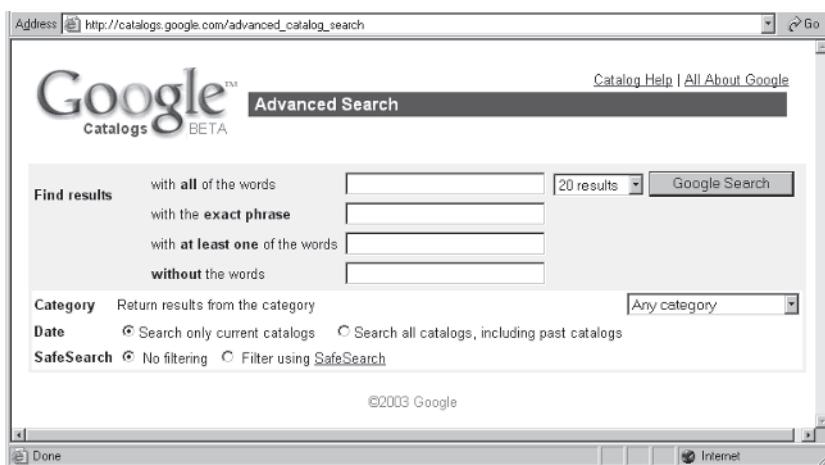


Figure 2.11: Google Catalogs Advanced Search



A search on the word “computer” returns page 1369 from Tessco Technologies’ catalog as the first search result, as shown in Figure 2.12. Left-click on the image shown at the right in the figure. Tessco Technologies offers for sale on page 1369 a Notepad IV Universal Computer Mount. (See Figure 2.13.)

Part No.	Description	Unit (\$)	Weight (lb.)	Qty/Unit
20637 EACH	Gumber Johnson Notepad III notebook computer mount, with short (3" 4" - 13.8") hold down clips	\$29.00	141.98	Continuing I EACH
20638 EACH	Gumber Johnson Notepad III notebook computer mount, with medium (1" - 1.5") hold down clips	32.00	137.40	
20639 EACH	Gumber Johnson Notepad III notebook computer mount, with long (1" - 2") hold down clips	32.00	135.11	
20641 EACH	Gumber Johnson Notepad III notebook computer mount, with Long (1"- 2") hold down clips	22.00	137.40	
20651 EACH	Gumber Johnson Cell Phone bracket for use with NP3 or UPR-24 computer mounts	19.00	11.78	
20652 EACH	Gumber Johnson NP-CRLL3 2 Yrs.	0.45	11.40	
20653 EACH	Gumber Johnson 12" map light with dimmer control for use with NP3 or UPR-24 computer mounts	99.00	11.21	
		0.26	28.41	<b>451.90</b> I EACH

Figure 2.12: Google Catalogs Advanced Search

Part No.	Description	Unit (\$)	Weight (lb.)	Qty/Unit
455320	NotePad IV Universal Computer Mount			
	Universal computer mount is a key component for any mobile computing application. Features a patent pending rugged steel design with black powder coating. Easily adjusts to fit most computers.			
	• New curved base design			
	• Offers front/rear and rear support brackets for added forward/backward resistance			
	• Swivels 60 degrees for driver or passenger			
	• Compression closure latch allows for fast size adjustments, quick removal/insertion with a key lock			
	• Adjusts to fit laptops 8.25" - 13.38" wide, 11" - 16.31" deep, 5" - 2.44" thick			
	• Accessories available for customizing			

Figure 2.13: Google Catalogs Advanced Search



A Notepad IV Universal Computer Mount is a computer accessory. I want to find real Dell, Compaq, and HP computers — you know, the things with circuit boards, microprocessors, power supplies, etc., not notepad mounts. If I wanted to find notepad mounts, I would have searched on “notepad mounts.” I want to search again and see if I can eliminate this page from the search results.

Note that page 1369 includes the word “Universal Computer Mount.” So, being the clever individual I am, I go back to the Google Catalogs Advanced Search page and I enter the phrase “Universal Computer Mount” in the **Without** the words text field and hit Search again. See Figure 2.14.

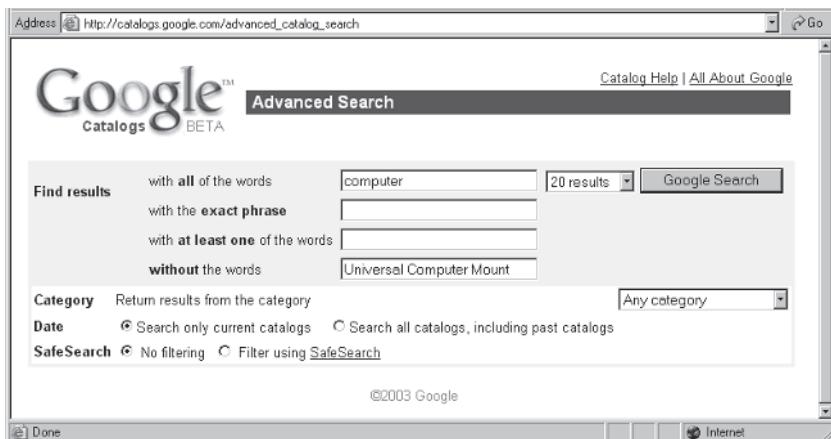


Figure 2.14: Google Catalogs Advanced Search

My new search results should not include Tessco Technologies’ page 1369 at all.

If the Google Catalogs search technology is mature and doing what search technology should be doing — namely what the owner says it should be doing — I will get back as the first search result a catalog page other than Tessco Technologies’ page 1369.

Guess what? I get back a page that says “Your search — computer –Universal –Computer –Mount — did not match any documents!” Amazing!

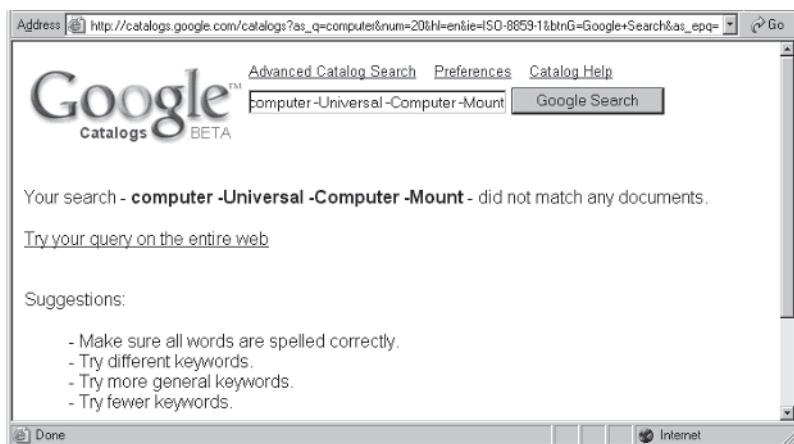


Figure 2.15: Google Catalogs Advanced Search

Well, one example does not a search engine make, so I tried five other examples and achieved varied results. But none of the results I achieved were the results expected.

You give it a go. Search on a favorite item, such as auto, travel, hotel, health, Red Cardinal, etc., from the Top 10 list (see Chapter 1). Try to narrow the field of search returns by selecting keywords that would eliminate a specific page. For example, if you search on “auto” in Google Catalogs you will get a lot of auto parts stores in the search returns. So, key in “parts” in the **Without** the words box on the Advanced Search page and see what you get. Interesting, eh? I did this and I got the Tessco Technologies page again. Google does state Google Catalogs is a beta version at this time. I imagine these types of issues will be addressed. When Google Catalogs search works the way it should, it will be a boon to shoppers.

Vendors can add their catalog (free!) to the database by adding Google to their subscriber list. If you are interested in this opportunity, view the Google Catalogs vendor FAQ (given earlier in this section) for the address.



## Google Groups

Google Groups is the Google incarnation of the Usenet discussion forums, also called newsgroups. You can post and read comments in the various discussion newsgroups. Discussion groups exist that cover virtually every conceivable topic. If you cannot find a discussion group that is already immersed in a subject of interest to you, you may easily start a new discussion group.

Google Groups provides a means for searching the Usenet groups that include both current and past postings. Google bought Usenet from Deja News in 2002 and has archived all posts, except binaries, back to 1981. That is over 700 million messages. Many researchers will find the ability of Google Groups to search the newsgroups a bonus in their research efforts.

Searching newsgroups is an interesting and entertaining affair. To give the subject the treatment it deserves, I have written an entire chapter on the topic; please see Chapter 10.

## Google Image Search

Google claims that its Google Image Search feature is the most comprehensive on the web with 425 million images currently indexed. You can access Google Image Search by left-clicking on the Images tab on Google's home page or going to <http://images.google.com>. You initiate an image search by entering a keyword in the search box. Then, left-click on the Search button. Google will display the thumbnails of images that it found, along with the associated web page link. To view the image, just left-click the thumbnail of interest displayed on the results page, and you will see a larger version of the image. Additionally, you will see the web page where the image is located.



Figure 2.16: Google Image Search

Photos and artwork are copyrighted, regardless of your personal perspective of the function of the Internet and web. Although you may gain the ability to view and save images found in your search, Google cannot grant you the right to use the images in any manner. If you desire to use the image in some manner, you should contact the owner of the web site where the image was found and ask for permission to use the image.

Google Image Search displays five rows containing four thumbnails per web page. Setting the number of results to display per page on the Preferences page does not affect the number of thumbnails. Setting the results per page is discussed in the Preferences section of Chapter 3.

The Google search engine dutifully displays as thumbnails all images it finds associated with the keywords that you used for the search, regardless of your sensibilities. You can turn SafeSearch on and have a modicum of filtering to remove potentially offensive material. For the obvious four-letter type words, SafeSearch works



pretty well. However, be aware that you can have some unexpected results when searching seemingly innocent keywords. SafeSearch is discussed in more detail in Chapter 3. Note that SafeSearch is available only from an English interface. It is worth repeating the Google warning about the possible results of your image search:

**Warning:** The results you see with this feature may contain mature content. Google considers a number of factors when determining whether an image is relevant to your search request. Because these methods are not entirely foolproof, it's possible some inappropriate pictures may be included among the images you see. (The mature content filter is only available from an English interface.)

At this point, you should know that you must go to the Preferences page at <http://www.google.com/preferences?hl=en> and set the level of filtering that you prefer.



### **Caution:**

If you are thinking about searching for images with a 56K dial-up modem, I would suggest that you forget the idea, unless you have some amount of patience. DSL/cable modems are sometimes slow to download the graphics-intense search results web pages. Although the thumbnail images are not large, there are 20 of them per search results page. In the following example search for "travel" images, my 1.5MB DSL modem took five seconds to load the search results web page. Your 56KB modem would take approximately one minute, depending, of course, upon your server traffic.

To explore Google Image Search, let's search on the word "travel." See Figure 2.17. Key the word in the query box and then left-click on the **Google Search** button just below the query box. Make sure you have spelled the keyword correctly! Note that we did not append the typical image suffix ".jpg" on the end of "travel." If we did, we would get a different search result. Leaving .jpg off the keyword gives us a search result that includes the image formats jpg, gif, and png.



Figure 2.17: Google Image Search for travel

The search results for “travel” are displayed in Figure 2.18.

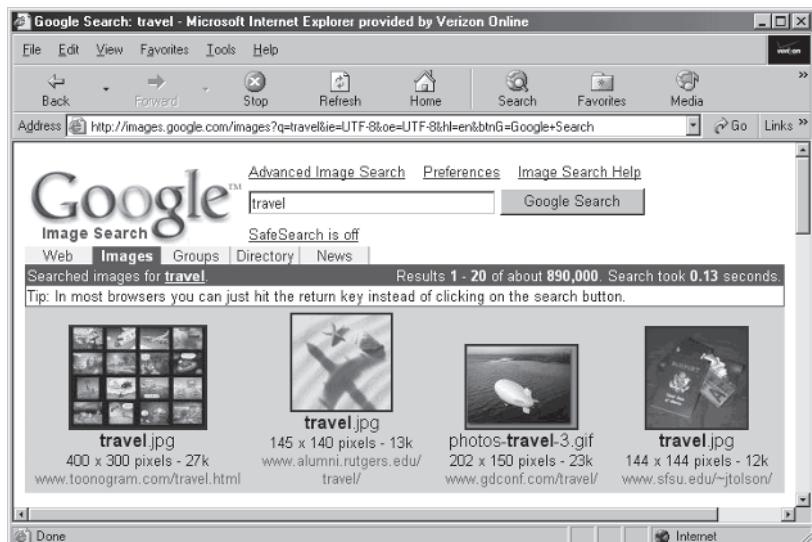


Figure 2.18: Google Image Search results for travel



Notice that underneath the search box there is a link (the underlined text) that allows you to set the level for filtering adult content. If you left-click on SafeSearch is off, the Preferences web page is displayed. From this page, you can set the filter level that you prefer. You do not have to click on the browser's Return/Back button to get back to the results page. As soon as you left-click on the Save Preferences button, the Google Image Search results will again be displayed.

To view the web page where an image is located, left-click on the travel image of your choice. Let's click the second image from the left in Figure 2.18 and see what we get.

The web page that includes the image we selected is shown in Figure 2.19.

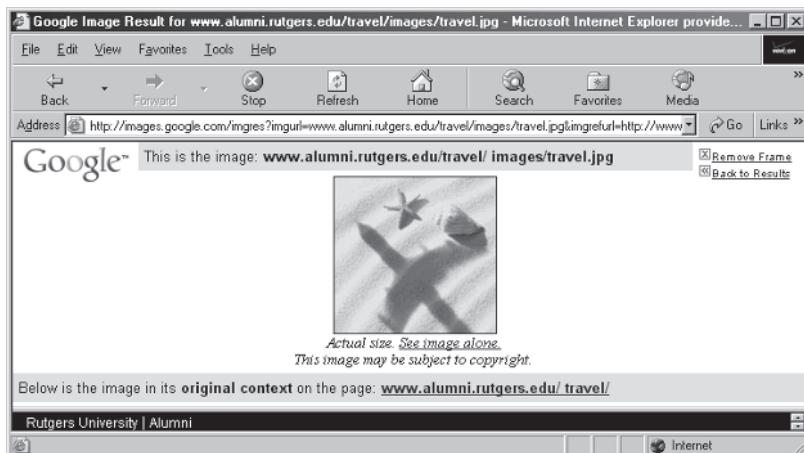


Figure 2.19: Google Image Search results for travel

There are several noteworthy items to point out while we are viewing this page. Note that the image we clicked on is shown at the top of the page surrounded by a frame. You can remove the frame by left-clicking on Remove Frame at the top right-hand side of the page. This image is shown in actual size. Sometimes the image is shown in a reduced size that you can enlarge by left-clicking on the image. The address of the web page where the image is found is shown below the image. It is noteworthy that this is not the actual



web site of the page. You are still inside a Google web page. Google provides a link from its page to the web site. Below the image there is a blue bar with the link [www.alumni.rutgers.edu/travel/](http://www.alumni.rutgers.edu/travel/). A left-click on that link will take you to the web site containing the image.

There is a Back to Results link below the Remove Frame link at the top right-hand side of the page. Most people are probably familiar with their browser's Back button and will navigate their way back using it. Having a Back link on a web page is a holdover from the early 1990s when web browsers were simple and had no Back or Return button.

If you want to see the image by itself and in its original size, left-click on it and the image will be displayed in its original size in a new window. See Figure 2.20.

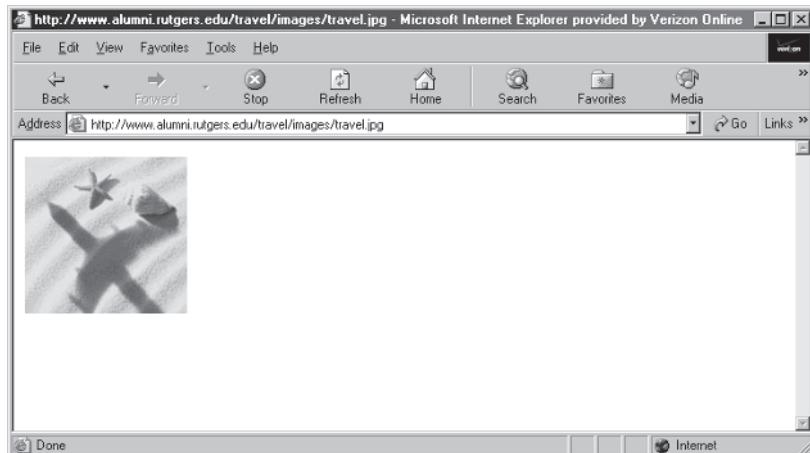


Figure 2.20: Google Image Search results for travel

Finally, you have found an image that is to die for, or at least one that you would like to keep for some purpose. How do you save that scrumptious image? Just right-click on the image and select Save Picture As in the drop-down menu that “magically” appears. See Figure 2.21.

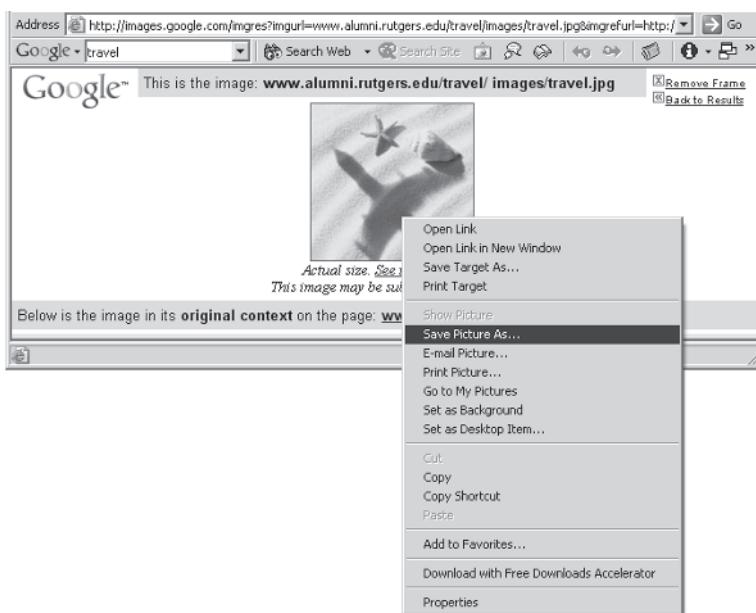


Figure 2.21: Saving an image from Google Image Search

This drop-down menu is an Internet Explorer service. This is not a book on Internet Explorer, but I point out that there are other neat, perhaps illegal, things you can do with the images. The drop-down menu includes choices that let you print the image, e-mail the image, or set it as the background on your computer. One option not shown on the menu in Figure 2.21 is available if you have purchased and installed a McAfee product called Visual Trace. Visual Trace is a program that allows the owner (me in this case) to determine who the hackers are (hopefully not you) trying to get inside my/his/her computer.



Did you perform an image search on the keyword “travel”? Note that the search did not yield any images of travel destinations in the first 20 images returned. To see images of a place we might contemplate for our next vacation, we must include a little more information in the search. Let’s pick Orlando, Florida, for our next vacation destination. Now let’s see what images we can find. Search on “travel Orlando.” See Figure 2.22. We received numerous images of Orlando, including two city maps (very useful).

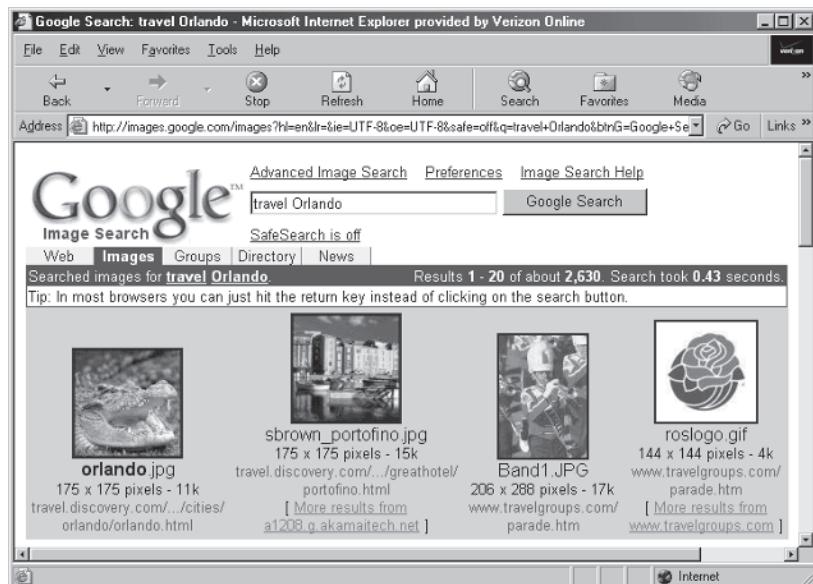


Figure 2.22: Google Image Search results for travel Orlando

Suppose that we want to see images of hotels in Orlando so that we can select one for our vacation. To do that, we search on “Orlando hotel.” The search results are shown in Figure 2.23.

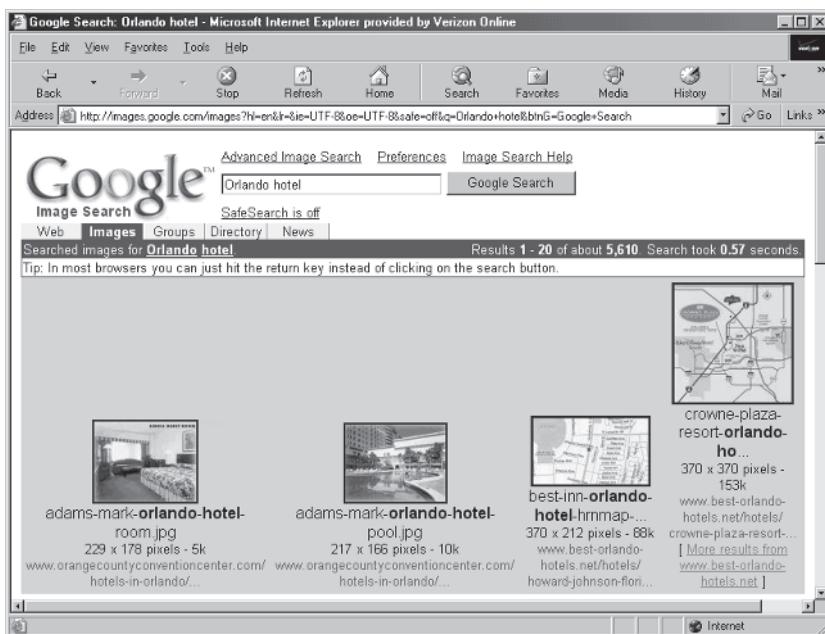


Figure 2.23: Google Image Search results for Orlando hotel

If we want to zero in on a specific hotel, we can search again and include the name of the hotel in the keyword set. Let's say that we are interested in the Disney World hotels. Then our search keywords would be "Orlando hotel Disney World." I have left this search for your own practice.

Google has an Image Search FAQ page available by selecting the Image Search Help link at the top of the page.

Google says it analyzes the text on the web page adjacent to the image, the image caption, and "dozens of other factors" to assess the image subject or content. Google also claims to utilize "sophisticated algorithms to remove duplicates and ensure that the highest quality images are presented first in your results."



Sometimes, you may inadvertently (or on purpose in some cases) wind up at a web site that will not let you close the window. Every time you click to close the window, it just opens another. Also, this type of web site tends to overwrite the monitor screen with its web page, removing your task bar from view and any other windows you may have open. This type of heavy-handed control of a computer by a third party is extremely annoying.

If you find yourself in a web site that will not let you close the window, do not despair. Just hold down the Ctrl, Alt, and Delete keys simultaneously. Then release the keys. Wait until you see a Close Program dialog box. Click on the line that includes Internet Explorer. See Figure 2.24. Next, click on the End Task button. That should close all IE windows that are open. If not, repeat the process until all of the annoying pages are closed.

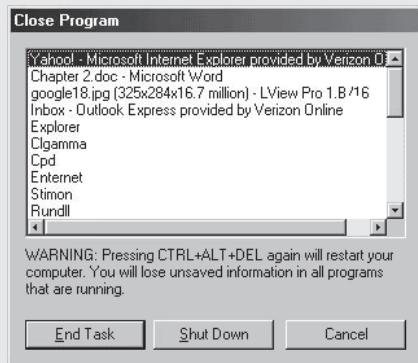


Figure 2.24: Close Program dialog

If you want to restrict your image search to specific file types, you can do it the hard (Google) way, the wrong (bright spark) way, or the easy (my) way.



## The Hard (Google) Way

Google, in its Image Search FAQ, specifies the following procedure for restricting an image search to a specific file type. Note that searching specific file types is an Advanced Image Search feature. The Advanced Image Search features are discussed at length in Chapter 7. Continuing with our search, to view images of a specific file type, such as jpg, gif, or png, use the Google Advanced Image Search web page located at [http://www.google.com/advanced\\_image\\_search?hl=en](http://www.google.com/advanced_image_search?hl=en). (Bookmark it? Yes.) Or you can navigate to the page by left-clicking on the Advanced Image Search link on the Google Image Search page. The link is located on the right side of the page. (Refer back to Figure 2.17.) To begin our search, enter “flower filetype:png” in the search box. Then left-click on Search. Alternately, you can just enter “flower” in the search box and then select PNG files from the Filetypes drop-down menu. See Figure 2.25.

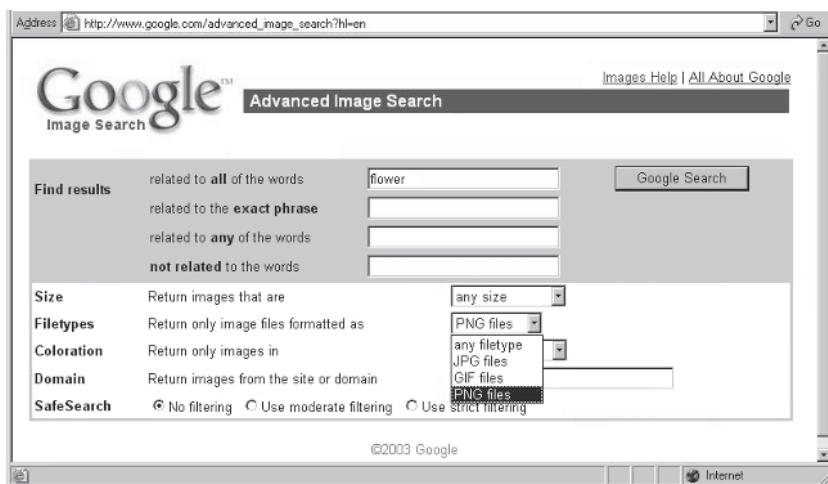


Figure 2.25: Advanced Image Search for file type png

The search results are shown in Figure 2.26. All images are restricted to the png file type.

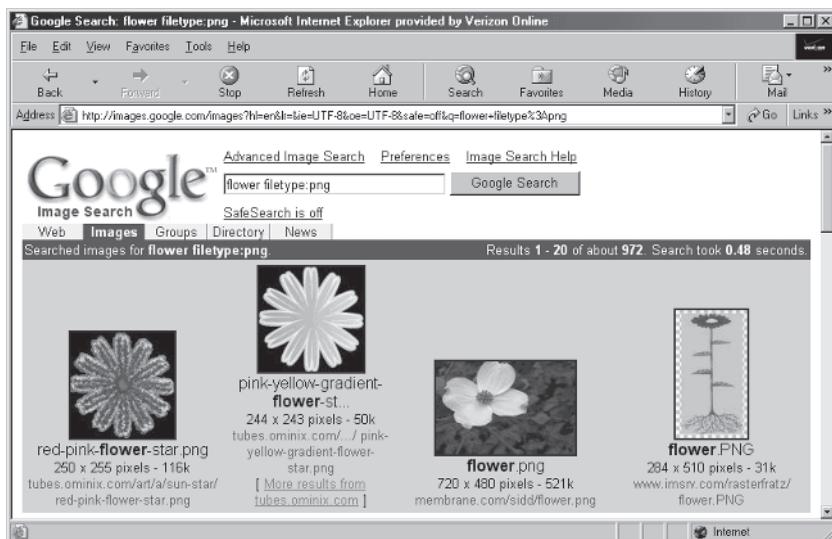


Figure 2.26: Advanced Image Search results for file type png

## The Wrong (Bright Spark) Way

I am including this procedure here because I know some bright spark is going to try it and then write either me or Google (or both) and complain that he did not get the correct results. How do I know this? Because I am one of those bright sparks who has tried it! Someone is going to try to restrict an image search using the Google Advanced Search web page as opposed to the Advanced Image Search page. So, let's go ahead and try it here to see what we get.

I entered the information as a file type search in the Advanced Search query box. See Figure 2.27. The results I achieved are shown in Figure 2.28. I did not get any images nor did I get any links to images. I have links to web pages that have the word "flower" and the suffix "png" somewhere in the page but no links to images. So, if you are searching for image types, do not try to search using the file type in the Advanced Search page.

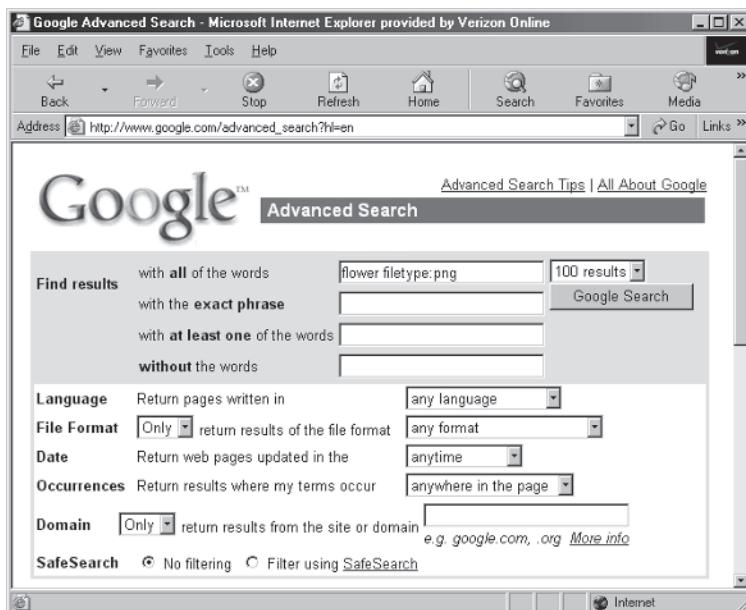


Figure 2.27: Search for file type png in Advanced Search

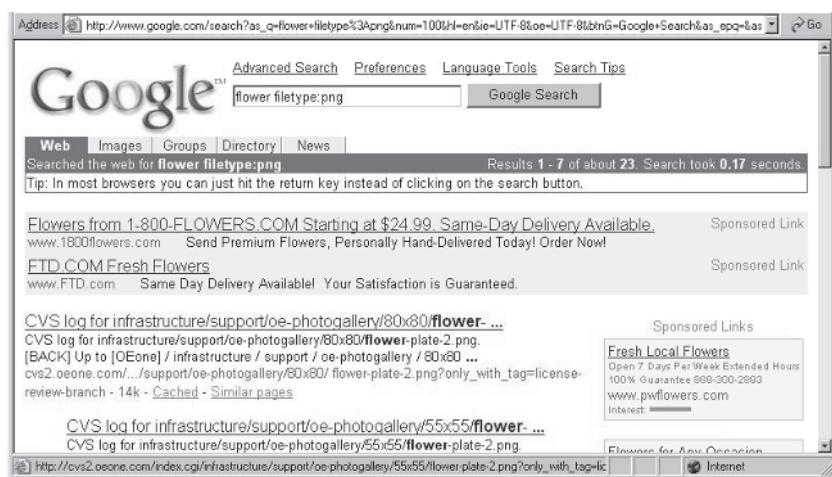


Figure 2.28: Search results for file type png



## The Easy (My) Way ... That Works!

While still in the Google Image Search web page, just enter “flower filetype:png” in the search box. Then left-click on Google Search. The following two figures illustrate how easy it is, and the results you get searching for “png”-type image files are identical to using the Advanced Image Search page. Compare Figures 2.26 and 2.30.



Figure 2.29: Search for file type png in Google Image Search

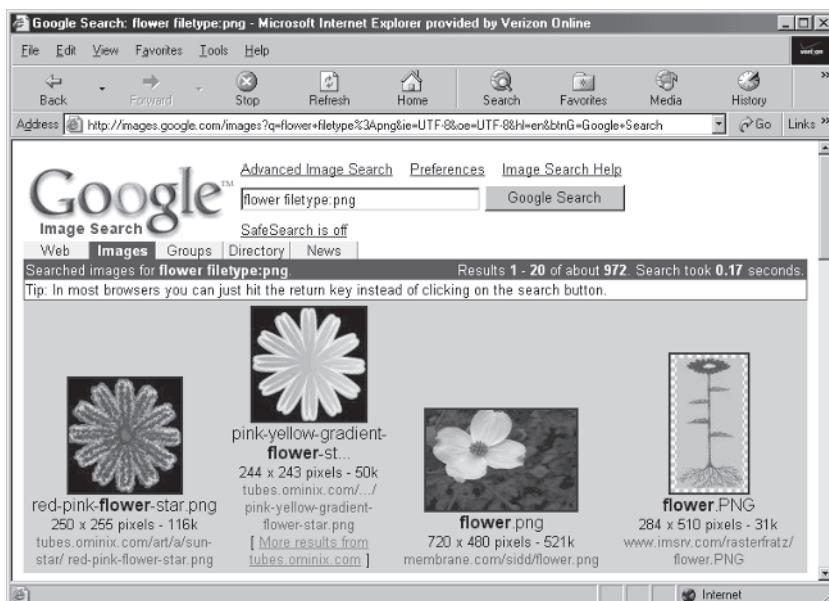


Figure 2.30: Search results for file type png

The common image file types are jpg, gif, tif, png, and bmp. Google only indexes images of file type jpg, gif, and png. A search on file type tif or bmp will return zero (that is, nada) search results. See Figures 2.31 and 2.32 to see what nada search results looks like.

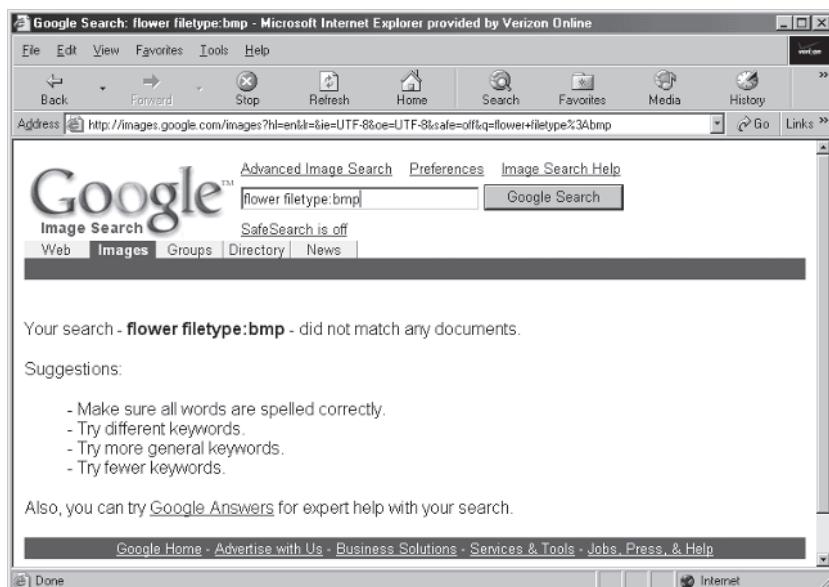


Figure 2.31: Search results for file type bmp

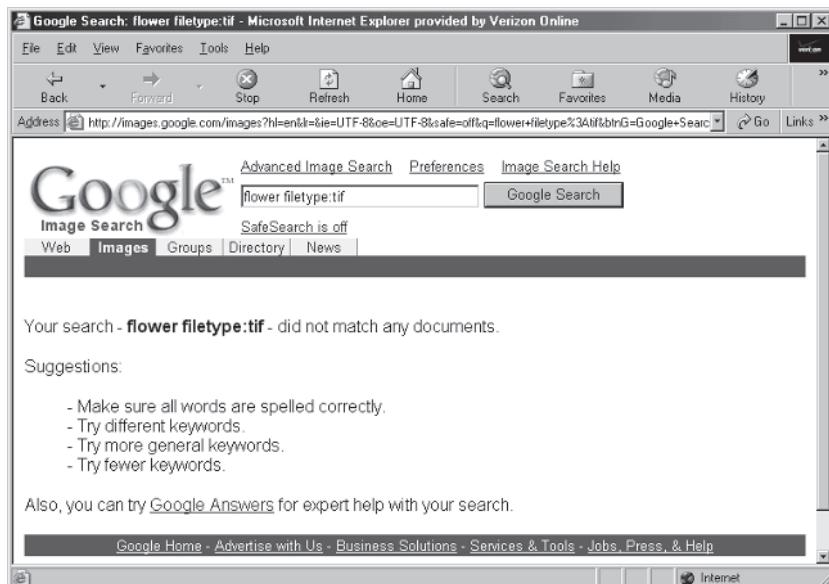


Figure 2.32: Search results for file type tif



Note that the suggestions to find a search match do not match the search circumstances. The suggestions refer to the keyword(s) and not the unsupported file type. The keyword, “flower,” is not the issue; it is the search for a file type, tif, that Google does not support. Google would be more informative by telling the user to search for a supported file type.

Additional image search issues are discussed in Chapter 7, “Advanced Image Search.”

## Google Labs

Google Labs is a showcase for ideas whose implementation is still being developed. Google Labs is located at <http://labs.google.com/>. (Bookmark it? Yes!) Google warns that these features are in the development stage and may perform erratically. Google also warns us that these features may disappear without warning. Everyone is encouraged to report any issues regarding any feature to the specific Google developer responsible for the feature. A developer’s lab is a fun place to visit. Come on, let’s have some fun!



### Note:

My experiments indicate that the features discussed in the following sections do not use SafeSearch. You may, especially if you search using the usual four-letter terms, see material intended for “mature, adult audiences.”

## Google News Alerts

News Alerts are useful for monitoring breaking news stories about international, national, regional, or local hot topics including industry or trade events, business and stock market issues, sporting events, weather issues, celebrity and/or entertainment info, governmental processes (new laws, regulations, elections, etc.), defense/military issues, religious events, and any other category of news you can imagine.

To begin receiving your News Alerts you must create a Google News Alert. To do this, go to <http://www.google.com/newsalerts>.



(Bookmark it if you want to use it.) See Figure 2.33. Enter the topic of interest in the News Search text box, then select the frequency — either Once a day or As it happens — by left-clicking on the down arrow of the How Often box and left-clicking on your choice. Next, enter the e-mail address where you wish to receive the News Alerts. Then, left-click on the Create News Alert button.

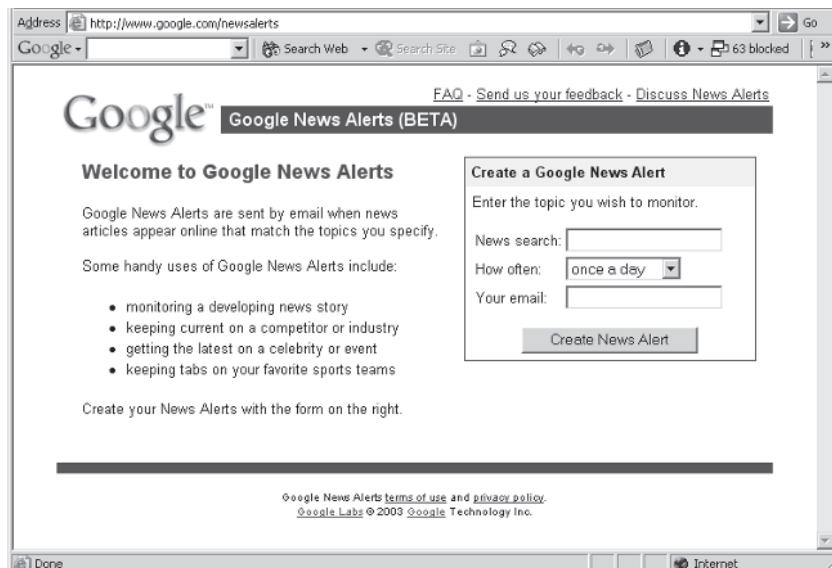


Figure 2.33: Google News Alerts

Next, you will receive a notification that a confirmation e-mail has been sent to the e-mail address you specified when you created the News Alert. See Figure 2.34. At this point, you can return to the News Alerts page and add another topic or continue doing something else.

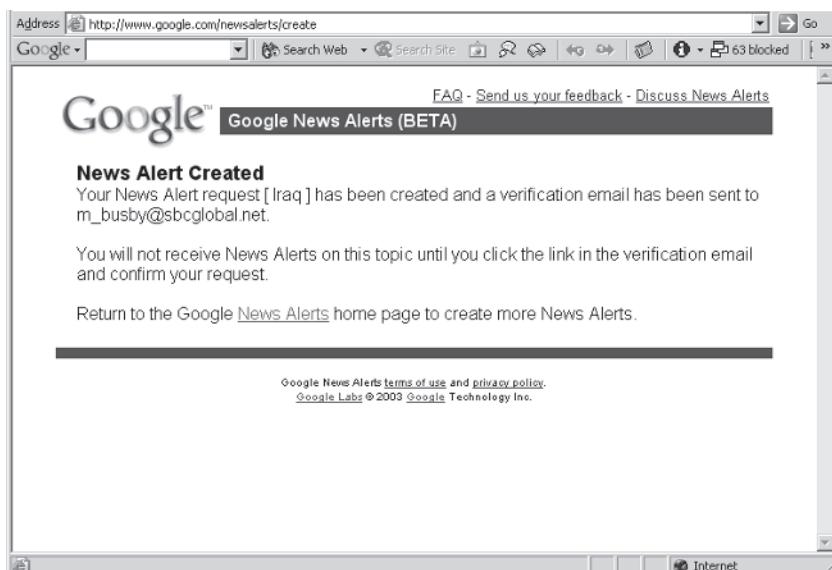


Figure 2.34: News Alert created

The confirmation e-mail you receive from Google will give you three choices. You can verify the News Alert, cancel it, or return to the Google News Alerts page. See Figure 2.35.

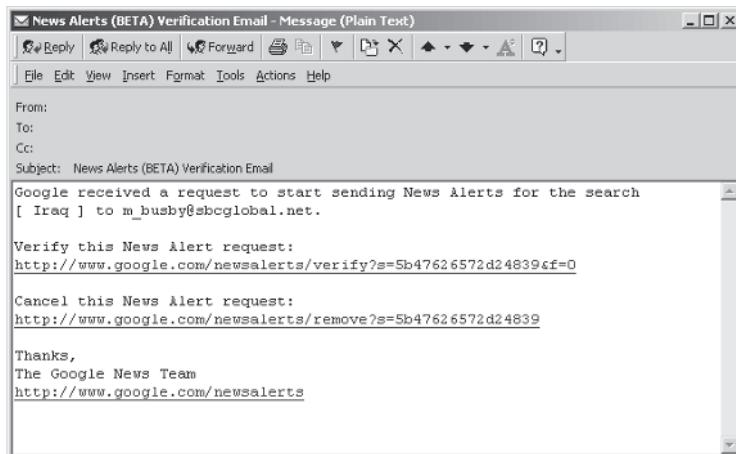


Figure 2.35: News Alerts verification e-mail



If you decide to cancel, left-click on the Cancel this News Alert request link. You will see Figure 2.36. Either left-click on the Return to the Google News Alerts home page link or go do something else.

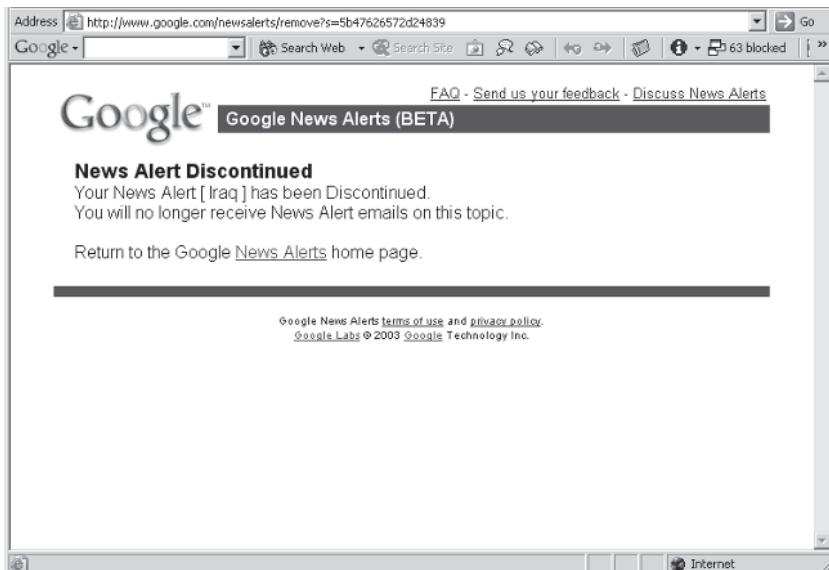


Figure 2.36: News Alert canceled/discontinued

If you decide to verify the alert request, left-click on the Verify this News Alert request link. You will see Figure 2.37.

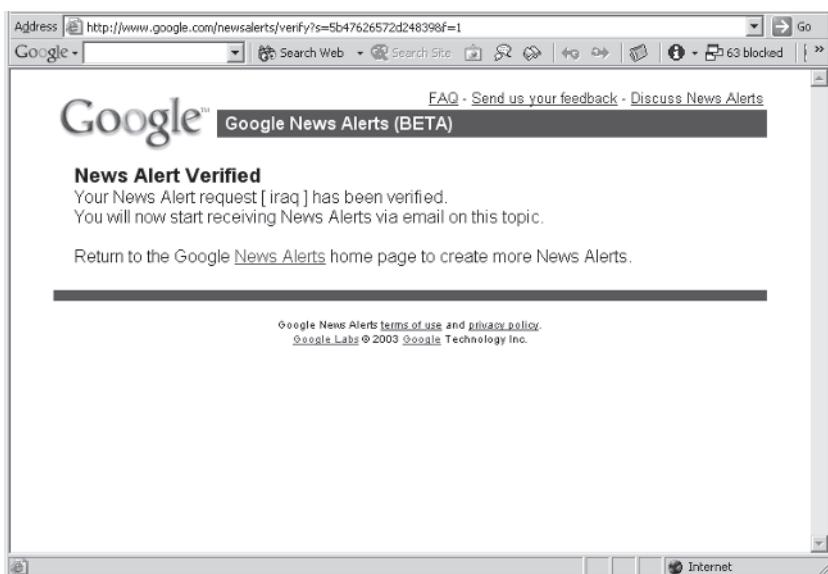


Figure 2.37: News Alert verified

Left-click on the Return to the Google News Alerts home page... link to set another News Alert, or go do something else.

Notice in Figure 2.37 that there is no means for canceling a news alert after you have verified one. However, when you are notified by Google of a news story, the e-mail will contain not only a link to the story but also a link allowing you to discontinue the alert, if you so desire.

Try it!

You can use advanced search features with the keywords/topics for which you want to receive alerts. The Google News advanced search page provides various ways to fine-tune a search to eliminate unwanted results. You can incorporate the search criteria of interest into your Google News Alerts settings. Select the criteria you want on the Google News advanced search page, then left-click the Google Search button. When the search results page appears, copy the resulting text that appears in the query box on that page. Paste it into the News Search text box on the Google News Alerts home page.



Currently, there's a limit of ten unverified News Alerts per e-mail address. If you've reached your limit, you will receive an error message, so check your e-mail and verify your pending News Alerts by clicking on the verification link.

## Google Compute

Google Compute enables your computer to share the computing needs of some endeavor that Google has identified as worthwhile. That is, Google Compute enables your computer to participate in selected collaborative computing opportunities. Google reviews and approves the collaborative computing opportunities. Such opportunities include many scientific research projects that require huge computing resources. Currently, Google has only approved one opportunity, a nonprofit academic research project at Stanford University that is trying to understand the structure of proteins so that better treatments for a number of illnesses can be developed. You can find Google Compute at <http://toolbar.google.com/dc/offerdc.html>. (Did you bookmark it?)

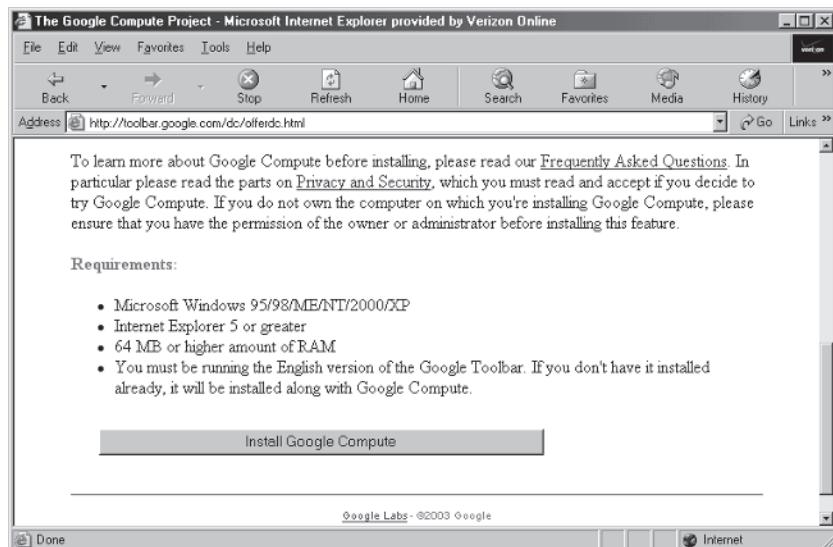


Figure 2.38: Google Compute installation



This is such a worthwhile endeavor that will benefit all of mankind, and you can be a part of it simply by “loaning” a part of your computer’s idle time to the project. Google Compute will install a button on the Google Toolbar. By clicking the button, you can enable or disable the collaborative computing program. To install this button on the toolbar, left-click on the Install Google Compute button at the bottom of the Google Compute page. The next four figures show the installation process.



Figure 2.39: Google Toolbar installation

Over my DSL connection, it took about two minutes to download the Google Toolbar.

In the dialog box shown in Figure 2.40, click on Yes to install the toolbar.

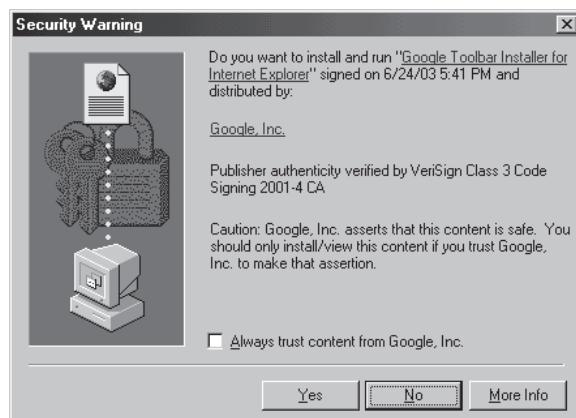


Figure 2.40: Google Toolbar installation



Be patient. Even though the little green speed bars stop moving as they approach the right side of their box, the download is still progressing.

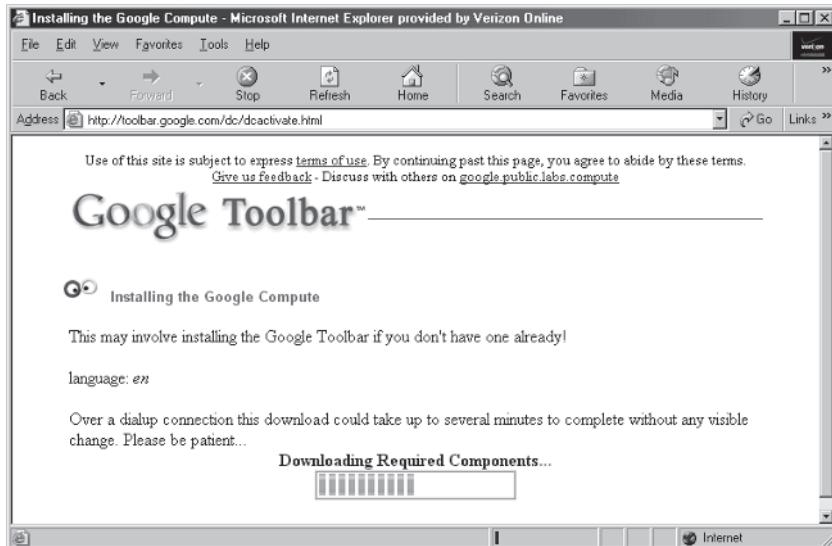


Figure 2.41: Google Toolbar installation

My firewall wanted to know if I would let the toolbar install program access the Internet. I selected Yes, allow this time, as shown in Figure 2.42.



Figure 2.42: Google Toolbar installation firewall



Finally, the download is complete. Notice the new Google Toolbar at the top of the page under the Address line. See Figures 2.43 and 2.44.

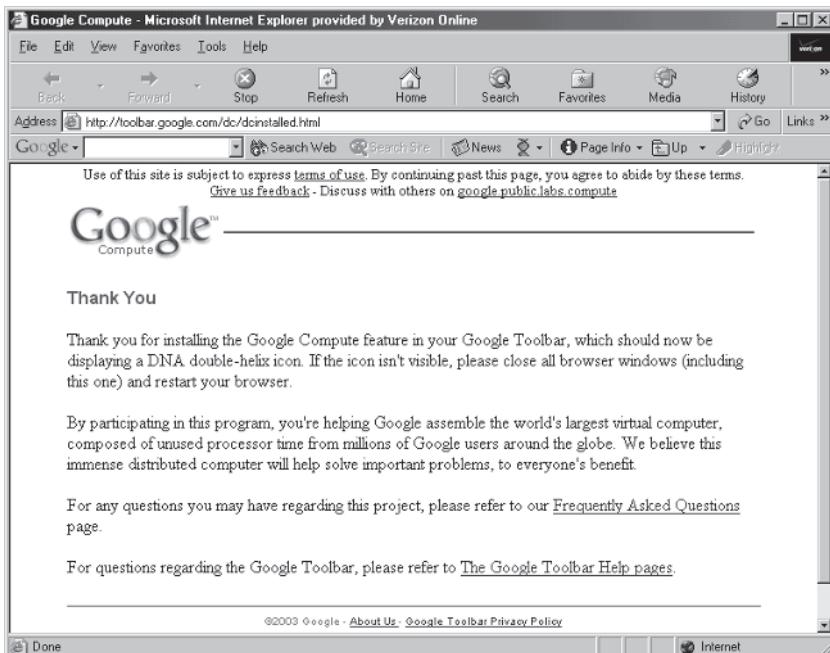


Figure 2.43: Google Toolbar installation



Figure 2.44: Google Toolbar

If you use your computer in an intense computing environment (that is, if you have a lot of windows open and you are working your processor to death or you are playing the latest whiz-bang game), do not start Goggle Compute. Leave it off. Turn on Google Compute only when you are not using the computer to do your work or have fun. A good time to use it is after your work for the day is finished. Let the company pay for the little electricity that will be used overnight. In that manner, the miserly curmudgeon of a boss you work for will benefit mankind even against his or her wishes. Just turn off your monitor so that you do not burn it out. For those



individuals with managed resources (boo-hoo!), ask your network administrators to add the Google Toolbar, without necessarily specifying why you want or need the toolbar — other than it will increase your productivity. You can outfox the foxes guarding the henhouse!

Turn Google Compute on by clicking the double helix icon in the Google Toolbar and selecting Start Computing from the drop-down menu. See Figure 2.45. Turn it off by clicking on the double helix icon and selecting Stop Computing. Easy, eh? Just a click before you leave work (or go to bed if at home) and a click in the morning before you start the day. Google Compute has just enabled everyone in the country to be a hero or heroine.

**Note:**

When you install Google Compute, it defaults to on.



Figure 2.45: Google Compute drop-down menu

Note the Switch to Conservative Mode option on the drop-down menu. Google Compute includes two modes of operation. One is standard mode, and the other is conservative mode. You may tend to forget to turn Google Compute on or off at the end of the day. If so, click on the conservative menu choice. Standard mode allows the collaborative computing software to run in the background, using computer cycles even while you are working at the computer. Conservative mode, on the other hand, keeps the software from running until it detects that you are not using the computer. The detection mode that the software uses is the time passed since you last hit a key on the keyboard or moved the mouse. The best choice is to use conservative mode with Google Compute turned off when you are working at the computer.



Google specifies the following requirements for running Google Compute:

- Microsoft Windows 95/98/ME/NT/2000/XP
- Internet Explorer 5 or greater
- 64MB or higher amount of RAM
- The English version of the Google Toolbar

If you decide collaborative computing is not for you and you must remove the Google Toolbar, enable Add/Remove Programs for your machine. The steps to do this are operating system dependent. If you need help, left-click on Start in the Windows toolbar. Select Help and search on “Add/Remove Programs.” Once you have the dialog box open (see Figure 2.46), select Google Toolbar for Internet Explorer. Left-click on Add/Remove. Then follow the remaining instructions that appear in the dialog boxes.

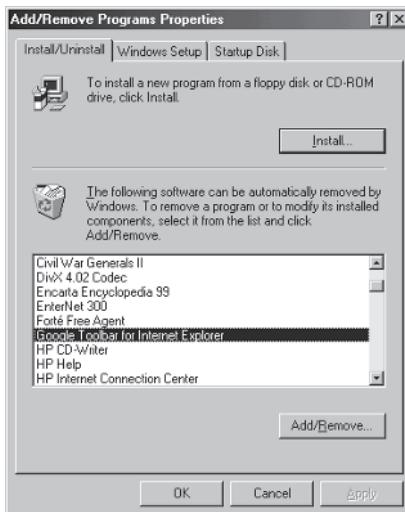


Figure 2.46: Removing the Google Toolbar

There is an excellent Google Compute FAQ page at [http://toolbar.google.com/dc/faq\\_dc.html#using1](http://toolbar.google.com/dc/faq_dc.html#using1). (Did you bookmark it?) Install Google Compute and help the world!



## Google Viewer

Google Viewer is a Google experiment to show you the search results in a scrolling slide show. Google Viewer is located at <http://labs.google.com/gviewer.html>. (Bookmark it now!) Google Viewer displays the web pages found in a search as an image in the slide show. Each image, or “slide,” is accompanied by a short text description. Figure 2.47 shows the Google Viewer page. Note that this Google Labs feature is not integrated into the general suite of Google Services.



Figure 2.47: Google Viewer

To check out Google Viewer, I searched on “flower.” The search results seem to be the same as the results I get using Google’s basic search tool (<http://www.google.com/>). The slide show uses standard audio/video icons for the start, stop, rewind, and fast-forward controls. See Figure 2.48. The control icons are placed just above the web page text in a sleek toolbar. You can speed up the slide show by left-clicking on the rabbit icon on the right side of the toolbar or slow it down by left-clicking on the turtle.

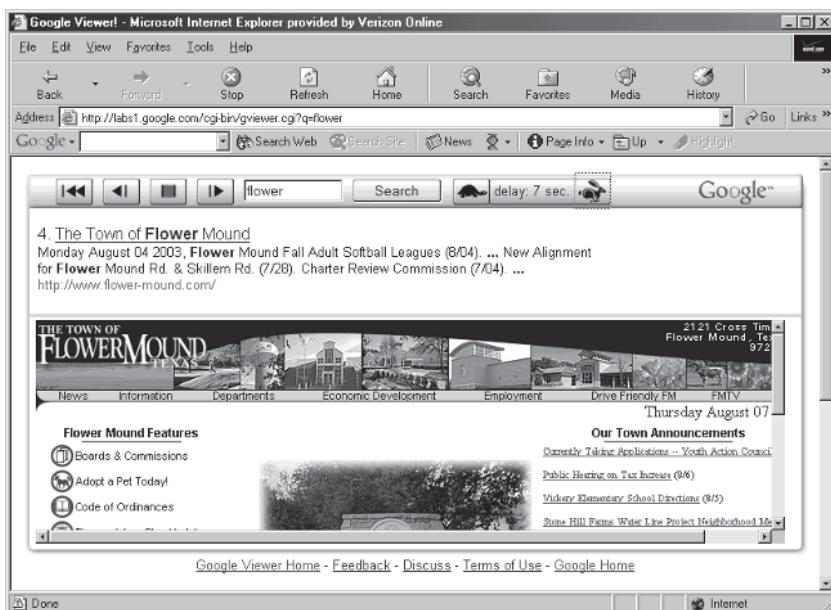


Figure 2.48: Google Viewer search results

There are a couple of requirements for Google Viewer:

- JavaScript enabled
- PC and Mac: Internet Explorer 5 and above or Netscape 6 and above
- Unix: Mozilla

Go ahead — experiment and have some fun!

## Google WebQuotes

When I first saw Google WebQuotes, I jumped to the conclusion that it was just another stock market quoting service. Boy, was I wrong! Google WebQuotes adds links to pages that comment on the web pages in your search results. This is a really neat way to find out what other people are saying about web sites. Google Web-Quotes is found at <http://labs.google.com/cgi-bin/webquotes>. (Bookmark it? If you like to read what other people say about your favorite web pages, including your own, then yes; otherwise, no.)



You can view all WebQuotes for a given or particular site by searching on that web site's keywords. To find out what people say about eBay, just search on "eBay."

To check out Google WebQuotes, I entered "travel hawaii" in the query box and left-clicked on Google WebQuotes Search. The result is shown in Figure 2.49. The first search result is "Hawaii Vacations Vacation Packages." To the right of the first search result is a link to two WebQuotes. Since I had selected 10 in the Find up to text box, the two WebQuotes are displayed directly under the search result (not visible in the figure).

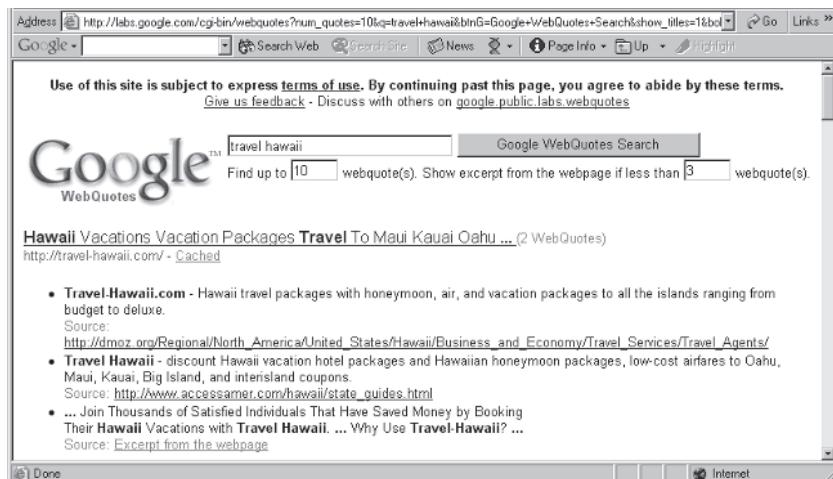


Figure 2.49: Google WebQuotes

In another example, not shown, I searched on "civil war." The first result, "The American Civil War Homepage," had 41 WebQuotes. When I clicked on the 41 WebQuotes link, all 41 quotes were displayed underneath the search result. I am not sure why the Find up to box is limited to 10, yet 41 WebQuotes were displayed for this example.

Is there a web site that you love? Go ahead and search for it and find out if others love it as much as you do. Hate a site? Search for it and see if you are not alone. Do you own a site? Search for your site and see what people are saying about it. Are you interested in



making a purchase from a web site but reluctant because the site's reputation is unknown to you? You can quickly determine a site's credibility by using Google WebQuotes.

I wanted to find out what people were saying about me on the web, so I keyed in my name in the query box and hit the WebQuotes search button. Unfortunately (or maybe fortunately), there were no search results. However, if I search on my name in the basic Google query box, I get 58,100 hits. So Google filters the search results in some manner. To discover what Google uses as search criteria, I next searched on the name of a favorite actor: Mel Gibson. I received a few hits, and from looking at them, it appears Google WebQuotes only returns search results when the keywords are in the web page's title.

This is another one of those "it's fun to experiment with" features of Google. Try your hand at it by searching on various company names. Find out what friends and foes are saying about the company you or your competition works for. Search on such companies as IBM, EDS, Microsoft, Ford, GM, and Martha Stewart. Enjoy.

## Google Glossary

Google Glossary returns definitions and related words and phrases of the keyword searched on.

Google Glossary is found at <http://labs.google.com/glossary>. (Bookmark!) Figure 2.50 shows the Google Glossary page. Enter the keyword in the query box, and left-click on Google Glossary Search.

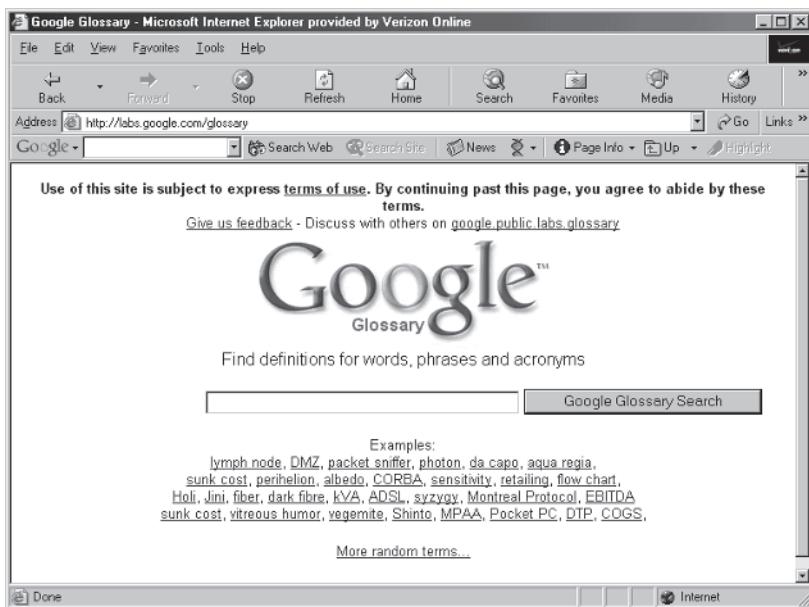


Figure 2.50: Google Glossary

I tested the word “travel” and received the result shown in Figure 2.51. Note that the definition is from a web site’s glossary. The other definitions not shown were from various web sites also. In other words, the definitions are not from a Google-owned or -maintained web site.



The screenshot shows a web browser window with the following details:

- Address Bar:** http://labs.google.com/glossary?q=travel&btnG=Google+Glossary+Search
- Toolbar:** Back, Forward, Stop, Refresh, Search Web, Search Site, News, Page Info, Up, Links, Go.
- Content Area:**
  - A message: "Use of this site is subject to express terms of use. By continuing past this page, you agree to abide by these terms."
  - A link: "Give us feedback - Discuss with others on [google.public.labs.glossary](#)"
  - The Google logo with "Glossary" below it.
  - A search bar with "travel" typed into it and a "Google Glossary Search" button.
  - A note: "Look up travel at [Dictionary.com](#) or [Merriam-Webster](#)"
  - A section titled "Related phrases" with links: Travel Glossary, Travel status, Travel Time to Work, travel agent, Travel trailer, Travel Time, Travel Agency, Travel manager, Annual Multi-Trip Travel, Latin America Travel Experts, Senior travel consultant, Class Of Service/travel.
  - A section titled "Definitions for Travel from the web" with a bullet point:
    - Sites that provide travel services and information. These include online travel guides, travel tools and online agents for purchasing tickets, hotel locators, and travel writing.  
[http://www.webbyawards.com/main/webby\\_awards/cat\\_defs.html](http://www.webbyawards.com/main/webby_awards/cat_defs.html)

Figure 2.51: Google Glossary search results

This is a useful site that should be in your bookmarks!

Be aware that SafeSearch does not work on this Google site.

## Google Sets

Google Sets will use the keywords that you enter to “predict” a set of related items. This feature seems to be a primitive thesaurus. Google Sets is found at <http://labs.google.com/sets>. (Don’t bother to bookmark.) The Google Sets page is shown in Figure 2.52.



Feedback Discuss Terms of Use

# Google Sets™

Automatically create sets of items from a few examples.

Enter a few items from a set of things. ([example](#))  
Next, press *Large Set* or *Small Set* and we'll try to predict other items in the set.

- 
- 
- 
- 
- [\(clear all\)](#)

[Large Set](#) [Small Set \(15 items or fewer\)](#)

Examples:  
green, purple, red stanford university, harvard university jerry maguire, mission impossible, top gun armani, versace

Figure 2.52: Google Sets

To see how it works, place a few keywords in the text boxes, one keyword per box. Then left-click either the Large Set button or the Small Set (15 items or fewer) button. The results are displayed in a new window. I tested the two keywords “cup” and “saucer.” The result was a list of eating appliances, such as spoon, fork, knife, plate, etc. I tested several other related keywords, and the results were not impressive. To be honest, I much prefer my *Roget’s Thesaurus*.



## Google Voice Search

Google Voice Search permits you to call a Google phone number and enter your keyword(s) verbally. The search result is displayed on your computer. The Voice Search page is located at <http://labs1.google.com/gvs.html>. (Bookmark? Definitely not.) Note that a toll call must be placed to utilize this feature. The Voice Search page is shown in Figure 2.53.

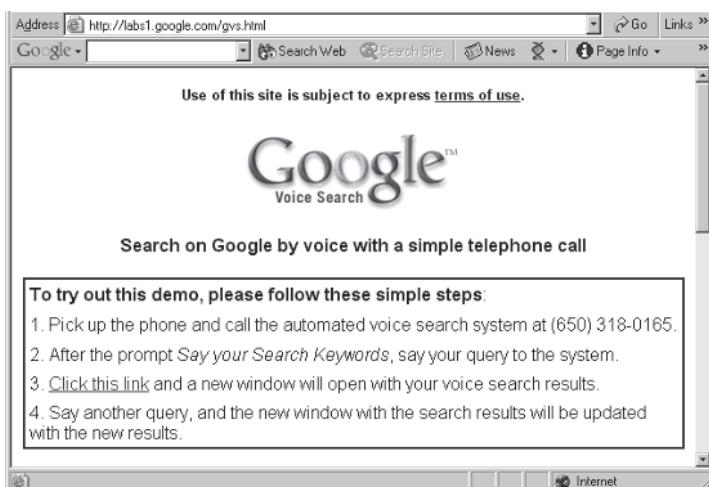


Figure 2.53: Google Voice Search

Instructions to use this feature are displayed on the Voice Search page. They are self-explanatory and easy to follow.

## Keyboard Shortcuts

Keyboard Shortcuts is a page detailing the Google shortcut keys that you can use when searching using the Keyboard Shortcuts query box. Keyboard shortcuts do not work in a search from the Google basic and advanced search pages. If you try to use them from another search site, you may experience the immediate termination of your browser. Keyboard Shortcuts is located at <http://labs.google.com/keys/index.html>. (Bookmark? Maybe.) Keyboard



shortcuts allow you to scroll through search results without taking your hands off the keyboard.

**Table 2.1: Keyboard shortcut keys**

Shortcut Key	Function
A	Search again
C	View Google's cached page
I	Scroll up
J	Move to the left
K	Scroll down
L	Move to the right
N	Next page
P	Previous page
S	Similar pages
Enter	Follow link
Backspace	Return to search results page
Alt + left arrow	Return to search results page (Netscape 6 only)
0-9	View corresponding result (0 = tenth result)
? or /	Displays a help page detailing the shortcut keys

When you press any key for the first time, an icon is displayed (➊). Press K to begin scrolling down. Have fun.



#### Note:

Press L to move the icon over to the paid advertising on the right side of the browser window. Press J to move back to the search results.



#### Note:

Java scripts must be enabled. If you have a pop-up killer active, including Google's own toolbar pop-up killer, Keyboard Shortcuts will not work.

To search from the Keyboard Shortcuts page, the system requirements are:



- Internet Explorer 5+, Netscape 6, or Mozilla
- JavaScript enabled
- Cookies enabled (optional)
- For Macintosh, Google recommends Netscape 6 or Mozilla. Google Keyboard Shortcuts is not fully functional on Mac Internet Explorer.

Shortcuts are fun to play with but not very useful to a mouse-dependent computer user like me. However, if you like keyboard shortcuts, this is just the thing. Power users who can remember shortcut keys find them much quicker to use than the mouse.

## Google News

Google News is located at <http://news.google.com/>. (Bookmark it for sure!) Instead of one source for a story, Google News provides news stories from many different sources (as many as 4,500), so you get different perspectives of an event. Google News is much better than a single-viewpoint newspaper. The sources include newspapers, news services such as CNN, and wire services such as UPI. Google News organizes the news by categories. See Figure 2.54.

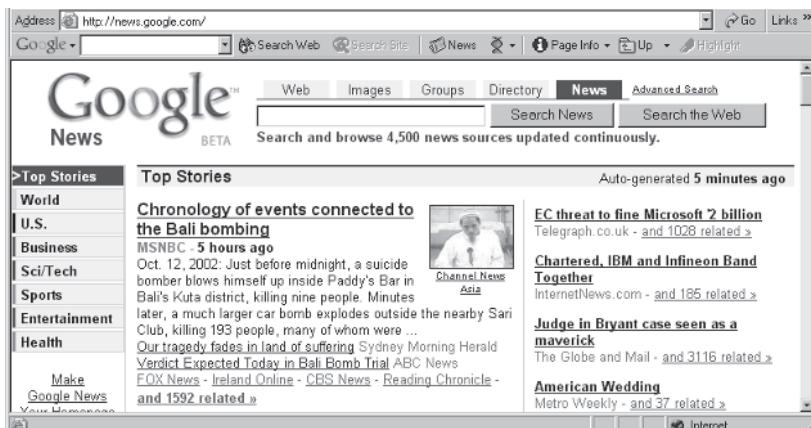


Figure 2.54: Google News



A nice feature of Google News is the ability to search within the news stories or the entire web. A search query box with search buttons are positioned at the top of the Google News web page. Simply enter your keyword(s) into the query box and choose your search preference.

I suggest that you try Google News as your home page. Then, every time you open your browser, you are presented with the current top stories of the moment. To make Google News your home page, left-click on the Make Google News Your Homepage link on the left side of the Google News page.

Finally, if you have dial-up Internet access, you may want to speed things up a bit by selecting a text version of Google News. Google provides a link on the left side of the Google News page to view the text version.

## Google Special Searches

Google Special Searches confines a search to a particular topic. Google Special Searches is located at <http://www.google.com/options/specialsearches.html>. (Yes, bookmark it!) The topics of Google Special Searches are:

- U.S. government
- Colleges and universities
- Linux
- BSD
- Apple Macintosh
- Microsoft

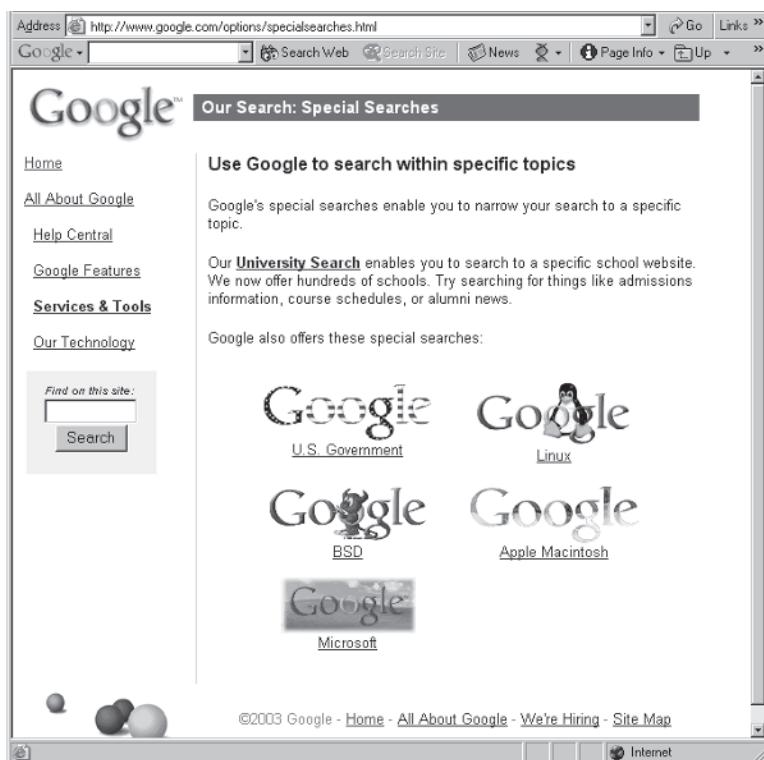


Figure 2.55: Google Special Searches

To search any of these special topics, just left-click on the link identifying the topic. I found the U.S. Government Special Searches to be very informative and entertaining. Find out what the government is talking about and doing by searching on keywords such as "legislation," "vote," "Clinton," "revolution," "war," "White House," "Army," "Navy," "Marine Corps," "Air Force," and "Iraq." Are you tired of foul language in the media? Want to find out what the government is saying about obscene words used in the media? Just key in the word of interest and left-click on Google Search. See Figure 2.56.



Figure 2.56: Google Special Searches

## Google University Search

Google University Search is an alphabetical listing of colleges and universities where you can select one to search. See Figure 2.57.

Google University Search is located at <http://www.google.com/options/universities.html>. (Bookmark it — even if you have no college-age or -bound children. Your niece, nephew, or neighbor might need the help.) This is a powerful tool for recruiters and salespeople. They can use the search feature to mine e-mail addresses and other contact information, such as telephone numbers. Then, using whatever piece of information they have acquired (e-mail addresses or telephone numbers), they can do a reverse lookup to determine other contact information. It also lets you impress the prospective candidate by feeding back information that you have found in your research. Everyone likes to boast about the university that they graduated from.

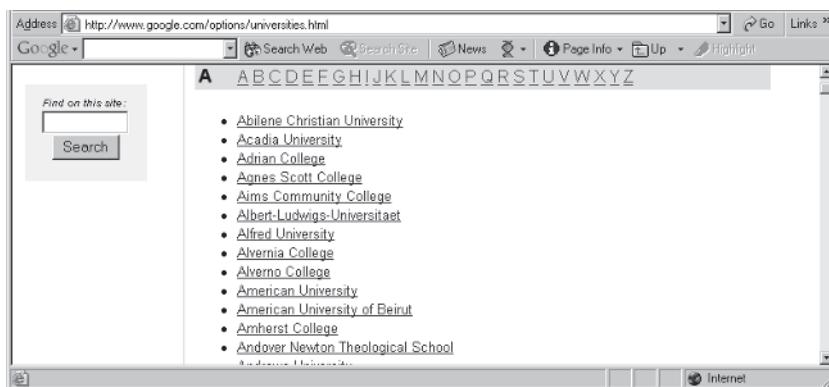


Figure 2.57: Google University Search

After you select a college/university to search for from the list, click on the college/university name, and you will be taken to the search page containing the search query box. See Figure 2.58. Key in your search word(s) and then left-click on Google Search.

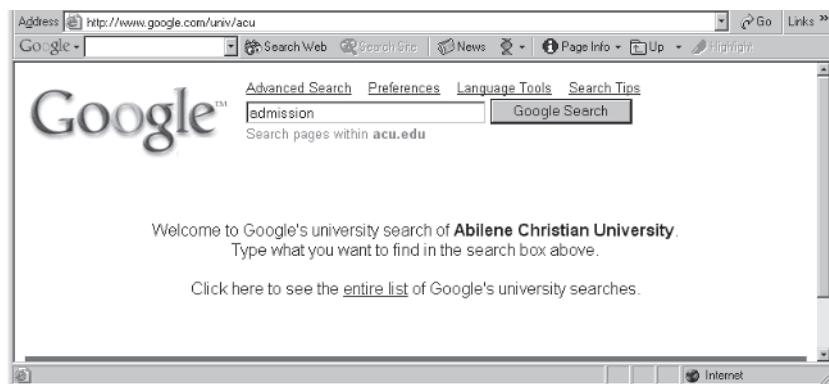


Figure 2.58: Google University Search

Google University Search is useful for zeroing in on your search objectives if you are looking for college/university information. A search restricted to the specific college/university domain (if you know the college/university domain name) will accomplish the same task as Google University Search. Performing a domain-restricted search is discussed in Chapter 5, “Advanced Google Search Techniques.”



## Google Web Directory

Google Web Directory is a yellow pages-style directory of web sites. Listings are segregated by category and subcategory. The directory indexes virtually every conceivable type of web page, including personal, business, commercial, government, academic, institutions, and not-for-profit organizations. Web Directory Search is located at <http://directory.google.com/>. (Bookmark it!)

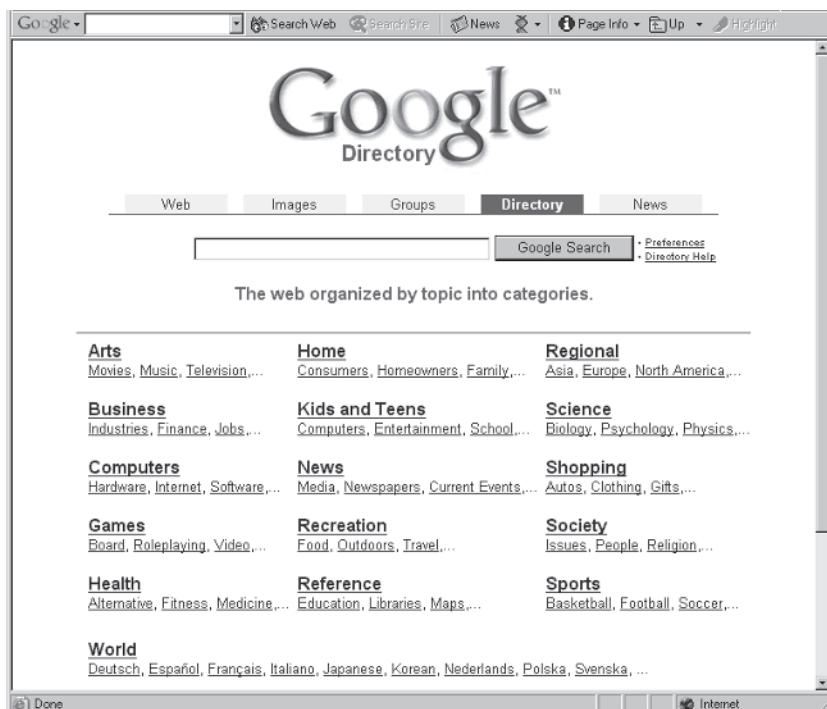


Figure 2.59: Google Directory

The content of Google's directory comes from the Open Directory Project (<http://dmoz.org/>). The Open Directory Project is a directory service staffed by volunteer editors from around the world. Open Directory differs from Yahoo in that Open Directory charges no fee to list a URL or site (yahoo!). Open Directory accepts any web site listing that meets its modest criteria (nothing illegal, slanderous, etc.).



You navigate through the directory by clicking on the colored links to go down the tree and using your browser button to go back up the tree. It is simple and easy to use.

## Google Web Search

Google Web Search is the Google home page located at [http://www.google.com](http://www.google.com/webhp?tab=dw&q=&ie=UTF-8&oe=UTF-8&hl=en&cat=gwd%2FTop). (Yes, bookmark it!) See Figure 2.60.



Figure 2.60: Google's web search page

Google Web Search is Google's graphical user interface that is powered by the Google web search technology behind the scenes. Google Web Search searches 3 billion web pages that Google spiders have indexed for your search word(s) and returns the number of matches it finds with links to the first 1,000 or so pages ranked in order of "link popularity" by PageRank. Recall from Chapter 1 that Google's search technology ranks pages according to "link popularity," and Google calls the software that performs that task PageRank.

To use Google Web Search to mine your nugget of information, you key your query words into the query box. Then left-click Google Search.



Google Search will return a list of web pages found by Google's search engine in a results page. The number of links that Google finds for any given search is dependent, of course, upon the number of web pages containing the search words and the number of those pages that the Google spiders have indexed. However, if you left-click on the I'm Feeling Lucky button, your browser is taken to the web site that Google has ranked as number 1. You bypass the results page and go directly to the numero uno web site. I hope you do not click this button too often, as this approach to finding information empowers Google with the opportunity to play games with everyone's searches.

To illustrate Google Web Search, I keyed in the word "Iraq" (without the quotes) and left-clicked on the Google Search button. The results of this search query are shown in Figure 2.61.

The screenshot shows a web browser window with the following details:

- Address Bar:** http://www.google.com/search?q=Iraq&num=100&l=en&rls=IE=UTF-8&oe=UTF-8&safe=active&output=search
- Search Bar:** Iraq
- Toolbar:** Includes buttons for Back, Forward, Stop, Refresh, Home, and a link to '63 blocked'.
- Google Logo:** Large Google logo on the left.
- Search Options:** Advanced Search, Preferences, Language Tools, Search Tips.
- Search Results:**
  - Web:** Categories: Society > Issues > Economic > Sanctions > on Iraq | Regional > Middle East > Iraq > Government
  - News:** London, Berlin, Paris seek joint Iraq stand - The Age - 32 minutes ago | Suit Claims Iraq Ties To Al Qaeda - CBS News - 38 minutes ago | Key European leaders to discuss Iraq - MSNBC - 1 hour ago
  - Try Google News: [Search news for Iraq](#) or [browse the latest headlines](#).
- Result Preview:** CIA - The World Factbook -- Iraq
- Description:** ... Introduction, Iraq, Top of Page, Background: Definition Field Listing Formerly part of the Ottoman Empire, Iraq became an independent kingdom in 1932. ... Description: Provides information about geography, people, government, and economy. Includes map. [Kids/Teens/Mature...]
- Category:** Kids and Teens > School Time > ... > World Cultures > Middle East
- Link:** www.cia.gov/cia/publications/factbook/geos/iz.html - 101k - Sep 15, 2003 - [Cached](#) - [Similar pages](#)
- Other Results:** Iraq.net Information Network | Towards a better future |
- Description:** Offers selected news and articles in English, links for other news sources and news coverage in Arabic....
- Category:** Regional > Middle East > Iraq > News and Media
- Link:** www.Iraq.net/ - 1k - [Cached](#) - [Similar pages](#)

Figure 2.61: Search results (top) page

This is about as simple as web searching can get. Just plug in the words, and select Google Search! The search results page, however, is not so simple. A lot of information is packed onto the page, and a searcher with low blood pressure knows what that information means to him or her. If you have search-induced blood pressure



issues, then read the following information carefully and make sure that you understand what is being discussed. Your path to search nirvana is here.

The top of the results page displays links to other Google features, as shown in Figure 2.61. Below the Google links is another search box that you can use to search the web again for other keywords; there is no need to return to the previous page to begin a new search. Below the search box there are a couple of framed lines of messages and tips. You are told what range of search results you are viewing. In this case, we are looking at results 1-100 of “about 21,300,000.”

We are told that the search performed with SafeSearch turned on took 0.69 seconds.

We can go to the Preferences page and set the number of pages to view and also turn SafeSearch off/on. We discuss setting Preferences in the next chapter. Below the message and tips bar, we are advised in what category we can find the search topic in Google’s directory. Below the category listing, we find links to recent news stories regarding the search topic “Iraq.” Below the News tag, we have the web pages identified by the search.



---

**Note:**

The search results include the titles, Google Web Directory categories, and domain names of the found web pages that have a match somewhere within the page to the keywords.

When we scroll to the bottom of the search results page, we find several interesting and useful things. There is a Result Page line that allows you to move to another set of search results. Since we are looking at 100 search results per Google results page, each number in the Result Page line represents 100 more pages in our search results. Just click on any number (2 through 10 in this case) to see the additional results. Keep in mind that PageRank has chosen these links to web pages out of the 21,300,000 total pages found and sorted them into the ranking shown. Left-clicking on



Next will select the next page displaying another rank-ordered 100 search result links.

Below the Result Page line is another search box. This search box does not initiate a new search of the entire web. Instead, it allows you to search *within the results already found*. This is a good way to narrow your search focus and quickly zero in on a web page containing the specific information that you are seeking. However, I find that it is not necessary to go this extra step after you gain a modicum of web searching savvy.

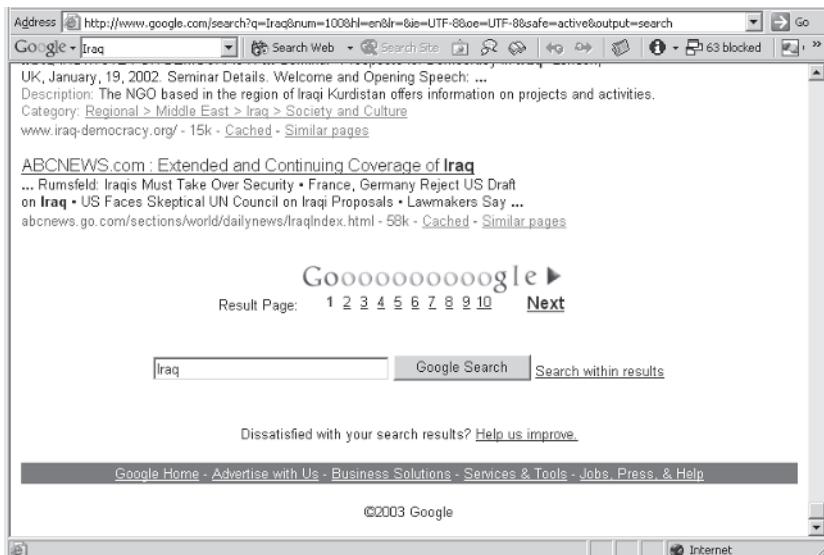


Figure 2.62: Search results (bottom) page

Below the search box are Google messages and links. If you have not previously installed the Google Toolbar, you will see a link (not shown in Figure 2.62) to the toolbar at the bottom of the page. The Google Toolbar is useful, and I suggest that you try it. Just click on the Get the Google Toolbar link and follow the instructions.

Unemployed? Or looking for another job? Click on the Jobs link in the bottom right-hand side of the window. This link is a Google resource and might land you your dream job at Google.



## Wireless

Wireless is a web page instructing people who connect to the Internet via their cell phones, Palm hand-holds and organizers, and the Handspring Organizer how to connect to Google. Wireless is found at <http://www.google.com/options/wireless.html>. (If your cell phone company charges as much as mine for wireless access, then don't bookmark it.)

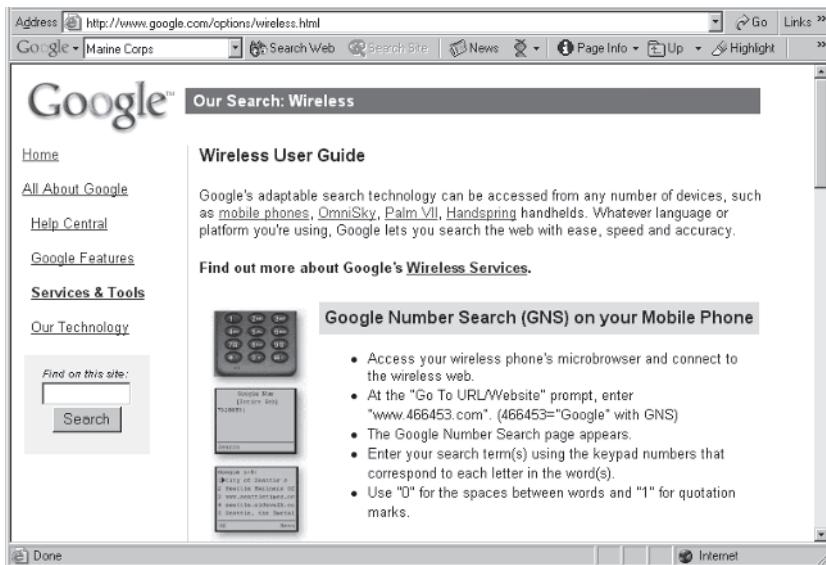


Figure 2.63: Wireless page

If you desire to connect to Google via any of the methods mentioned above, visit the Wireless site for instructions on how to connect.



## Summary

Google is feature rich, offering numerous products and services to the user. Many of these features, such as Froogle, Google Answers, Google Catalogs, and Google Labs, are not offered by competitors. While some of the elements of these products and services are incomplete, Google is sure to get the bugs worked out in the near future. In just a short period of time, Google has become the leader in offering web-based information services in addition to its search engine and directory services. Considering the web's youth, it is exciting to think what Google may bring to the information marketplace in the future.

This page intentionally left blank.

# Chapter 3

## Google Preferences

### Introduction

---

In the previous chapter, we discussed Google features in some depth. Most Google features are not closely coupled to Google's search engine, but rather they are stand-alone products. On the other hand, Google Preferences is closely coupled to the search engine. Preferences is very useful for organizing and displaying the search results. Preferences provides a means for us to enhance our search experience by customizing certain search features. So on to Preferences!

### Preferences

---

Google provides a Preferences web page where you can customize your search choices. Google's Preferences web page is found at <http://www.google.com/preferences?hl=en> or by left-clicking the Preferences link to the right of the search box on the Google home page. The choices on the Preferences page are:



- Interface Language
- Search Language
- SafeSearch Filtering
- Number of Results
- Results Window

The screenshot shows the Google Preferences page. At the top, there's a navigation bar with links for "Preferences Help" and "All About Google". Below the bar, a message says "Save your preferences when finished and [return to search](#)". There are "Save Preferences" and "Cancel" buttons.

**Global Preferences** (changes apply to all Google services)

**Interface Language**: Display Google tips and messages in:  English. If you do not find your native language in the pulldown above, you can help Google create it through our [Google in Your Language program](#).

**Search Language**:  Search for pages written in any language ([Recommended](#)).  Search only for pages written in these language(s):

<input type="checkbox"/> Arabic	<input type="checkbox"/> English	<input type="checkbox"/> Indonesian	<input type="checkbox"/> Romanian
<input type="checkbox"/> Bulgarian	<input type="checkbox"/> Estonian	<input type="checkbox"/> Italian	<input type="checkbox"/> Russian
<input type="checkbox"/> Catalan	<input type="checkbox"/> Finnish	<input type="checkbox"/> Japanese	<input type="checkbox"/> Serbian
<input type="checkbox"/> Chinese (Simplified)	<input type="checkbox"/> French	<input type="checkbox"/> Korean	<input type="checkbox"/> Slovak
<input type="checkbox"/> Chinese (Traditional)	<input type="checkbox"/> German	<input type="checkbox"/> Latvian	<input type="checkbox"/> Slovenian
<input type="checkbox"/> Croatian	<input type="checkbox"/> Greek	<input type="checkbox"/> Lithuanian	<input type="checkbox"/> Spanish
<input type="checkbox"/> Czech	<input type="checkbox"/> Hebrew	<input type="checkbox"/> Norwegian	<input type="checkbox"/> Swedish
<input type="checkbox"/> Danish	<input type="checkbox"/> Hungarian	<input type="checkbox"/> Polish	<input type="checkbox"/> Turkish
<input type="checkbox"/> Dutch	<input type="checkbox"/> Icelandic	<input type="checkbox"/> Portuguese	

**SafeSearch Filtering**: Google's SafeSearch blocks web pages containing explicit sexual content from appearing in search results.  
 Use strict filtering (Filter both explicit text and explicit images)  
 Use moderate filtering (Filter explicit images only - default behavior)  
 Do not filter my search results.

Figure 3.1: Preferences page

Interface Language gives you the choice of the language in which you prefer to view Google messages, tips, and text. This choice does not translate web pages.

Search Language lets you choose the language in which web pages are written. Note that this is not a domain-restricted search. In



other words, selecting German will result in a search of all web pages written in the German language, regardless of the physical location of the web page. So, a web page written in German that is located on a server in the United States will be a part of the search results, as will web pages written in German located on a server in Germany. A web page written in English and located on a German server will not be a part of the search results.

SafeSearch Filtering allows you to control the level of sexual content to which you are exposed in your search results. The three levels are no filtering, moderate filtering, and strict filtering.

Number of Results allows you to set the number of search results displayed per web page. The choices are 10, 20, 30, 50, and 100. The greater the number of search results per page, the less clicking you must do to see all of the results. I prefer to view 100 at a time. In that manner, I can quickly scroll through the approximately 1,000 search results that Google provides.



#### Note:

I do not like the 1,000-result limitation. Before Google, several search engines that I used allowed me to view every single search result. I liked that, as I sometimes found a very useful piece of information far beyond the 1,000 limit. That is, I did not look at thousands of search results (say every result from 1,000 to 10,000); I just hopped around in the hinterlands to see what I could find. I found interesting and entertaining information far beyond the 1,000-result limit.

The goal of search engines and this book is, of course, to reduce the search issue to a time-efficient solution so we do not spend hours (or even minutes) finding the answer to a question. But there is another popular aspect of searching that does not reduce to a quick fix for a problem (answers to questions). The popular aspect I am talking about is recreational searching.

We have many interests in life, such as hobbies and other forms of recreation. Many of us enjoy searching the web and reading about those interests. Individuals and small companies with home pages containing information of special interest can get blown over the



1,000-result wall, and you will never get the opportunity to read what they have to say. This 1,000-result wall is the same as a library telling you that you can only have five books about a subject. That's it. That's all you will ever get to read, even though there are hundreds of books written on the subject. So, your total exposure to the subject is confined to those five books, and the library chooses those five, not you. Does that sound kind of intimidating? It is exactly the same situation with search engines limiting what search results we can view. Search engines have become the portal to the information library, and they are now, in effect, telling us what we can and cannot see. They are also saying, in a manner, "Trust me." Scary, isn't it? Our search mantra should be: No limits!

Results Window selects a new browser window to display search results. This option can be useful if you perform multiple searches and do not want to lose the search result from one or more previous searches. However, there is a trade-off. Every time you open a new window, you are consuming machine resources. If you open too many windows, your computer is going to respond slowly to commands. How many is too many? Only you can judge that, as the number is based upon the configuration of your specific machine. If you have a fairly new machine and a whopping amount of physical memory, you may never notice any change in performance, regardless of the number of windows open. But if you are still utilizing an older machine, you may find that just a few windows bog down performance. Experiment and find out what your machine's limitations are.

When you change your preferences and left-click on Save Preferences, you naturally expect your browser to recall those exact preferences when you go back to the Preferences web page. However, some browsers may not. I use IE 6.0, and it does not actually save my preferences until I close the browser. When I return to the Preferences page, the old settings are displayed, not the new settings I chose in the last visit.

These Preferences options are discussed in great detail in the following subsections.



## Interface Language

Google can display its own tips and messages in any one of 89 different languages. To select a language other than whatever default is selected (English in most cases), go to <http://www.google.com/preferences?hl=en>. Left-click on the Interface Language drop-down menu and select the language you prefer. Be sure to select Save Preferences, or the next time you open Google the default language will be used.

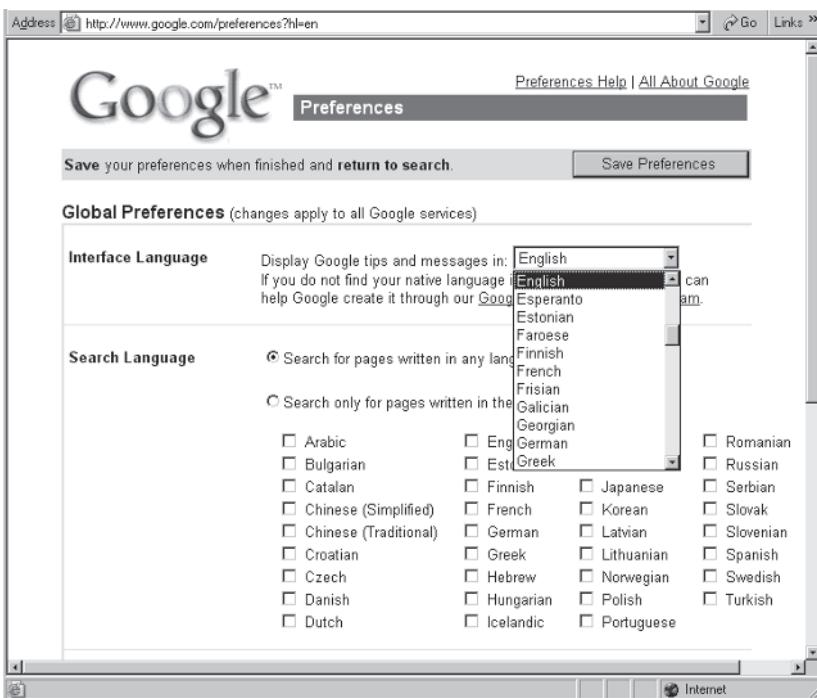


Figure 3.2: Selecting the interface language



## Search Language

You can search Google for web pages written in any one of 35 different languages.

This is not a search restricted to any domain. In other words, if you selected Portuguese and then performed a basic search, Google would search every web page that it has indexed written in Portuguese, regardless of the physical location of the web page. To perform a search for web pages in a language other than English, go to the Preferences page at <http://www.google.com/preferences?hl=en> or select the Preferences link on the Google home page. See Figure 3.3. Note the default Search Language setting is Search for pages written in any language. Now select the language of interest. If you want to search more than once, select Save Preferences.

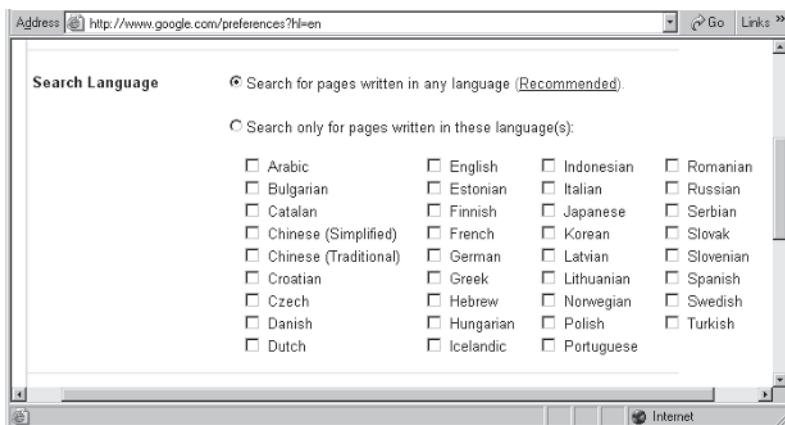


Figure 3.3: Selecting the search language

For an example, I chose to select German and search for the word “Berlin.” My search results are shown in Figure 3.4. Notice the search results are pages written in German, but Google tips and messages are still displayed in English. While viewing the Preferences page, we forgot to select German in the Interface Language section. So, let’s go back and select German as the interface language and perform the search again. See Figure 3.5. Again, select Save Preferences if this is a choice you want to use more than once.



We perform the search again, and now we have Google tips, messages, and web pages all in the same language, German. See Figure 3.6.

Address: http://www.google.com/search?num=20&hl=en&ie=UTF-8&oe=UTF-8&q=de+Berlin&tn=Google+Search&lr=lang\_de

**Google™** Advanced Search Preferences Language Tools Search Tips

Search the Web Search German pages

Web Images Groups Directory News

Searched **German pages for Berlin**. Results 1 - 20 of about 389,000. Search took 0.49 seconds.

Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Category: World > Deutsch > Regional > Europa > Deutschland > Berlin

Berlin - Offizielle Seite der Hauptstadt Deutschlands - [Translate this page]  
Offizielle Informationen des Landes Berlin, der Landesregierung und nachgeordneter Bezirke und Behörden sowie interessante Informationen für Bürger und ...  
Description: Offizielle Informationen des Landes Berlin, der Landesregierung und nachgeordneter Bezirke und Behörden...  
Category: World > Deutsch > Regional > Europa > Deutschland > Berlin  
www.berlin.de/-32k - Cached - Similar pages

TU Berlin - Willkommen! - [Translate this page]  
TU Berlin Straße des 17. Juni 135 10623 Berlin Tel.: +49-30-314-0, TU Berlin, ... Suchen  
Suchmaschine, Index AZ, Adressen & Telefon, Alumni, mailbox.tu.berlin.de, ...  
Description: Website mit Infos über Organisation, Fakultäten, Forschungsvorhaben, internationale Beziehungen,...  
Category: World > Deutsch > ... > Berlin > Technische Universität Berlin  
www.tu-berlin.de/-16k - Cached - Similar pages

Sponsored Links

**Berlin Hotels**  
Large collection, local broker, get 30.65% discount!  
www.euroStays.com  
Interest:

**Accommodation In Berlin**  
Discount Accommodation in Berlin  
Compare prices and Book Online.  
www.247-shopping.co.uk/berlin  
Interest:

See your message here...

Figure 3.4: Google search of German pages only

Address: http://www.google.com/preferences?hl=en

**Google™ Preferences**

Preferences Help | All About Google

Save your preferences when finished and **return to search**.

**Save Preferences**

**Global Preferences** (changes apply to all Google services)

**Interface Language** Display Google tips and messages in: **German**  
If you do not find your native language help Google create it through our [Good](#) [can](#) [am](#)

**Search Language** Search for pages written in any language: **German**  
Search only for pages written in the: **Cyrillic**  
 Arabic     English     Japanese  
 Bulgarian     Hebrew     Korean  
 Catalan     Finnish     Latvian  
 Chinese (Simplified)     French     Slovenian  
 Chinese (Traditional)     German     Romanian  
Interest:

Figure 3.5: Changing the interface language to German



Address: http://www.google.com/search?q=berlin&num=20&hl=de&tz=0&utf8=UTF-8&oe=UTF-8&safe=off&output=search

Erweiterte Suche Einstellungen Sprachtools Suchtipps

berlin Google Suche

Web-Suche Suche Seiten in Deutsch

Web Bilder Groups Verzeichnis News Neu! Seiten auf Deutsch wurden nach berlin durchsucht. Ergebnisse 1 - 20 von ungefähr 421,000. Suchdauer: 0.09 Sekunden.

Kategorie: World > Deutsch > Regional > Europa > Deutschland > Berlin

News Berlin führend in junger Kunst - Berliner Morgenpost - vor 5 Stunden gefunden  
Vorschlag für Doppelflughafen: Senat lehnt Achse Berlin-Leipzig ab - Die Welt - vor 14 Stunden gefunden  
Lesen Sie Google News: Nachrichten nach berlin durchsuchen oder in den neuesten Schlagzeilen blättern

Anzeigen

**Berlin Hotels**  
Large selection, local booking, get 30-65% discount!  
www.euroStays.com  
Interesse: —

**Accommodation In Berlin**  
Discount Accommodation Berlin  
Compare prices and book online.  
www.247-shopping.co.uk/berlin  
Interesse: —

Sehen Sie Ihre Anzeige hier...

TU Berlin - Willkommen!  
TU Berlin Straße des 17. Juni 135 10623 Berlin Tel. +49-30-314-0, TU Berlin, ... Suchen  
Suchmaschine, Index AZ, Adressen & Telefon, Alumni, mailbox.tu-berlin.de, ...  
Beschreibung: Website mit Infos über Organisation, Fakultäten, Forschungsvorhaben, internationale Beziehungen, ...  
Kategorie: World > Deutsch > ... > Berlin > Technische Universität Berlin  
www.tu-berlin.de/ - 16k - Im Cache - Ähnliche Seiten

Figure 3.6: Google search of German pages only with Google page in German

## SafeSearch Filtering

When searching the Internet, you are at risk of exposure to pornographic material, including web sites, images, and other explicit sexual content. Perhaps your sensibilities are such that you would prefer not to see such material. SafeSearch is a software filter intended to keep pornographic web sites, images, and explicit sexual content from appearing in your search results.

There are three levels of SafeSearch filtering:

- **Do not filter** — This is self-explanatory; what you see is what you get!
- **Moderate filtering** — Excludes most explicit images from Image Search results.
- **Strict filtering** — Applies SafeSearch filtering to all search results (web pages and images).

Depending upon the level of filtering you have selected, all, some, or none of the filtered sites will appear in your search results.



Google makes the disclaimer that “while no filter is 100% accurate Google’s filter uses advanced proprietary technology that checks keywords and phrases, URLs and Open Directory categories.”

Google’s warning concerning what you may see when viewing images is also applicable to the results you may get when searching for any material on the web.

**Warning:** The results you see with this feature may contain mature content. Google considers a number of factors when determining whether an image is relevant to your search request. Because these methods are not entirely foolproof, it’s possible some inappropriate pictures may be included among the images you see. (The mature content filter is only available from an English interface.)

This disclaimer is necessary for several reasons. You should be aware that there is an ongoing struggle between search engine designers and web-based pornographers, one trying to keep adult material from appearing in search results and the other trying to place their material in every search possible. So, regardless of Google’s best intentions, you will (if you do any amount of searching) come across such material from time to time. Please be sympathetic to the Herculean task of keeping such objectionable material out of search results.

The three levels of filtering are strict, moderate, and none. Strict filtering will filter both images and text-based material. Moderate filtering only filters images, while none does not filter text or images. You can adjust SafeSearch settings from the Preferences page, the Advanced Search page, and the Advanced Image Search page. You can only save your choices on the Preferences page. You set your choices on the other two pages on a search-by-search basis. The default is no filtering.

I found that with the particular version of Internet Explorer I use, I had to close all Internet Explorer windows and then open a new browser window for my preferences to be activated and SafeSearch to work.



Let's try a search using the word "sex" as the test word. We enter the test word in the search query box with SafeSearch off and hit the Search button. The search returns "about 8,780,000" results. See Figure 3.7. Now, let's do it again with SafeSearch on. The results are totally unexpected — 8,500,000 hits. See Figure 3.8. Hmm...one would think a filtering program with "sex" as the test word would return very few search results.

Web Images Groups Directory News  
Searched English pages for **sex**. Results 1 - 100 of about 8,780,000. Search took 0.23 seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 3.7: Search results without SafeSearch

Web Images Groups Directory News  
Searched English pages for **sex** with Safesearch on. Results 1 - 100 of about 8,500,000. Search took 0.18 seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 3.8: Search results with SafeSearch

Okay, every good search engine can have a "bad hair day." I performed the same experiment the following day using "sex" as the test word, and the results stunned even me. See Figure 3.9.

Web Images Groups Directory News  
Searched the web for **sex**. Results 1 - 100 of about 203,000,000. Search took 0.27 seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 3.9: Search results without SafeSearch

I received an incredible 203,000,000 hits! Sex must be a popular topic on the Internet, eh? It's not just us men talking about it, either! I checked a random sample of sites in the first 1,000 results and discovered many sites were pornographic, but many sites were also directed toward women's sex-related issues.

Okay, back to the issue at hand. We received an amazing 195 million more hits by waiting a day between searches. What is going on here? Did 195 million web pages get added to the Internet in one day? I don't think so! Then what is the explanation for the 250 percent increase in web sites? The first search that I performed yielded 8,780,000 results and was allotted 0.23 seconds, while the second search was allotted 0.27 seconds. The first set of searches



(with and without SafeSearch on) was performed on a Friday morning. The second set of searches was performed on a Saturday morning. This means the low-level software search routines working without benefit of union representation or compensation for your exclusive benefit can identify more web pages on weekends than weekdays. The point is, Google traffic is lighter on a Saturday morning than it is on a Friday morning, allowing the low-level routines more machine cycles per search to find keyword matches.



Figure 3.10: Search results with SafeSearch and strict filtering

Notice that even with strict filtering set in the Preferences page, a search on the word “sex” still returns “about” 197 million results. Want to guess how many of those sites contain sexually explicit material? Me neither.

If we have 197,000,000 web pages as a result of searching on “sex,” what material will SafeSearch filter? SafeSearch does block the obvious four-letter-type word searches. I will not list the words here, but I imagine you know which ones I mean. When you key in a filtered four-letter (or five or six, etc.) word, you will receive the message, “No standard web pages containing all your search terms were found.”

Unless we have children around, it is not this type of filtering that we are really interested in. After all, most of us control our own behavior, and if we want, we can avoid keying in four-letter words. We are more interested in blocking web pages containing offensive material when we search using what we believe are innocent terms. SafeSearch attempts to filter web pages in this type of search, but Google emphasizes on its web site that SafeSearch filtering is not 100 percent foolproof.



## Number of Results

The Number of Results drop-down menu allows you to select the number of results that you want displayed on each search results page. As shown in Figure 3.11 you can choose 10, 20, 30, 50, or 100. I cannot think of any good reason to choose a number less than 100.

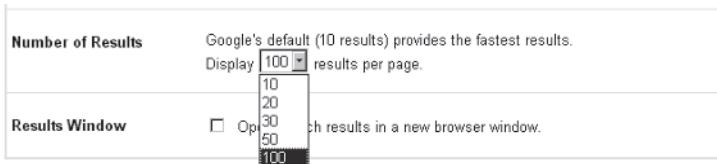


Figure 3.11: Number of Results option

When you choose less than 100, you just have to click more often to get the remainder of the results. Save yourself some clicking and go for the max!

## Results Window

The Results Window option, shown in Figure 3.12, lets you choose whether the search results will open in a new browser window. This is convenient if you are doing multiple searches or if you just like to open everything in a new window so you do not have to use the browser's Back button.



Figure 3.12: Results Window option

If you want to open a new browser window every time you search, left-click in the box adjacent to Open search results in a new browser window. Left-click on the box again to clear your choice.



## Save Preferences

Be sure to left-click on the Save Preferences button before you exit the Preferences window. See Figure 3.13.

Save your preferences when finished and **return to search**. Save Preferences

Figure 3.13: Save Preferences button



### Note:

Cookies must be enabled in your browser in order to save your preferences. When you click on Save Preferences, you will get a message that says your preferences have been saved. See Figure 3.14. Left-click on OK to move on to your search. After you click on OK, Google will display its home page where you can begin your search.



Figure 3.14: Confirming that preferences have been saved

On my computer and my version of Internet Explorer (IE), I had to close IE and then reopen it to get all of the preferences to change to the new choices.

## Summary

Google Preferences includes several search options that allow you to customize your search experience. Visit Preferences early in your search and set the options to suit your sensibilities and search needs. Be sure to save your selections.

This page intentionally left blank.

# Chapter 4

## Basic Google Search Techniques

### Introduction

---

Before we could really begin our discussion on searching, we needed to understand what search services Google provides. Now that we have reviewed Google Services in Chapter 2 and we know how to set our preferences, we are ready to do a little serious searching. Recall from Chapter 1 that 29 percent of us are “very frustrated” when we search the web for our nugget of information. Also, 77 percent of us experience some degree of frustration. The reason so many of us are frustrated is that we do not use the wide variety of tools available to assist us in our searches. We just perform a very basic search and then sit there staring at the hundreds of thousands of search results we get and ponder which one of those innumerable URLs has the information we seek. We are now going to study the tools that will remove searching the Internet from the realm of Colchis (frustration) to the land of Iolcus (nirvana).



Let's examine the Google home page for a moment so that we are on the same page (literally speaking). The Google home page is shown in Figure 4.1.



Figure 4.1: Google home page

The query box, also called the search box, is shown in Figure 4.2. This is the box where you enter the keywords. You may enter as many words as you like, up to a maximum of 256 characters.



Figure 4.2: Google query box

After you key in the search words, you left-click on the **Google Search** button. After a moment, the search results page will be displayed. The search results page contains hyperlinks to web pages containing the search criteria that you specified.

Alternately, if you want to view only the first search result, you can left-click on **I'm Feeling Lucky** to go to that web page.

When you are buried in the midst of search results and decide that you want to go back to the Google home page, just left-click on the Google icon:





and you will find yourself back home! Now on to the good stuff.

There are three fundamental types of searches that one may perform when searching the Internet or the web, or some portion of either. For simplicity's sake, from now on I refer to searching the Internet, the web, or some portion of either simply as "searching the web," unless a specific reference is required to understand the context of the reference. There is the simple search, the basic search, and the advanced search. When searching the web, we typically key the text that we are searching for in the search box and then hit the Google Search button. This type of search is referred to as a simple keyword search, regardless of how much text is keyed into the query box. A simple keyword search (from now on called a simple search) typically yields thousands, if not millions, of search results. Few people have time to review the first 200 search results, so it is not likely that anyone is going to take on 10,000 search results. However, note that Google provides access to the first 1,000 search results only.

The next type of search is called the basic search. A basic search employs a few basic operators within the text string with the objective of including wanted, or excluding unwanted, material. The advanced search has the same objectives as the basic search, namely to include or exclude specific material, but it differs from the basic search in the quantity, range, and complexity of its operators. We discuss advanced searches in the next chapter.

None of the search operators, tools, and services used in a Google search are complicated in terms of your ability to understand how to employ them. In fact, they are very comprehensible. They are only complex in terms of how Google's search engine uses them to identify the results that you are searching for. Fortunately for us, we do not have to have a software developer's understanding of search engines to effectively use the operators and tools available to us. We begin delving into the search operators and tools by examining a simple search.



Figure 4.3 illustrates a simple search on the keywords “cheap air fare.” This search will return around 246,000 results. See Figure 4.4.



Figure 4.3: A simple search



Figure 4.4: Results of a simple search

Obviously, simple searching is not the way to information nirvana. In fact, simple searching is the cause of most web search frustration.

From this point forward, we are going to work on developing additional search skills that will reduce or even totally eliminate your search frustration by increasing the accuracy of your searches. We accomplish this by employing *basic search operators*, also called *search modifiers*, in the search criteria. You can think of basic search operators as filters that prevent unwanted URLs from appearing in your search results, giving you better quality search results than ever before. If you employ even a few of the basic search operators in your everyday searches, you will soon find that you seldom, if ever, experience search rage again.

There are several classes of basic search operators. Boolean and arithmetic search operators are similar, while text operators are in a unique class of their own. Boolean and arithmetic operators employ Boolean or their equivalent mathematical expressions in the search criteria. Depending upon the application (Google or some other search engine), Boolean or arithmetic operators may be employed but not both.

When I mention quality from this point forward, I am not talking about the graphical presentation of the material but rather the quality of the information related to the search query. In other words,



the quality of the search results has nothing to do with how pleasing the material may be presented to you in its graphical format but how well the information contained in the web page relates to the text you searched on. A high-quality page will contain more than one reference to the search text, and the search text will be the subject matter of the page. The query words you searched are the reason the web page exists. Google does play around with its rank listings some in order to favor certain organizations or types of organizations, so search results are not set in concrete.

## Keywords

We need to discuss searching the web for a moment before we examine search operators. One simple act can reduce the frustration that most people feel when searching the web to a fairly manageable level. It is so simple that you may wonder why someone must write a book to advise people about it. However, the vast majority of people have little or no experience searching for information, especially searching for information on the web. That is why 77 percent of us are frustrated when we search! But the simple act of stopping for a moment and thinking about what it is you are trying to find when you have a search need will save you considerable time and frustration.

An example of a search need would be to assist a child with a history homework assignment. I will use an example from my own experience. I receive numerous questions from moms, dads, and schoolchildren regarding the American Civil War, as a result of owning an American Civil War web site. A typical question is, “Where and when was the first Civil War battle fought?”

This question when posed to a human being is understandable, and a knowledgeable individual could easily answer the question. However, such a question posed to a search engine would result in many thousands of erroneous search results. This is because of the way search engines parse (connect) text strings. When a search engine



sees the text string “When was the first American Civil War battle fought” with the quotes, it will return only those pages, not with the answer but with a list of pages that posed the same question! Repeat the search without the quotes, and every web page that included the terms “first,” “American,” “civil,” “war,” “battle,” and “fought” anywhere within the contents of the page, including the title and page body, would be identified as a valid search result.

When we examine the question, we find several terms that are extremely important to the nature of the question. The terms “American,” “civil,” “war,” “first,” “battle,” and “fought” are known as *keywords*. The verb (“was”), adverb (“when”), article (“the”), and conjunction (“and”) in the question are superfluous. By removing verbs, adverbs, articles, conjunctions, gerunds, and pronouns, you focus on your true search objectives with a resulting increase in the quality of your search results. From this discussion, you may determine that keywords tend to be nouns and adjectives. The ability to zero in on important keywords is critical to successfully reaching search nirvana.

Google ignores some common words, single characters, and single digits. Obviously, ignoring adverbs, articles, and conjunctions is almost a necessity. If Google has ignored an entry in the search text, it indicates so by displaying the appropriate details on the results page below the search box as in Figure 4.5. The informative message, “‘where’ is a very common word and was not included in your search,” is a typical message whenever a word is ignored.

Advanced Search Preferences Language Tools Search Tips  
Abraham Lincoln where civil war Google Search

"where" is a very common word and was not included in your search.  
[details]

Figure 4.5: Ignoring a common word

Leaving out common words in your search text results in a faster search but not necessarily a higher-quality search result. In any case, common words are not keywords, and keywords are the key to successful searching. Now we are armed with enough



information about simple and basic searches to move on to our first and simplest basic operator.

## The Quote ("") Operator

Perhaps the simplest basic search operator is the quote (""). It is a text operator. By placing quotes around your text, you “combine” the words into a phrase. Searches on phrases return only those URLs that contain that *exact* phrase either in the title, the URL, or the body of the page.

In the following basic search example, the words “civil war” without the quotes will return over three million search results.

Civil War  
Google Search I'm Feeling Lucky

Figure 4.6: A search on Civil War

Web Images | Groups | Directory | News |  
Searched the web for Civil War Results 1 - 10 of about 3,480,000. Search took 0.20 seconds.

Figure 4.7: Results of a search on Civil War

There are too many results because every web page in the world with the word “civil” or “war” is returned in the results. But, we are really interested in the American Civil War. So, let’s search again using the quotes around the search text. See Figure 4.8.

"civil war" Google Search

Figure 4.8: A “civil war” search

The number of search results that we get with “civil war” (with the quotes) is 1,950,000. I think that number of search results is still too many for me to review in a day or two. But we have eliminated about 1.5 million search results by applying a simple “basic” search operator. We are making progress.

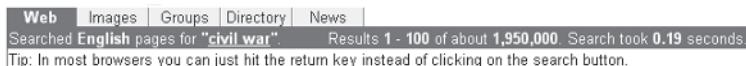


Figure 4.9: "Civil war" search results

Google may find a large number of web sites, but you can only access approximately the first 685 pages found unless you click on Repeat the search with the omitted results included. This link is only available when you go to the last result page — page 9 in this instance. See Figure 4.10. After you click on the link shown in Figure 4.10, you will have access to the first 985 or so results. My experiments show that the number of pages you can view varies from search to search, ranging anywhere from 750 to 1,000 maximum.

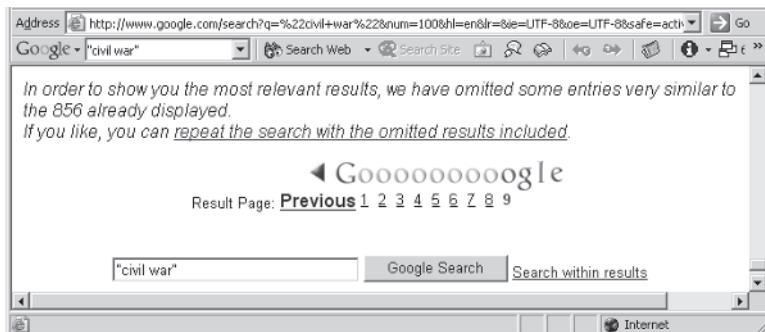


Figure 4.10: "Civil war" search results

Note that a search of "civil war" on Yahoo returns only 571 viewable results. If Yahoo is "powered" by Google, then why the differing search results? I cannot answer that question at the moment.

Now, let's think about our search for a moment. What Civil War are we interested in? The French Civil War (revolution)? The Russian Civil War? How about the American Civil War? The search results that we have obtained so far will include every web page that contains the phrase "civil war." But we are only interested in the American Civil War, so let's search on that phrase. See Figure 4.11.

Figure 4.11: An “American Civil War” search

The results that we get when we search on “American Civil War” are shown in Figure 4.12. The 243,000 results represent only those URLs that contain the phrase “American Civil War” somewhere within the web page. The use of capital letters in this search query is for instructional purposes only. The same results are obtained regardless of capitalization since Google is not sensitive to capitalization.

The results we obtain for an “American Civil War” search is much more manageable than just a search on “civil war.” But 243,000 search results are still too many to view.

[Web](#) [Images](#) [Groups](#) [Directory](#) [News](#) |  
Searched English pages for “**American Civil War**”. Results 1 - 100 of about 243,000. Search took 1.67 seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 4.12: “American Civil War” search results

If 243,000 search results are still too many for you to peruse, other advanced search operators may be combined with the quote operator to further reduce the number of search results.

Perhaps this is a good place to interject a side note about searching the web. To reduce the number of search results as much as possible, it is advisable for you to consider what it is that you are trying to find. “American civil war” is a very broad topic. We can expect to get many thousands of web pages containing that phrase, as it is a very popular subject. But what is it about the American Civil War that we really want to know? The first battle? The last battle?

When it began? The questions are innumerable, but the point is made. When you search, try to determine as specifically as possible what it is you are trying to find. The narrower the search topic, the fewer results you will obtain, and the quality of the search results will increase proportionally.



Remember, the quotes combine everything inside into a single phrase, and that is the text string, or phrase, searched for. Using quotes is the simplest basic operator that you can use to eliminate unwanted URLs from your search results. Try it. Search on my name, Michael Busby, first without quotes and then again using the quotes. Now, perform the same exercise on your name. You may be surprised to discover that you are mentioned in a web page somewhere.

## The Arithmetic (+, -) Operators

---

A little more advanced than the quote operator, arithmetic operators include or exclude text and phrases in the search results. The arithmetic operators are the familiar “plus” symbol (+) and the “minus” symbol (-). The plus symbol instructs the search engine to include in the results all web pages that *include* the text to the right of the plus symbol in addition to the text on the left side of the symbol.

The minus symbol instructs the search engine to return in the results all web pages that *exclude* the text to the right of the minus symbol and *include* the text on the left side of the symbol. Some search engines require there be no space between arithmetic operators and their operands (search words). Google does not care if you place a space between the plus operators and the operands. For example, “flower + pot” and “flower+pot” will be viewed by Google as the same search. However, the minus operator works a little differently. To exclude a query word from a search, you must put a space between the wanted query word and the minus operator, then have no space between the minus operator and the excluded query word. So, to exclude “Lincoln” from all “Abraham” searches, we would search on “Abraham -Lincoln” — without the quotes.

Let’s search using the plus symbol and see how it operates on the results. Let’s search on “American Civil War” + “Abraham



Lincoln.” First we search without any operators (the quotes and the plus symbol) and then compare that result with the search results using the operators. See Figure 4.13.

**Note:**

It does not make any difference if the characters in the search string/text are uppercase or lowercase. The search engine converts all characters to the same case. So “American Civil War” will yield the same search results as “american civil war.”

American Civil War Abraham Lincoln

Figure 4.13: A search on American Civil War Abraham Lincoln

Searching on the text “American Civil War Abraham Lincoln” (without the quotes) returns 161,000 search results. See Figure 4.14.

Web | Images | Groups | Directory | News |  
Searched English pages for **American Civil War Abraham Lincoln**. Results 1 - 100 of about 161,000 Search took 0.29 seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 4.14: Results of a search on American Civil War Abraham Lincoln

Now, we repeat the same search and use the two operators that we have discussed so far, the quotes operator and the plus operator. See Figure 4.15.

“American Civil War” + “Abraham Lin

Figure 4.15: “American Civil War” + “Abraham Lincoln” search text

The search on “American Civil War” + “Abraham Lincoln” yielded 16,200 results. Now, you may wonder what real difference it makes to have nearly 3,500,000 results and 16,200 results when all you have the patience to review is about 200 (if that many)! Well, the fact is, the quality of the first 200 web pages of the 16,200 search results will be much better than the quality of the first 200 web pages of the 3,500,000 results.



[Web](#) [Images](#) [Groups](#) [Directory](#) [News](#) |  
Searched English pages for "[American Civil War](#)" + "[Abraham Lincoln](#)". Results 1 - 100 of about 16,200. Search took 0.29 seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 4.16: "American Civil War" + "Abraham Lincoln" search results

Google makes searching a little simpler by automatically assuming a "+" when search text is keyed into the search box. That is, Google will automatically put the "+" between the words for you. Google interprets "American Civil War" + "Abraham Lincoln" as "American + Civil + War" + "Abraham + Lincoln." So, "American Civil War" + "Abraham Lincoln" will give you exactly the same results as "American Civil War" "Abraham Lincoln" and "American + Civil + War" + "Abraham + Lincoln."

Other search engines may or may not operate on the "+" symbol in exactly the same way. If you use another search engine, play around with and without the "+" and see whether or not you get the same results.

While the plus symbol (+) is inclusive, the minus symbol (-) is exclusive. That is, you can use the minus symbol to exclude unwanted web pages from your search results. The Google search engine also implements the Boolean operator NOT to exclude unwanted terms. We cover the NOT operator in the next section. The "-" symbol is included here for your information. Other search engines may or may not implement this. If you search using another engine, try the minus symbol to see if it works.

Let's repeat the same search and use the quotes operator with the minus operator. See Figure 4.17. Note the minus operator has a space between itself and the left keyword/phrase but has no space between itself and the right keyword/phrase.

Figure 4.17: "American Civil War" – "Abraham Lincoln" search text

The search on "American Civil War" – "Abraham Lincoln" yielded 385,000 results. Contrast that to the search on just "Civil War" that returned nearly 3,500,000 results.



Web | Images | Groups | Directory | News |  
Searched the web for "American Civil War" - "Abraham Lincoln" with Safesearch on. Results 1 - 100 of about 385,000. Search took 0.32 seconds.

Figure 4.18: "American Civil War" –"Abraham Lincoln" search results

The search results using the minus operator include all of the web pages containing the text "American Civil War" and exclude any web page that contain the text "Abraham Lincoln." Obviously, a search using this helpful operator will reduce the number of search results tremendously.

The number of results you get when you duplicate the search examples in this book are dependent upon many factors. You may or may not get the same number of results shown in these figures. Notice the "Search took 0.xx seconds" text in the results page information bar. The time allotted per search is dependent upon the amount of traffic Google is experiencing moment to moment. During slower periods of the day or night, more time is allotted per search. Usually, the longer the time allowed to search, the greater the number of search results.

Google ignores some common words, characters, and single digits. If you want to include such text in your search, use either the double quotes for exact phrase matches or the plus operator. A (silly) example is "where are you" or "+ where + are + you." Note the leading "+" before "where." Do not confuse adding the plus (+) symbol to search on common words with Google's default search that assumes a plus symbol between words. *Without the physical presence of the plus symbol, the common words will be ignored.* If you have any doubt, include the plus symbol.

From the above examples, you can clearly ascertain the benefits of incorporating the three basic operators in your search queries. If you have not used them in the past, I strongly encourage you to employ them at every opportunity. I can see that your blood pressure is going down now.



## Boolean Operators

---

The Boolean operators are the three words AND, OR, and NOT. These three words are called Boolean operators because they have a very specific meaning in a field of math called Boolean algebra. No, we will not discuss any algebra in this section — thank God! However, a little web history is in order to have an appreciation for the origins of these three words. Behind the web, underpinning the foundation of both the hardware and software comprising the web, is the Boolean field of math. So, what does that have to do with searching for our piece of information? AND, OR, and NOT are three basic Boolean operations, so it is relatively easy to incorporate these three operators into the search engine's bag of tricks. Note that the Boolean operators are capitalized.

The OR operator operates to find web pages containing any of the search terms that are OR'ed together. That is, "American OR Civil OR War" (without the quotes) will return every web page with either "American" or "Civil" or "War" anywhere within the web page. Google recognizes the "|" symbol when substituted for the term "OR." That is, "American|Civil|War" is the same as "American OR Civil OR War." While this book was in production, Google introduced the tilde (~) symbol to be used instead of "OR" or "|". Try all three and see what you get. My quick check does indicate there are some differences. The tilde operator seems to be more of a wildcard than the "|" symbol.

The AND operator works very similarly to the plus symbol. It is inclusive. Using the AND operator includes the search text on both sides of the operator. The NOT operator is exclusive, as it excludes the text listed on the right side of the operator. Therefore, the NOT operator is similar to the minus operator. But the two Boolean operators are not the same as the two mathematical operators. The following examples illustrate that there are differences between the two types of operators. I use the same search text as used to illustrate the mathematical operators, except the Boolean operators are used in lieu of the mathematical operators.

**Note:**

Google automatically assumes the AND operator between keywords. That is, a search on “cheap travel London” will return the same search results as “cheap AND travel AND London.” I include the AND operator in these examples for instructional purposes. Also, note that Boolean operators are applicable to all search engines and search situations, including searches conducted within a site. In that case, the AND may or may not be assumed.

In Figures 4.19 and 4.20, the Boolean operator AND is used in lieu of the mathematical plus operator. Note that the search using the Boolean operator returns 16,300 search results, whereas the same search using the mathematical plus operator returns 16,200 results.

"American Civil War" AND "Abraham Lincoln"

Google Search    I'm Feeling Lucky

• Advanced Search  
• Preferences  
• Language Tools

Figure 4.19: “American Civil War” AND “Abraham Lincoln” search text

Web Images Groups Directory News

Searched English pages for **“American Civil War” AND “Abraham Lincoln”**. Results 1 - 100 of about 16,300. Search took 0.28 seconds.

Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 4.20: “American Civil War” AND “Abraham Lincoln” search results

In Figure 4.21 and 4.22, the Boolean operator NOT is used. Note that the search using the NOT operator returns 10,900 search results.

"American Civil War" NOT "Abraham Lincoln"

Google Search

Figure 4.21: “American Civil War” NOT “Abraham Lincoln” search text

Web Images Groups Directory News

Searched English pages for **“American Civil War” NOT “Abraham Lincoln”**. Results 1 - 100 of about 10,900. Search took 0.24 seconds.

Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 4.22: “American Civil War” NOT “Abraham Lincoln” search results

**Table 4.1: Comparison of Boolean and arithmetic operators**

Operator	Search Results
AND	16,300
+	16,200
NOT	10,900
-	385,000

One would expect the difference in the search results to be the same for both Boolean operators. That is, if one Boolean operator yielded fewer results than its mathematical counterpart, then you might expect the other to yield fewer results also. That seems a reasonable conclusion given the nature of mathematics. However, due to nuances of search engine algorithms, you cannot expect such results. So, what is the real difference between mathematical and Boolean operators in terms of search results quality? Try both counterparts (AND and +, NOT and -) and compare the results.

To demonstrate the power of Boolean operator searching, I used a history-based search example. The number of us who search the web looking for answers to our children's homework assignments is unknown, but given the number of children in this country and the propensity of teachers to assign homework, we must be numerous. You can employ the power of Boolean searches in virtually any web search that you perform. Travel is the number one search query. Using Boolean operators to include or exclude travel keywords will help eliminate unwanted results. Of course, you can apply the same technique to any of your search needs.

## Complex Boolean Operations

---

You can use the basic Boolean operators (AND, OR, and NOT) to form more complicated search expressions. By combining the Boolean operators, you can exercise more control over the search algorithm, thereby reducing the number of hits while receiving



better quality results. The use of complex Boolean operations may appear daunting, but a little practice will yield terrific results.

To get the best possible results, think about your search goals and plan your search around those goals. To illustrate the approach to complex Boolean searching, let's use an example. Suppose we are interested in purchasing a Hewlett-Packard (HP) desktop or laptop computer. At this point, we do not care if the computer is new or used, a laptop or desktop. So, we have several keywords that we can search on: desktop, laptop, computer, and HP. The arrangement of the keywords in the Boolean operations is important. Some arrangements of keywords might be nonsensical, so make sure you arrange them in sensible patterns. The complex operations are shown in the leftmost column of the following table. The second column illustrates the Boolean operation performed by the Google search algorithm, the third column illustrates how you enter the text in the Google search box, and the rightmost column illustrates the possible combinations of keywords that will be present in the search results.

In Table 4.2, we use the characters a, b, c, and d to represent the following keywords:

- a = desktop
- b = laptop
- c = computer
- d = HP

**Table 4.2: Complex Boolean operations**

Boolean Operation	Equivalent Expression	Key This in Search Text Box	Search Results
a AND b	a b	desktop laptop	Every web page that contains both “desktop” and “laptop.” The possible combinations are: desktop laptop
a OR b	a OR b	desktop OR laptop	Every web page that contains either “laptop” or “desktop.” The possible combinations are: desktop laptop



Boolean Operation	Equivalent Expression	Key This in Search Text Box	Search Results
c AND (a OR b)	a c OR b c	(computer desktop) OR (computer laptop)	Every web page that contains either “computer” and “desktop” or “computer” and “laptop.” The possible combinations are: computer desktop computer laptop
(a AND b) OR (c AND d)	a b OR c d	(desktop laptop) OR (computer HP)	Every web page that contains either “desktop” and “laptop” or “computer” and “HP.” The possible combinations are: desktop laptop computer HP
(a OR b) AND (c OR d)	a OR b c OR d	(desktop OR laptop) AND (computer OR HP)	Every web page that contains “desktop” or “laptop” and either “computer” or “HP.” The possible combinations are: desktop computer desktop HP desktop computer HP laptop computer laptop HP laptop computer HP
c AND (a OR b OR d)	a b OR c OR d	(computer desktop) OR (computer laptop) OR (computer HP)	Every web page that contains “computer” and “desktop” or “laptop” or “HP.” The possible combinations are: computer desktop computer laptop computer HP
c AND (a OR b) AND d	a c d OR b c d	(desktop computer HP) OR (laptop computer HP)	Every web page that contains “desktop” and “computer” and “HP” or every web page that contains “laptop” and “computer” and “HP.” The possible combinations are: computer desktop HP computer laptop HP



### Note:

Parentheses are used in the table and in the following discussion to show the order in which Google operates on the search words and operators. The parentheses are not part of the operation.



At this point, I should mention that there is a discrepancy between how Google says it interprets the Boolean expression “c AND (a OR b) AND d” and how the expression is interpreted in Boolean algebra:

a c OR b d: Google interpretation

a c d or b c d: Boolean algebra interpretation

Examples of the two different interpretations are:

c = shirt

a = white

b = blue

d = polo

shirt AND (white OR blue) AND polo

shirt white polo OR shirt blue polo: Boolean algebra interpretation

shirt white OR blue polo: Google’s interpretation

Recall the admonition above to arrange your text into search patterns that made sense. Note that the Google interpretation is nonsensical on the right side of the OR. Yes, this implementation of the Boolean expression will contain all web pages that pertain to blue polo shirts, but the results will also contain every web page with the words “blue” and “polo,” which will include every web page about the sport polo if the word “blue” is anywhere on the page.

As of publication, Google has not implemented the NOT operator. On the Google web site, Google claims that the NOT operator is functional, but my experiments show that it is not a functioning operator at this time. Perhaps by the time you read this material, it will be functioning.



## Search Issues

---

An annoying Google tactic is the manipulation of search results. When we searched on the word “cat” in the Double Words section (see Chapter 9), the first search result was “Caterpillar — Heavy Equipment & Engine Manufacturer.” See Figure 4.23.

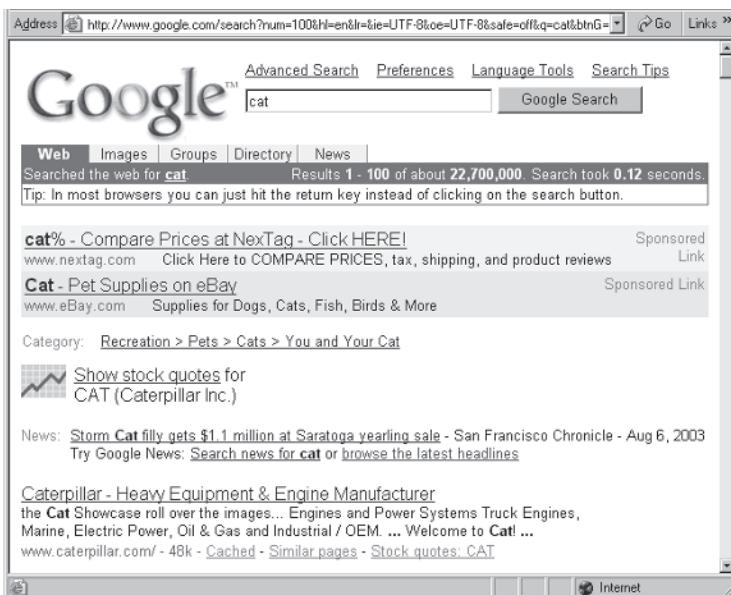


Figure 4.23: Results of a search on cat

It is apparent that Google massaged the search results to include Caterpillar Equipment as the first result. The search results appeared under the category Recreation > Pets > Cats > You and Your Cat. I understand that the word “cat” is an abbreviated term for Caterpillar, but Caterpillar has nothing whatever to do with Recreation > Pets > Cats > You and Your Cat.

A search engine based upon inviolate machine-dependent rules, such as those found in software programming, cannot arbitrarily decide to include a web page result just because it “thinks” the



result should be there. No sir. The only way “Caterpillar — Heavy Equipment & Engine Manufacturer” could appear in the category Recreation > Pets > Cats > You and Your Cat is if a human placed it there, either on purpose or by mistake. I do not like anyone manipulating my search results. Such activity can easily lead to a Big Brother approach to sharing information.

## Summary

---

In this chapter we examined the use of basic search operators to reduce the number of search results with the intention of increasing the quality of the remaining results. The basic operators are the quotes (""), the arithmetic operators, and the Boolean operators. Use of these search operators will enhance your search experience considerably and reduce your search frustration to virtually nil. With a little practice, you can become proficient searching the web and witness a tremendous increase in your productivity.

Are there other operators that we can use to reduce our search results even further? Of course. Google is chock full of operators that we can use, and we explore the advanced operators in the next chapter. However, you must keep in mind that the quality of your search results is always dependent upon how you define your search criteria and identify your keywords. All of the operators in the world will do little to help a search based on ill-defined search criteria and poorly chosen keywords. Think about what you want to find and then apply these operators to help narrow the search. As you search, use the operators to refine and narrow your search even more. Now on to the next chapter, where the advanced operators are examined.

This page intentionally left blank.

# Chapter 5

## Advanced Google Search Techniques

### Introduction

---

In the previous chapter, we discussed basic search operators and search queries employing those operators. Building upon what we have learned, let's now examine advanced search capabilities and their associated search queries. Google employs two advanced search capabilities called advanced search operators and advanced search features. Both capabilities bring to the search issue unique and interesting ways of finding the nugget of information you seek. Although they are called "advanced," these capabilities are easy to use. They are called "advanced" simply because they have limited scope and because most people do not use them every day.

Advanced search features and operators fine-tune your search, narrowing the possible choices and focusing in on the substance of your search. Most of the operators can be entered into the basic search query box, or you can select them from the Advanced Search page.



Advanced search features differ from search operators in that the features can be utilized in conjunction with the basic and advanced search operators. The features limit searches to some particular characteristic or aspect of web pages. If you have thought about your search and you know the results that you are seeking, then employing advanced search capabilities can dramatically reduce the time spent searching.

The advanced search features are:

- **Find Results:** Four query boxes that allow you to tailor the search criteria
- **Language:** Specify the language the search results are displayed in
- **File Format:** Returns results only in the specified file format
- **Date:** Restrict the search results to the past three, six, or 12 months
- **Occurrences:** Specify where the search terms occur on the page — in the title, in the URL, or anywhere on the page
- **Domain:** Search only a specific web site or exclude that site completely from the search
- **SafeSearch:** Screens for sites that contain sexually explicit-type information and eliminates them from search results. Advanced Search allows you to turn on or off the SafeSearch feature.
- **Similar:** Returns search results “similar” to a specified web page
- **Links:** Searches for pages that link to the specified web page
- **Topic-Specific Searches:** Allows you to select specific topics to search



The availability of these advanced features is not limited to the Advanced Search page located at [http://www.google.com/advanced\\_search?hl=en](http://www.google.com/advanced_search?hl=en) and available by clicking on the Advanced Search link on the Google home page. See Figure 5.1. Many of these operators can be used in the basic search query box. The specific section that discusses each in detail identifies those operators that can be used in the basic query box.

Okay, let's face it. How often are you going to search for something that is restricted to a specific URL? Well, probably never ... unless you read the remainder of this chapter and discover that all of the above search features have great utility (if you know how and when to use them). Each of these advanced search capabilities brings to the search issue the ability to focus your search in specific relief, thereby reducing the amount of low-quality search results and increasing your search efficiency. Yippee! That is just what we need to accomplish in order to reduce our search frustration to a manageable (that is, nonexistent) level.

Achieving a better position in search results is an objective that most web page designers set when they begin to design a web page/site. Knowing that these search capabilities exist and understanding how they work can help designers achieve a better position in a search result, perhaps landing the page/site in the coveted top 30. If you design web pages, careful consideration of the usage of advanced search capabilities will help you in achieving that top 30 spot.

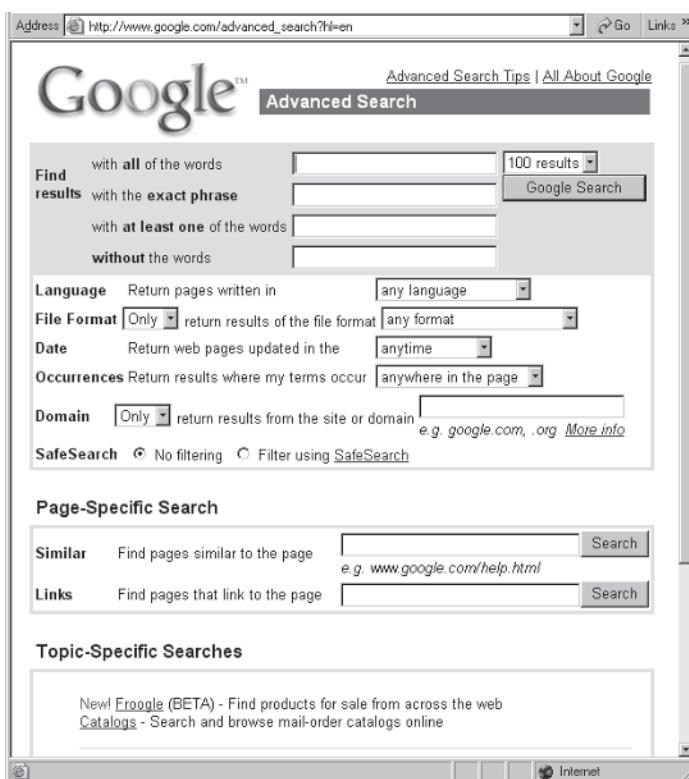


Figure 5.1: The Advanced Search window

## Advanced Search Features

---

Advanced search features differ from advanced search operators in that advanced search operators “operate” within the search engine to include or exclude web pages (that is, they are “query words that have special meaning to Google,” whereas advanced search features are boundaries placed upon the search field). To illustrate the difference, I give an example of each. The quotes operator used to search for a phrase is an operator. A feature is being able to specify that only web pages in a specific language will be searched. Strictly speaking, there is some overlap between operators and features. To make a greater distinction, operators are words or symbols used



within the query box to include or exclude web pages, whereas features are implemented outside the query box via special mouse clicks, option selections, or special text boxes.

## Find Results

Find Results is a set of four text boxes in the Advanced Search web page that allow you to include and/or exclude keywords in your search. See Figure 5.2.

The uppermost text box, With **all** of the words, allows you to specify one or more keywords that must be somewhere within the web page. The keywords do not have to appear in the order in which you placed them in the query box and neither do they have to appear in any sequence. As long as each word is somewhere in the web page, it will be identified as a search result. This approach to searching, as in other things in life, is called the “shotgun” search approach.

The screenshot shows a set of four text input fields labeled from top to bottom: "with all of the words", "with the exact phrase", "with at least one of the words", and "without the words". To the right of the first field is a dropdown menu set to "100 results" and a "Google Search" button.

Figure 5.2: The Find Results text boxes

The With the **exact phrase** text box is used to specify a sequence of keywords that form a phrase you want to search for. This is the same as placing double quotes around a phrase in a general query box or the With **all** of the words text box in the Advanced Search page (see the above comments about this text box). An example of this type of search is “civil war.” You may recall that we illustrated the use of quotes as an operator in a previous chapter. Placing the words “civil war,” without the quotes, in the With the **exact phrase** text box and searching is the same as placing the text with quotes in the With **all** of the words query box above or in a general query box such as the one on Google’s home page.



A query using the With **at least one** of the words query box requires that at least one of the keywords entered in the query box appear on the web page. Such a search on “civil war” would return every web page containing the word “civil” and every page containing the word “war.” Using this query box is the same as using a general query box with the OR operator between each of the keywords.

Using the **Without** the words query allows you to exclude web pages that do not pertain to the specific search goals that you have. To illustrate the usefulness of this box, let’s say I searched on “civil war” and received thousands of search results for web pages about Abraham Lincoln. But I am not interested in Abraham Lincoln; rather, I am searching for information about battles. In this case, I could key in “Lincoln” and those pages containing “Lincoln” would be excluded from the search results. Neat, eh? Using this query box is the same as using a general query box with the NOT operator between each of the keywords.

## File Format

Recall from Chapter 1 that a list of file types was identified for which search engines could not access the content. At the bottom of the list, I mentioned there were certain types for which Google has developed the expertise to search the content. This section discusses those file types.

File Format allows you to select a specific file format that will exclude all search results without the specified format. Google is currently the only search engine with a user interface specifically designed to search for file formats. To select a particular file format, click on the drop-down menu (see Figure 5.3) and highlight the file format of your choice. As you can see, there are currently six different formats you can choose from. The default format is Any format. When you left-click on a format to choose it, the choice is highlighted, signifying that it is selected. Then the menu will close.

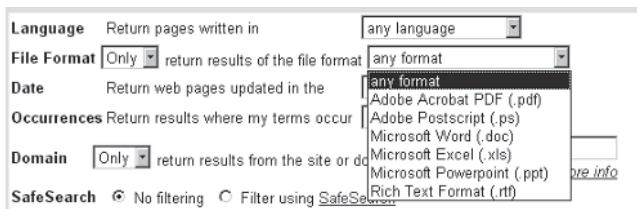


Figure 5.3: The File Format menu

On the same line as File Format, the Only drop-down menu allows you to select “only” that file format, or you can exclude that file format by selecting Don’t. See Figure 5.4.



Figure 5.4: The File Format choice

When you have some experience searching the web, you may begin to find that there are certain document types that you can exclude because experience has taught you that search results of a particular document type are useless for your search goals. However, I think the only time that I would exclude a document type is if I was searching for the technical specifications of a semiconductor component. Technical specs are typically written using Word, whereas technical information (user guides, applications, service information, etc.) is written in Acrobat. Since I am seeking specs only in a .doc format, I would exclude the numerous .pdf files from my search objectives. In pursuit of your profession, hobbies, or interests, you may also benefit from such an approach to customizing your search criteria.



## Date

Date restricts your results to web pages that have been updated recently. The choices you can make are:

- Anytime
- Past 3 months
- Past 6 months
- Past year

Date is particularly useful when searching for news reports and other timely matter. To use Date, just click on the Date drop-down menu on the Advanced Search page and select the range you are interested in.

The screenshot shows the Google Advanced Search form. On the right, the 'Date' dropdown menu is open, displaying four options: 'anytime', 'past 3 months', 'past 6 months', and 'past year'. The 'anytime' option is currently selected. Other dropdown menus for Language ('any language'), File Format ('any format'), and SafeSearch ('No filtering') are also visible.

Figure 5.5: Date search

## Occurrences

Occurrences specifies where your search terms occur on the page. You can select these choices:

- Anywhere in the page
- In the title of the page
- In the text of the page
- In the URL of the page
- In links to the page

Anywhere in the page is self-explanatory. The search will be conducted by examining the complete contents of every web page. Actually, the contents are already “examined” and indexed by the



search engine before you search. The search engine examines its index of words to locate those web pages containing your search words.

The screenshot shows the Google Advanced Search interface with the 'Occurrences' dropdown menu open. The menu options are:

- anywhere in the page
- in the title of the page
- in the text of the page
- in the URL of the page
- in links to the page

Figure 5.6: Occurrences search

The In the title of the page option means the search will be confined to web page titles only. This is useful for zeroing in on the subject matter, but it leaves out web pages with ancillary information that may or may not be useful in your search. If you seek specific information about concrete subjects, this may be the most useful method of performing a general search and keeping your blood pressure to a manageable level. Setting this preference is probably the wisest decision that you can make concerning your approach to searching the Internet. If you experience “search rage,” I recommend selecting this choice and saving it as a personal preference. The relationship between your search terms and your search objective, assuming you have chosen your search terms properly, will be a one-to-one correspondence, meaning you should hit the bull’s-eye in your search quest. However, if you are an adventuresome soul and enjoy finding the unusual or odd thing now and then, I suggest that you leave this preference set to Anywhere in the page.

In the text of the page means the search is confined to the body of web pages. If you are searching for a phrase or clause contained within the body of a page, this search option will help you zero in on your search objective. I am not sure how it can otherwise help in eliminating unwanted web pages from your search.

In the URL of the page means the search will be confined to web page URLs only. This limits the scope of your search to just a few



web pages that contain the keyword(s) in the URL only. This option should help you zero in on companies, businesses, organizations, schools, government entities, and individuals with alacrity.

The In links to the page option searches only web pages that contain the keyword(s) in a URL link. A URL link is a string of characters that define a path to another web site. Hypertext Markup Language (HTML) documents place these URL links within the body of a web page so users can “jump” to another web page, either within the same domain (web site) or in another new domain.

The text “<A HREF=”<http://www.adobe.com/products/ebook-reader/register.html>”>” in a web page is an example of a link to Adobe’s eBookReader web site where a user can register and download a copy of eBookReader. A keyword search for “Adobe” confined to the In links to the page option would only return web pages with links such as this example buried within the page.



#### Note:

I tested each of the Occurrences options using well-known company names. I did not discern any significant differences in the number of search results. The results I achieved using “Google” as the keyword are shown in Figures 5.7 through 5.11. Note that each test returned essentially the same number of results, “about” 20,000,000. There is an insignificant variation (+/-400,000) in the results due to the time of search (0.13 seconds vs. 0.27 seconds). I am not sure what is going on here. The first few search results in every test case shown are the same web pages. It appears from this limited test that this feature of Google does not work, at least as advertised. Perhaps there is a significant difference in the search results if the later search results are examined.

Web Images Groups Directory News  
Searched the web for **Google**. Results 1 - 10 of about **20,300,000**. Search took **0.13** seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 5.7: Anywhere in the page

Web Images Groups Directory News  
Searched the web for **allintitle: Google**. Results 1 - 10 of about **19,900,000**. Search took **0.27** seconds.  
Tip: In most browsers you can just hit the return key instead of clicking on the search button.

Figure 5.8: In the title of the page



**Web** Images Groups Directory News |  
Searched the web for **allintext: Google**. Results 1 - 100 of about **21,000,000** Search took **0.96** seconds.

Figure 5.9: In the text of the page

**Web** Images Groups Directory News |  
Searched the web for **allinurl: Google**. Results 1 - 100 of about **21,000,000** Search took **0.39** seconds.

Figure 5.10: In the URL of the page

**Web** Images Groups Directory News |  
Searched the web for **allinanchor: Google**. Results 1 - 100 of about **20,700,000** Search took **0.39** seconds.

Figure 5.11: In links to the page

For some reason, Google places the tag “allinanchor” instead of “allinlink” in the message bar displayed with the search results when you search with the In links to the page option. See Figure 5.11.

When this feature works as it should, searching with Occurrences will be balm to all of us suffering from search rage.

## Domain

Sometimes, we may only want to search a specific web site, called a *domain*, or we may want to exclude a specific web site from our search results. We might want to perform this site-specific search at our place of employment, especially if the company is a large one and has many web pages and/or web sites. Or, we might want to restrict our search to a specific geographical region. On occasion, searching a domain is easier than trying to navigate through numerous web pages that do not apply to the circumstances. I have found restricting my search to a domain useful when looking for information about a particular person, place, or thing. As an example, I have found limiting my search to a domain is very useful when trying to find technical information about a product or service in the performance of my engineering profession.

Also, searching a consumer catalog for a particular product can be trying, but when the search is restricted to the domain, finding your gadget becomes elementary. So, there are times when you can benefit from a domain search. The difficulty lies not in using the tool



but knowing when to use it, and that comes with a little thinking and even less practice. Of course, if you do not know the domain name of the site you wish to search, you cannot use the domain feature until you search for and discover the domain name. Also, you must include the word “site,” as shown in the search query box.

We illustrate the power of domain searching with two different examples and illustrate the power of domain searching combined with Google’s language translation service with a third real-life example later in this chapter. Catalog shopping is a national pastime that probably outranks baseball. There is no telling how many hours are spent across the nation and world thumbing through catalogs on a Sunday afternoon. No doubt, such activity is chicken soup for the soul. But what if you are at work and all of a sudden you remember that a nephew or niece has an upcoming birthday or special event? Do you want to spend a couple of hours on company time searching for his or her gift? (Okay, so maybe you do, but does your boss want you to?) Using domain-restricted searches will narrow the field and help you zero in on your search objectives very quickly.

The first example we use is from the L.L. Bean catalog. Currently, the L.L. Bean catalog is 64 pages. Let’s say that we have a nephew who likes to camp, and he needs a new tent. We know L.L. Bean sells good-quality camping equipment that we can afford to purchase; we know this from spending Sunday afternoons thumbing through the L.L. Bean catalog! Or maybe we are just guessing that L.L. Bean can meet our needs. In either case, this approach to solving the search objective is faster than just entering “tent” in a basic search and getting a couple of million returns to muddle through.

We enter the information in the search query box exactly as shown in Figure 5.12. Do not place a space between the colon and the domain name. After entering the search criteria, left-click on Google Search.



### **Caution:**

Do not place any characters, such as the forward slash (/), after the “.com.” If you do, you will not have a successful search.

Figure 5.12: Site search for tent

The search results will be displayed momentarily. See Figure 5.13.

Address http://www.google.com/search?as\_q=tent+site%3Awww.llbean.com&num=10&hl=en&ie=UTF8&oe=UTF8&btt=G+Google+Search Go Links >

**Google™** Advanced Search Preferences Language Tools Search Tips

tent site:www.llbean.com Google Search

Web Images Groups Directory News | Searched pages from www.llbean.com for tent. Results 1 - 10 of about 132. Search took 0.06 seconds. Tip: In most browsers you can just hit the return key instead of clicking on the search button.

**Tent At Sears** www.sears.com Save on Shipping with the in-store pick up option! Sponsored Link

**LL.Bean: Choosing A Tent** ... Choosing a Tent Choosing the proper **tent** depends primarily on where and how it will be used. For example, an expedition dome **tent** ... www.llbean.com/outdoorsOnline/outdoorSports/camping/tips/choosetent.html?featIn=23k - Cached - Similar pages

**LL.Bean: Summer Camping FAQs** ... Q. Do ground cloths belong under or inside the **tent**? A. Ground cloths are intended to protect the floor of your **tent** from excessive ... www.llbean.com/outdoorsOnline/outdoorSports/camping/FAQs/?featIn=27k - Cached - Similar pages

**Tent** Compare Prices at 30,000 Stores. Find the Best Deals at BizRate.com! www.BizRate.com Interest: ■■■■■

**High Quality Canopies** And Pop Up Tents at Wholesale Prices www.accecanopy.com Interest: ■■■■■

Figure 5.13: Site search results for tent

In the search results shown in Figure 5.13, notice that the web site at the top of the page is not the web site that we listed in the search query box, an L.L. Bean site, but rather a Sears site. This is a paid advertisement — what Google calls “Sponsored Links” — as indicated by the text at the right side of the page. To keep you from confusing search results and sponsored links, Google labels those links and places them in colored boxes. In this example, all of the actual search results below the sponsored links are from L.L. Bean.

For the next example, let’s try a little genealogy research. A popular web site for genealogy researchers is [www.rootsweb.com](http://www.rootsweb.com). Let’s say that I am interested in finding every web page with my surname at [rootsweb.com](http://www.rootsweb.com). To achieve this search objective without having to wade through thousands and thousands of web pages, I enter the domain-restricted site search criteria, as shown in Figure 5.14.


 10 results 

Figure 5.14: Site search for Busby

Address: http://www.google.com/search?as\_q=busby+site%3Awww.rootsw.com&num=10&hl=en&ie=UTF-8&oe=UTF-8

**Google™** Advanced Search Preferences Language Tools Search Tips

busby site:www.rootsw.com Google Search

Web Images Groups Directory News | Searched pages from www.rootsw.com for busby. Results 1 - 10 of about 3,270. Search took 0.10 seconds. Tip: In most browsers you can just hit the return key instead of clicking on the search button.

**ROBERT AUGUSTUS BUSBY**  
ROBERT AUGUSTUS BUSBY. Robert Augustus **Busby** was born a slave in 1822 in Kanawha County, West Virginia, which at the time was western Virginia. ... www.rootsw.com/~ilissda/text\_files/database\_intro\_files/oral\_histories/robert\_busby.htm - 8k - [Cached](#) - [Similar pages](#)

**Busby Info**  
David T. **Busby**, born Nov 16, 1807 in unk Co of KY. Son of Matthew **Busby** and Seluda  
Tinsley **Busby**, originally from VA, (Seluda was from Amherst Co, VA). ... www.rootsw.com/~txfalls/busbyInfo.htm - 2k - [Cached](#) - [Similar pages](#)

**JF Busby**  
... Yours father. JF **Busby**. I had found this letter and others in my mothers stuff Her name Betty Lee Newton. Submitted by SShorty5@aol.com. www.rootsw.com/~okbits/bettylenewton.html - 3k - [Cached](#) - [Similar pages](#)

Sponsored Links  
[Save up to 70% on Hotels](#)  
Great rates on Hotels & Packages  
[Save on International Airlines](#)  
www.AirfarePlanet.com  
Interest:   
[See your message here...](#)

Figure 5.15: Site search results for Busby

The results of my rootsweb.com site search are displayed in Figure 5.15. I could narrow the search results further by searching on a specific name, such as “Michael Busby.” When I performed such a search, 1,350 results were returned. The results included all web pages that contained both “Michael” and Busby” but not necessarily together. Let’s say that I am interested in only those web pages that contain “Michael Busby” as a matched pair of words. I can combine the restricted site domain search with the basic quotes search operator to further restrict my search to only those web pages at rootsweb.com containing “Michael Busby.” See Figure 5.16.

<input type="text" value='["Michael Busby" site:www.rootsw.com]"/> 10 results 

Figure 5.16: Using the basic quote search operator



The results of searching for web pages containing only “Michael Busby” on rootsweb.com are shown in Figure 5.17. The results demonstrate the power of restricted domain searching combined with other search operators. Go ahead and search on your first and last name. Interesting, eh?

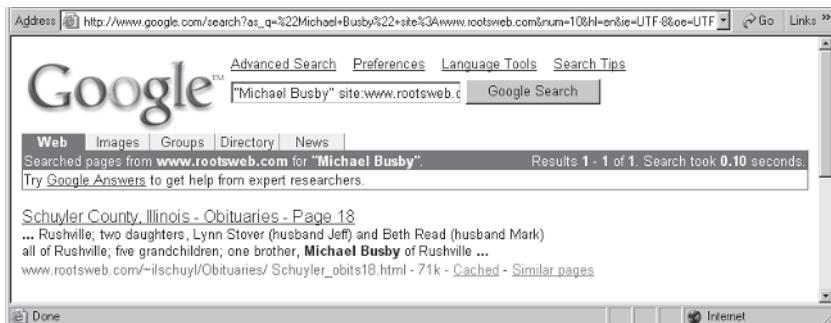


Figure 5.17: Results of a restricted domain search

For technical people, such as engineers, a restricted domain search can help you quickly find a manufacturer’s or vendor’s product. I cannot recall the untold hours I spent searching such sites as Motorola and National Semiconductor web page by web page, trying to find the technical details of a particular semiconductor that I wanted to use in a design. Using a restricted domain search to find such information reduces the effort from a Herculean task to a minuscule issue.

Here is an example of the need to do a restricted domain search that I came across while cruising the Internet. The individual was looking for Harry Potter information, and he did not want to spend time examining pages in the .com domain. Apparently, he was not aware of the power of restricted domain searching because he examined the .com web pages and “checked them out quickly.” If the search was [“Harry Potter” -.com], the search results would not include any commercial web sites. The text of his search experience follows:



“I was looking for Harry Potter web sites [a famous fictional book character] and noted *I was avoiding links that looked like corporations* in the search results page (I did a search on Google).”

Using restricted domain searching increases your productivity immensely. This is a work tool that undoubtedly will help you “work smarter, not harder.” That is the name of the game. When you can increase your productivity, you have a corresponding increase in efficiency, and that is the best time to go into the boss’s office and ask for a pay raise. If you don’t get it, go find another job that pays better! You deserve it.

## Language Tools and Services

Google’s language feature is multifaceted. You can search pages written in any of 35 languages regardless of the web page’s physical location. You can search for pages physically located in any one of 36 countries regardless of the language the web page is written in. Within the body of Google’s Translate page, you can translate words, text, documents, or web pages. You can set Google’s interface to display messages in any one of more than 80 languages. Finally, you can search specific country domains and easily view the web pages in English or the language in which they were written.

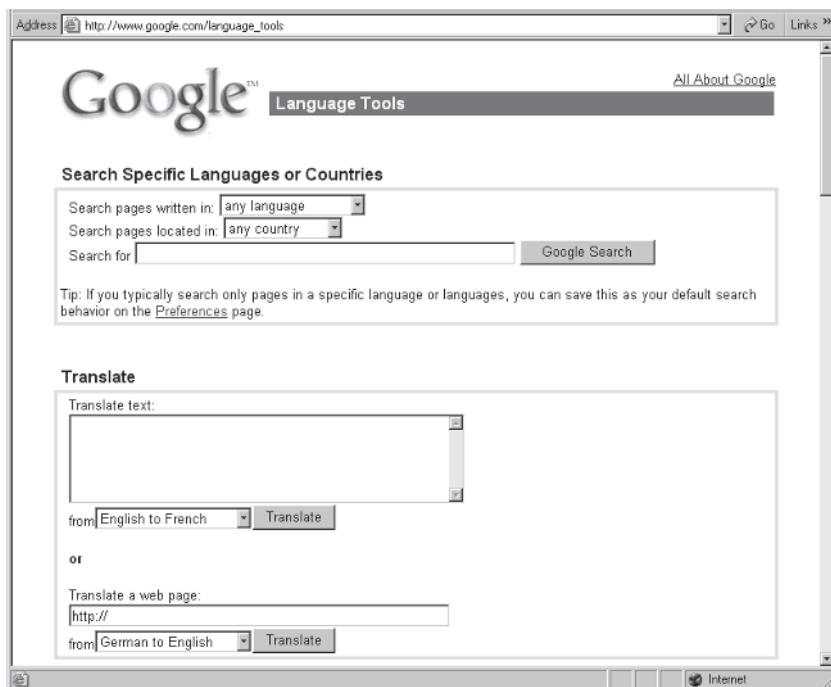


Figure 5.18: Language Tools page

You can find Language Tools at [http://www.google.com/language\\_tools](http://www.google.com/language_tools) or by left-clicking on the Language Tools link on the Google home page. Be sure to bookmark this page, as you will use it often if you do any searching outside the usual English domains. One of the neatest features of Google's Language Tools is translation. The Translate drop-down menu offers the following translation choices:

- English to German
- English to Spanish
- English to French
- English to Italian
- English to Portuguese
- German to English
- German to French
- Spanish to English



- French to English
- French to German
- Italian to English

You can translate words, phrases, sentences, documents, or whole web pages. The translation service is pretty good. If you want to translate a document, just copy the text to the clipboard and then paste the text into the Translate text box. See Figure 5.19. The translate service works on plain text only.

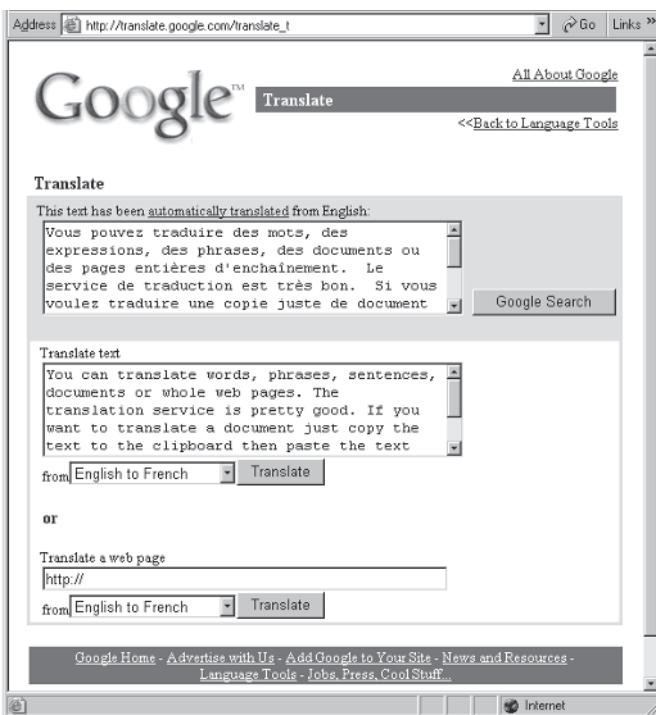


Figure 5.19: Translate box

I experimented by translating the text of a Word document from English to French back to English. The service faithfully translated the document word for word, resulting in some error in meaning by the time the document was translated back into English. But for words, phrases, and short text, the service works quite well. It is certainly much better than nothing!



As an example, I selected the short sentence “I love you” and translated it from English to French back to English. The results were:

I love you — Je t'aime — I love you

See Figure 5.20. Now left-click on Google Search and find “Je t'aime” on the web.

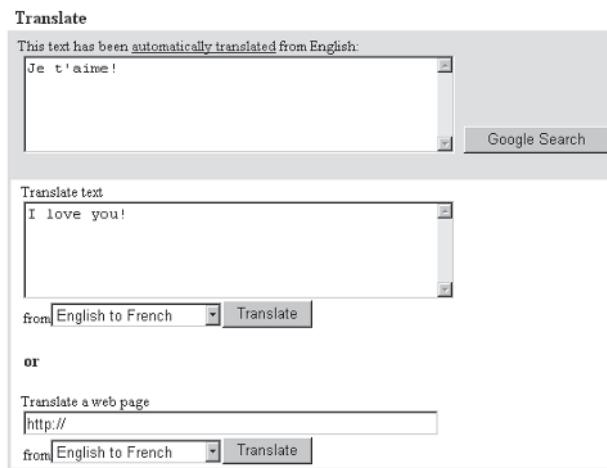


Figure 5.20: Translating “I love you”

Now, I decided to really put the service to the test. I selected the same sentence (“I love you”) and translated it into French, then German, and then back to English. Was anything lost in the multiple translation? Well, you decide:

English to French:

I love you — Je t'aime

French to German:

Je t'aime — Ich mag dich

German to English:

Ich mag dich — I like you

See Figures 5.21 and 5.22.

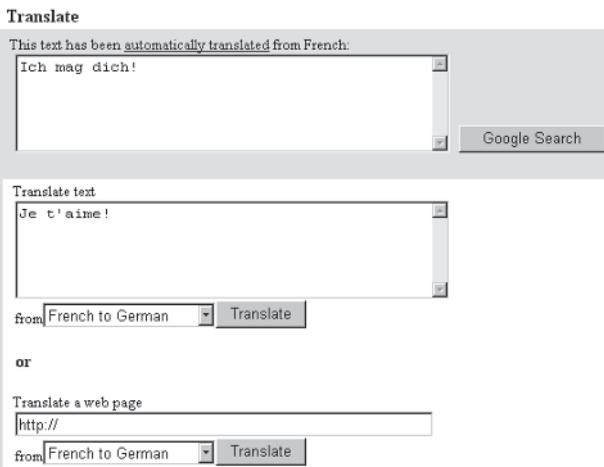


Figure 5.21: Translating “I love you”

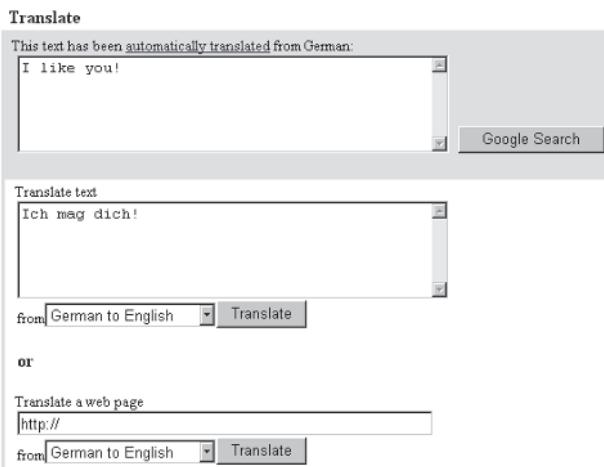


Figure 5.22: Translating “I love you”

From my French and German classes taken many years ago, I believe the translation should be “Je t’aime — Ich liebe die.” As we would say in computer programming, there is an interface issue here. For the record, “I like you” translates differently from English to French:

I like you — *Je vous aime* vs. *Je t'aime* — I love you



Next, I decided to test Translate's ability to translate a more complicated sentence. I selected "The quick brown fox jumped over the lazy dog's back" as a test case, and the results are shown below.

Original English text:

The quick brown fox jumped over the lazy dog's back

English to French:

Le renard brun rapide a sauté par-dessus le chien paresseux en arrière

French to German:

Der schnelle braune Fuchs ist über dem faulen Hund rückwärts gesprungen

German to English:

The fast brown fox jumped over the lazy dog backwards

English to French to English:

The fast brown fox jumped over the lazy dog behind

English to German to English:

The fast brown fox jumped back over the lazy dog

Somehow we lost the information about the dog's back but otherwise, the translation is reasonably accurate.

An adventurous soul can have fun with just a little creativity. One would think that if you put an English language sentence in the Translate box and click on Translate with English to English selected that you would get back exactly what you requested translated. After all, English is English, right? Unfortunately, we do not have an option to translate English to English, but we can perform a similar, although not exact, translation by pretending the German translation of "The quick brown fox jumped over the lazy dog's back" is really an English sentence. We place the German translation of the sentence back in the Translate box, select English to German, and click on Translate. We copy that seemingly



unintelligible German text back into the Translate box, select German to English, and click on Translate again. We now have a nonsensical sentence that does contain a couple of interesting or entertaining phrases. See below. What an excellent way to create new dialogue for a best-selling book! Okay, the point here is that you need to exercise some caution when using Translate. Do not use this tool for legal purposes, unless you like to live dangerously.

Original English text:

The quick brown fox jumped over the lazy dog's back

English to German:

Der schnelle braune Fuchs sprang zurück über den faulen Hund  
zurück

German to German:

Schnellebraune Fuchs Der entsprang zurücküberden faulen  
zurück Hund

German to English:

Fast-brown fox rose practice-grounds putrefies back to dog

When you use Translate, ensure that the words are spelled correctly. When I use Translate, I key any text into a Word document, use Word's spell checker, and then copy/paste the text into the Translate text box. The difference can be significant. I intentionally misspelled "quick" in the next example to illustrate how the Translate tool manages misspelled words.

Original English text:

The qucik brown fox jumped over the lazy dog's back



English to French:

Le renard de brun de quicik a sauté par-dessus le chien paresseux en arrière

The English to French translation without the misspelled word is:

Le renard brun rapide a sauté par-dessus le chien paresseux en arrière

French to English:

The fox of brown of quicik jumped over the lazy dog behind

Always ensure that the text you are translating is spelled correctly, regardless of the language that you are translating. I checked to see if using punctuation, such as apostrophes, made any difference in the translation. I could not isolate any case where punctuation, whether present or missing, made any difference in the translation.

## Google in Your Language

Google in Your Language is a collaborative effort to translate the Google home page, messages, and tips into the various languages of the world. Google is recruiting volunteer translators. The languages that Google currently supports are shown below. I do not know if Google is still accepting volunteers for the languages shown. If you are interested, visit the Google in Your Language program at <http://www.google.com/intl/en/language.html>. (Bookmark it! You cannot find it again if you don't bookmark it.)

Afrikaans	Albanian	Amharic	Arabic
Azerbaijani	Basque	Belorussian	Bengali
Bihari	Bork bork bork	Bosnian	Bulgarian
Catalan	Chinese (Simplified)	Chinese (Traditional)	Croatian
Czech	Danish	Dutch	Elmer Fudd
English	Esperanto	Estonian	Faroese
Finnish	French	Frisian	Galician



Georgian	German	Greek	Gujarati
Hacker	Hebrew	Hindi	Hungarian
Icelandic	Indonesian	Interlingua	Irish
Italian	Japanese	Javanese	Kannada
Klingon	Korean	Latin	Latvian
Lithuanian	Macedonian	Malay	Malayalam
Maltese	Marathi	Nepali	Norwegian
Norwegian (Nynorsk)	Occitan	Persian	Pig Latin
Polish	Portuguese (Portugal)	Punjabi	Romanian
Russian	Scots Gaelic	Serbian	Sinhalese
Slovak	Slovenian	Spanish	Sudanese
Swahili	Swedish	Tagalog	Tamil
Telugu	Thai	Tigrinya	Turkish
Ukrainian	Urdu	Uzbek	Vietnamese
Welsh			

Google includes several uncommon languages and dialects of the English language in its suite of languages in which you can choose to display Google messages, tips, and text. Check out Bork bork bork, Interlingua, Elmer Fudd, Hacker, Klingon, and Pig Latin. I did, and two of the results are shown below. Bookmark Google's Preferences page so you can easily navigate back to it. Go ahead; it is fun, and you cannot mess it up. Go to Google's Preferences page at <http://www.google.com/preferences?hl=en> and select the language of your choice. See Figure 5.23. Then left-click on Save Preferences.

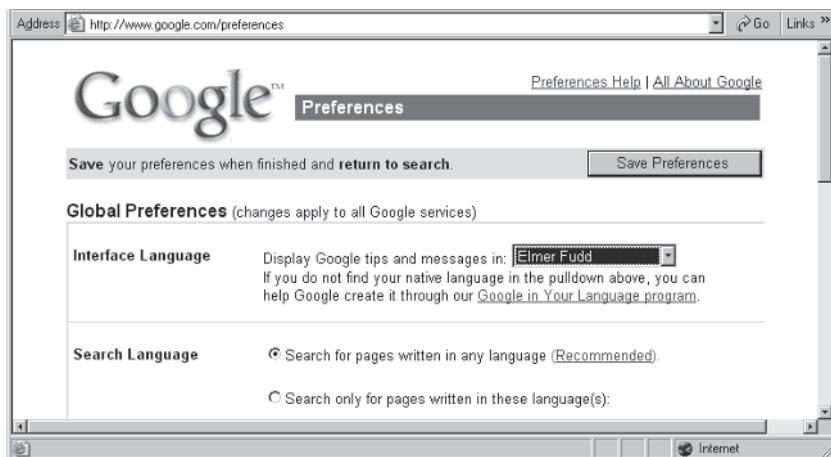


Figure 5.23: Global Preferences page

After you save your preferences, a pop-up box will tell you that they have been saved. See Figure 5.24. Left-click on OK, and you will be returned to the Google home page.



Figure 5.24: Saving preferences

If the resulting page is disconcerting, just click on the Google.com in English link at the bottom of the page. Voilà — now the page is displayed in English. Or return to the Preferences page via the bookmark that you saved above, and select another language. Save that choice and enjoy the Google home page in that language. Be adventurous and do a search or two in the language of your choice.



Figure 5.25 shows a Google search page using the Elmer Fudd dialect! Notice the peculiar spelling (and pronunciation) of some of the words. All Google messages and text are translated into Fudd, but web pages themselves are not translated.

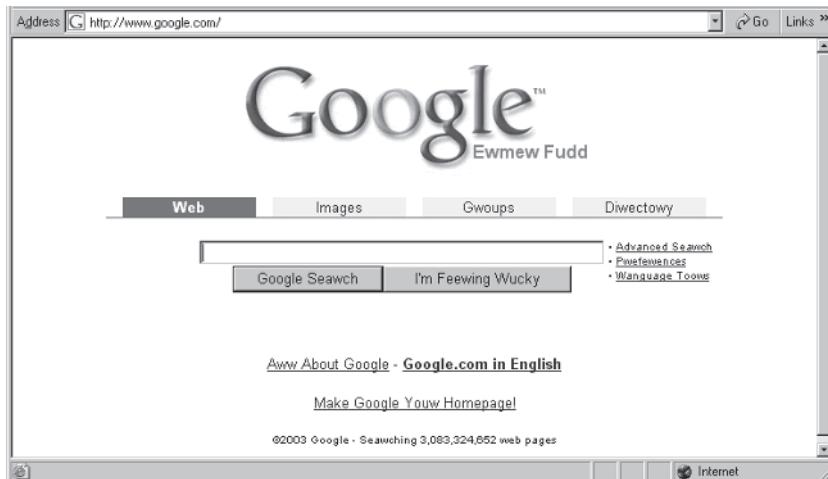


Figure 5.25: Google search page displayed in Elmer Fudd

The results of an Elmer Fudd “dialect” search are shown in Figure 5.26. Since the web page information is still in English, I did not show any search results. But, as you can see, Google text messages are in Elmer Fudd.

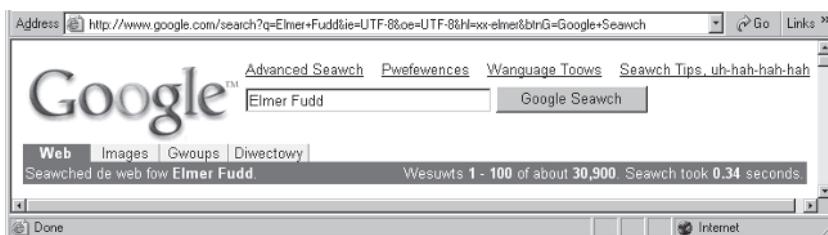


Figure 5.26: “Elmer Fudd” search results



Figure 5.27 shows a Google search page displayed in Hacker. Note the unintelligible text (even to hackers!). Google demonstrates a great sense of humor by including these “dialects” and odd “languages” in its suite of languages.



Figure 5.27: Google search page displayed in Hacker

## Searching for Hilux: A Real-life Search Example

Here is an excellent real-life example for restricting a search to a domain utilizing Google’s language translation tool and services. My niece and her husband, Brian and Laura Kelly, with their three sons, have been missionaries in Guyana, South America, for four years. This summer they were visiting in the States with a goal of registering their oldest son, B.J. (a spring high school graduate), in college and raising funds for their missionary work. Due to Guyana’s visa restrictions and work laws, they could not return to Guyana on a work permit, having stayed as long as Guyana permits. Instead, they are going to Brazil and will establish a base of operations in Boa Vista, visiting Guyana on six-week tourist visas. To get to Boa Vista, they will fly into Rio de Janeiro, purchase a vehicle,



and drive to Boa Vista, a long and hazardous journey. To minimize the length and cost of their stay in Rio de Janeiro, they asked me to see if I could find a four-wheel drive vehicle for sale in Rio de Janeiro. Using Google's Translate and Advanced Search pages in combination with a restricted domain search, I was able to locate several suitable vehicles for sale. I was able to manage this challenge and succeed by restricting my searches to the domain ".br," which is the country-wide domain for Brazil.

When the Internet was first organized, it was the exclusive domain of the United States government, defense contractors, and college campuses. There was no need for country domains in those early days, as the U.S. was the only domain. At some point in the evolution of the Internet, it was determined that the division of the Internet into country-wide domains would facilitate the ability to locate information in a more orderly fashion. So, the Internet-governing body at the time, InterNIC, assigned two-letter country-wide domain "names." Simple, eh?

How do you determine a country's domain? Well, the topic of this book is using Google to search the Internet, so why don't we just plug "country domain name" into the Google search box and hit Search? The first result I received when I performed the search was "UNINETT:name registries around the world" at [www.norid.no/domenenavnbasen/domreg.html](http://www.norid.no/domenenavnbasen/domreg.html). Note that the domain names are in alphabetical order, which do not necessarily correspond with alphabetical country names. See Figure 5.28.

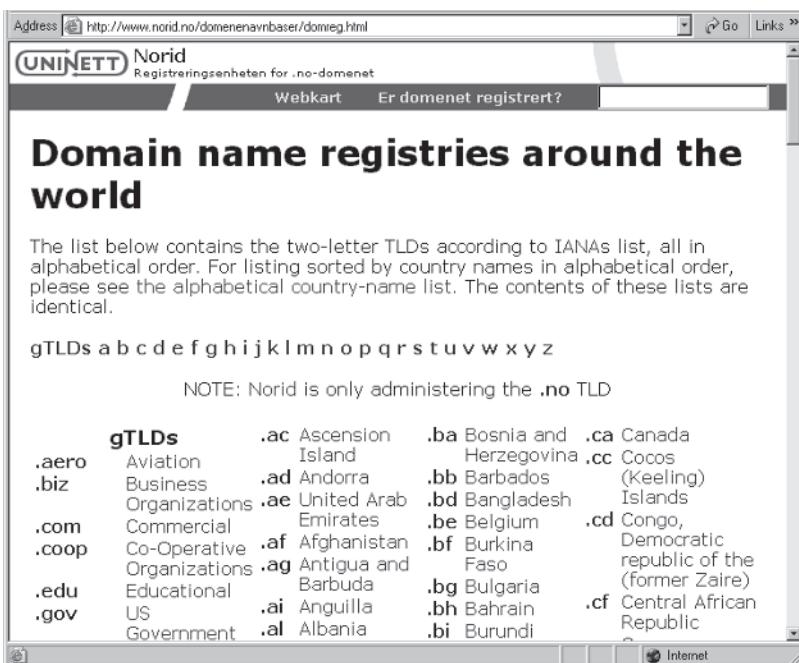


Figure 5.28: Top-level domains

In our example, we are interested in Brazil, so we select the country domain name .br. Before we go to the Advanced Search page to initiate our search, there is another task that we need to accomplish. Assuming people in Brazil advertise their products with the word “sale” (as in “for sale”), we need to translate the word “sale” into the official language of Brazil, which is Portuguese. We go to Google’s Translate page (hopefully by now you have started bookmarking all of these pages for easy access later). After entering “sale” in the Translate text box, we left-click on the Translate button, and magically the word “venda” appears in the This text has been automatically translated from English text box. See Figure 5.29.

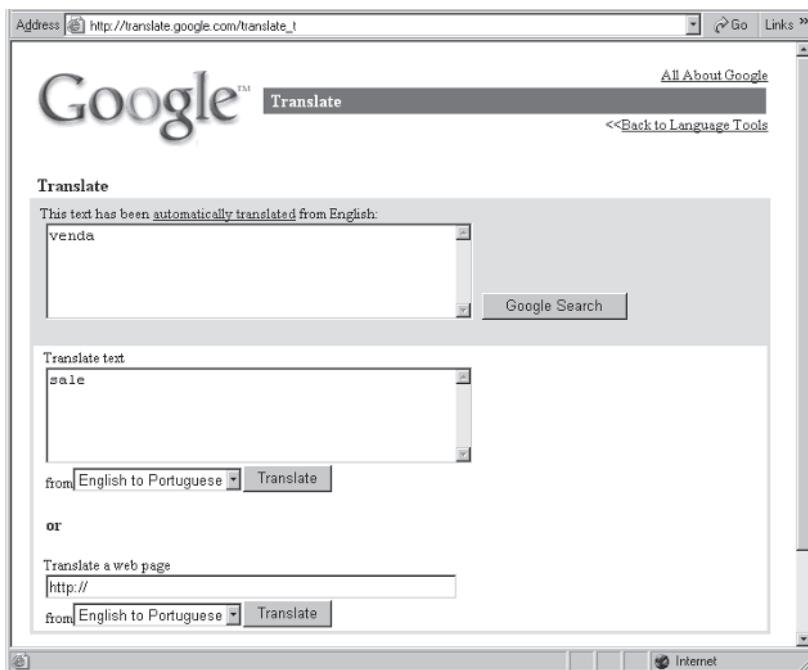


Figure 5.29: Translate page

We could translate “for sale” (the translation would then be “para a venda”), but the “for” (para a) is a common word that does not add any depth or quality to our search results. Next, we journey to the Advanced Search page and key the query words “Toyota Hilux venda Rio de Janeiro” into the search box. Then we go to the Domain line, select Only from the drop-down menu, and key in “.br” in the domain text box. Notice in Figure 5.30 that I selected 50 results per page. We are ready to search. Left-click on Google Search.



The screenshot shows the Google Advanced Search interface. The search query in the main search bar is "toyota hilux venda Rio De Janeiro". Below the search bar, there are several search parameters:

- Find results**:
  - with all of the words: toyota hilux venda Rio De Janeiro
  - with the exact phrase: (empty)
  - with at least one of the words: (empty)
  - without the words: (empty)
- Language**: Return pages written in: any language
- File Format**: Only return results of the file format: any format
- Date**: Return web pages updated in the: anytime
- Occurrences**: Return results where my terms occur: anywhere in the page
- Domain**: Only return results from the site or domain: .br e.g. google.com, .org [More info](#)
- SafeSearch**:  No filtering  Filter using SafeSearch

Figure 5.30: Searching for Hilux

We receive several search results. I have cropped all but one in Figure 5.31. The important point here is that all you need to do is left-click on the Translate this page link, and that web page will soon appear in your browser translated from Portuguese into English, of a sort. The translation is a word-for-word translation, and language nuances are ignored. So the web page will not measure up to Ms. Stern's high school English teacher standards, but it is sufficiently understandable to achieve our objective — identify and price a Toyota Hilux for sale in Rio de Janeiro, Brazil. Figure 5.32 is a shot of the page after translation.



Figure 5.31: Search results

Note that you can view the web page in its original language by left-clicking on View Original Web Page at the top of the page. See Figure 5.32.

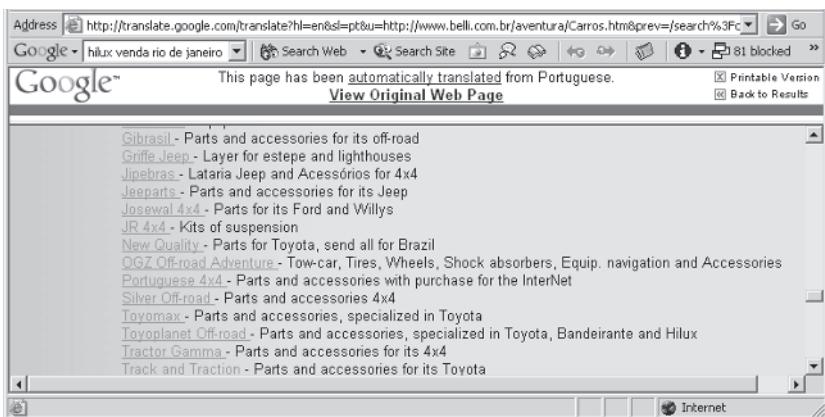


Figure 5.32: Translated web page from Portuguese to English

How long did it take for me to find a particular vehicle for sale that was thousands of miles from my home? The actual time I spent searching was just a few minutes. The contact information gleaned from the web page was sufficient for my niece and nephew to contact the seller and arrange the purchase a week before their arrival in the country.



Reading this book and practicing the examples herein does not (by itself) magically make anyone a search guru. Utilizing your skills to work more effectively, whether working for corporate America or yourself in your leisure time, is a conscious personal decision-making process. You must approach searching with a positive attitude.

Exercising what you learn in the spirit of adventure and discovery, you can become much more efficient, ergo productive, and much less stressed out.

## SafeSearch Filtering

The SafeSearch filtering option was examined in depth in Chapter 3.

## Similar

The Similar feature allows you to find web pages “similar” to a web page that you specify by entering the web page domain name in the Similar query box. See Figure 5.33. To test the utility of this feature, I keyed in the domain name of Advanced Search ([http://www.google.com/advanced\\_search?hl=en](http://www.google.com/advanced_search?hl=en)). Next I left-clicked on the Search button.

Page-Specific Search

Page-Specific Search	
Similar	Find pages similar to the page <input type="text" value="e.g. www.google.com/help.html"/> <input type="button" value="Search"/>
Links	Find pages that link to the page <input type="text"/> <input type="button" value="Search"/>

Figure 5.33: Searching for similar pages

Almost immediately, the results page was displayed with seven links to “similar” pages. Note that I performed this search with SafeSearch filtering off. The seven results were links to the following web pages:

1. Google Services & Tools ([www.google.com/options/](http://www.google.com/options/))



2. A poor imitation of Google's Advanced Search web page (which I am sure Google does not approve of) (<http://www.alltheweb.com/advanced>)
3. Google Toolbar ([toolbar.google.com/](http://toolbar.google.com/))
4. Something called "Porn Banner Showcase" (I went there to see what in the world porn banners could have in common with Google's Advanced Search web page! I found nothing but a porn site.) ([www.noapologiespress.com/pbs/](http://www.noapologiespress.com/pbs/))
5. A web page for a Virginia college professor who says he is "unavailable for immediate contact" ([http://www.cs.odu.edu/~bayle\\_j/](http://www.cs.odu.edu/~bayle_j/))
6. AltaVista's Advanced Search web page (AltaVista is a service similar to Google) (<http://www.altavista.com/web/adv>)
7. Google's Help Central page ([www.google.com/help/](http://www.google.com/help/))

Search results 1, 2, 3, 6, and 7 are understandable, as they relate in some manner to advanced searching, but results 4 and 5 are totally out in left field. Even though I had to look out of the corner of my eye, I could not see any reason that result number 4 was included in this search result as I slowly navigated the porn site. Result number 5 was just a single web page with a photo and a couple of words about being unavailable. What that has to do with advanced web searching escapes me.

In another test of SafeSearch's ability to filter pornography, I performed the same Similar search as above but this time with the Filter using SafeSearch option on. I received exactly the same results, including the link to the web page that didn't just promote porn banners but was a porn site itself. The home page included the slang word "porn," the word "pornographic," and one "four-letter" word. Looks like SafeSearch missed one that it should have caught. Your results may vary.



## Links

The Links feature allows you to find web pages that link to the web page you specify in the query box. This feature is extremely useful for web page developers who need to discover how many pages on the web link to their web page. Since Google uses this information (called link popularity) as the primary criteria for assigning the rank to a page in its search results, knowing who and how many web sites link to a page that you are responsible for is crucial to getting a higher rank in search results. Business people might use Links to determine who the competition is linking to. In this manner, you might uncover previously unknown business relationships. Using Links to discover information about a site is called “flipping” the site.

If you have a home page, it is also a fun endeavor to enter your domain and search Links to see if anyone found your web page interesting enough to link to it. Otherwise, I do not know what other use it might have, but it’s there if you figure out a need for it. We will look a little closer at links in the next chapter.

## Topic-Specific Searches

### Topic-Specific Searches

New! [Froogle](#) (BETA) - Find products for sale from across the web  
[Catalogs](#) - Search and browse mail-order catalogs online

[Apple Macintosh](#) - Search for all things Mac  
[BSD Unix](#) - Search web pages about the BSD operating system  
[Linux](#) - Search all penguin-friendly pages  
[Microsoft](#) - Search Microsoft-related pages

[U.S. Government](#) - Search all .gov and .mil sites  
[Universities](#): [Stanford](#), [Brown](#), [BYU](#), & more - Narrow your search to a specific school's website

Figure 5.34: Topic-Specific Searches

Topic-Specific Searches allow you to exclusively search the following companies or entities:



- Froogle
- Catalogs
- Apple Macintosh\*
- BSD Unix\*
- Linux\*
- Microsoft
- U.S. Government\*
- Universities\*

Searching a specific domain has the benefit of excluding all other domains from the search results. How can this help you? Let's say you need help with some Microsoft software update issue. By searching Microsoft's domain ([www.microsoft.com](http://www.microsoft.com)) and excluding all others, you can very quickly zero in on your objectives.

The five entities marked with asterisks are also found in Google Special Searches (<http://www.google.com/options/specialsearches.html>). Bookmark it.

Google Special Searches performs exactly the same service as Topic-Specific Searches ([http://www.google.com/advanced\\_search?hl=en](http://www.google.com/advanced_search?hl=en)). Why the duplication? We might as well ask why light is light — it just is!

---

## Summary

---

We have examined Google's advanced search techniques and determined how we can use various ways and means of reducing the scope of our searches to a mild roar (that is, a manageable level). We discovered that some Google tools and features work very well, while others are, shall we say, wounded? As you use these tools, services, features, and operators, bear in mind that Google has only had a few years to work on perfection, and perfection is



unattainable anyway. Deficiencies were pointed out where applicable in the interests of accurate reporting of the facts. I will hazard a guess that Google will work to improve all of its products as time passes. In the meantime, the Internet will continue to grow, and so will your search frustration, unless you use the knowledge that you have gained so far.

This page intentionally left blank.

# Chapter 6

## Advanced Alternate Query Search Operators

### Introduction

---

Google supports several advanced alternate query search operators. These are special query words that the Google search engine recognizes and then responds to accordingly. A discussion of advanced search operators is found at <http://www.google.com/help/operators.html>. (Bookmark it.)

The advanced alternate query search operators are:

- **cache:** Shows the version of the web page residing in Google's cache
- **link:** Lists the web pages that include links to the specified web page
- **related:** Lists web pages that are similar to the specified web page
- **info:** Displays certain information Google has about that web page



- **stock:** The query terms are interpreted as stock ticker symbols, and the search engine returns a web page displaying stock information.
- **site:** The search is restricted to sites within the specified domain.
- **allintitle:** The search is restricted to those web pages containing all of the query words in the title.
- **intitle:** The search is restricted to those web pages containing one or more of the query words in the title.
- **allinurl:** The search is restricted to those web pages containing all of the query words in the URL.
- **inurl:** The search is restricted to those web pages containing one or more of the query words in the URL.

## Advanced Alternate Query Search Operators

---

When using these operators, keep in mind that the syntax, or form, is `xxx:URL`, where `xxx` is the operator and `URL` is the domain name, also known as the web page's URL. For example, a search for links to one of my favorite web pages is `link:www.family-search.org/Eng/Search/frameset_search.asp`. Note that you do not need and should not include the “`http://`” portion of a web address.

### **cache:**

A search result using the `cache:` operator will show the web page version in the Google cache. But what is a cache? Originally, `cache` was a French word meaning “to hide.” Canadian and western frontier explorers and fur trappers adopted the word to mean “a hiding place,” especially “for concealing and preserving provisions or implements.” Today, `cache`, in the computer industry, refers to “a secure, temporary storage place.” So, Google’s cache is the



temporary storage of web pages on secure Google computers.

Storing web pages on Google's computers has several major advantages from our perspective (not Google's) and one great disadvantage. The advantages are:

- **Faster access time to web pages.** Since you are accessing a web page stored on a Google computer, you do not need to bounce around the Internet in multiple hops and electronic switches, gateways, and routers to get to the web page source that might be an overworked, slow computer or server.
- **Availability of web pages.** Google has a 99.9 percent uptime. Most web servers are single-machine enterprises that must go down occasionally for maintenance, either on schedule or as the result of an unplanned event (like a lightning strike). When a server is offline and unavailable, you can still see the last version of any of the server's web pages by viewing Google's cached version.
- **Longevity of web page.** When a web site is removed from the host server, Google's cached version remains available for about another four to six weeks.
- **Security.** Web pages can be infected with Internet flotsam and jetsam (i.e., viruses, trojan horses, worms, and other malicious program code). Google does provide a code-checking function on web pages before they are cached, so you are reasonably assured that the cached web page is free of malicious code.

I discuss the great disadvantage in a moment. But first, here is a note about malicious code: I have spent many, many days searching for and removing malicious code from my computers. These were days that I could have spent playing baseball or fishing with my three sons. I doubt anyone writing such code is reading this material at the moment, but if you are, know this: One day, in computer heaven, I will get the opportunity to torment you as much as you have tormented me in this life. The torment you inflicted lasted days; the torment I inflict will last an eternity. I eagerly look forward to our meeting in the great computer heaven in the sky!



My sincere apologies for the soapbox, but this is a good time to point out that you can get nasty stuff from web pages. In fact, that is where almost all of mine comes from nowadays. I have firewalls that keep out the e-mail-based malicious code (and keep me from sending such viruses onward), but I still get an occasional bug, either from a web page or a newsgroup. Be sure that your virus protection is up to date and *turned on* before you cruise the net. I like to disable mine for certain activities so that the processes run faster. Sometimes I forget to enable it again. Oh, well.

Here is a fact that is little known because universities do not want to take the heat for allowing irresponsible adults access to incredible computing resources on their campuses: Almost all hackers and virus originators are university graduate students. Don't believe me? Don't think anyone is hacking your computer? Buy an inexpensive copy of McAfee's Visual Trace program. You will have a tool to identify who is hacking your computer and the information you need to bust them.

Here is a little (hacking) story. I share it with you in the interest of education. An old man once told me as a young boy that everyone must pay for their education. He said some people pay money for their education in the public schools and colleges of the country, while others pay with blood and sweat in the "school of hard knocks." This is an education I received in the school of hard knocks.

I acquired DSL access several years ago. Almost immediately after I installed the modem, my phone began to ring in the morning. When I picked it up to answer, there was no one on the other end of the line. Assuming a wrong number or crank call, I hung up. Then I began to notice that my computer (at the time not a real fast one anyway) began to run much slower after such phone calls. It took me a couple of weeks to figure out that someone was calling my home and when I hung up, they seized my telephone line and accessed my computer via a software spoof.

To discover who was hacking my computer via the telephone ruse, I bought and installed McAfee's Visual Trace. Imagine my surprise



to discover 23 hackers from 23 different universities in the U.S., Canada, and Australia, all using my computer as a portal to go on to other sites and do who knows what to who knows who at the same time on the same day! With Visual Trace, I was able to identify the domain addresses of the offending hackers and the phone numbers, physical addresses, mail addresses, and names of network/system administrators, whom I contacted. This lesson taught me the wisdom of having an impenetrable firewall.

ZoneAlarm is a free firewall that works extremely well (as long as your kids do not defeat it to play computer games on the Internet or disable it to use file sharing programs such as Kazaa). If you do not have a firewall, I strongly urge you to try a free copy of ZoneAlarm, found at <http://www.zonelabs.com/store/content/company/products/znalm/freeDownload.jsp>. Click on Download FREE ZoneAlarm. Follow the directions to save and install. So maybe you don't have to learn via the "school of hard knocks."

Why all the discussion about hacking and computer security? What does that have to do with Google? Well, it really has nothing to do with Google, but it has everything to do with searching the web. To search the web, you must be connected to the Internet. When you are connected to the Internet, you are vulnerable to hacking and malicious code. Most people think that they only receive viruses via contaminated disks or e-mail. However, a lot of malicious code, not just viruses, trojan horses, and worms, comes from clicking on a link to a web site. What do we do when we search? We click on links to web sites. Keep the firewall up and the virus software up to date!

I mentioned (somewhere way back there!) that caching web pages has a singular disadvantage from a user perspective. The single most significant disadvantage is:

- **The information may be out of date.** Since a Google cached web page is updated only periodically, any changes to the original web page will not appear in the cached version until a Google spider updates the cached version. Since this happens only about once every four to six weeks, as my experience



indicates, the cached web page that you are viewing may not contain the latest information. If the information that you are perusing is not time sensitive, then cached pages can be useful in your searches.

The cached web page has an information and disclaimer box at the top of the page. See Figure 6.1. Unfortunately, the information box does not tell us when the web page was cached. It would be useful to know this information. A page cached just a day or two ago might be acceptable, whereas a page cached eight weeks ago might not be acceptable.

Note that the search term “travel” is highlighted in Figure 6.1.

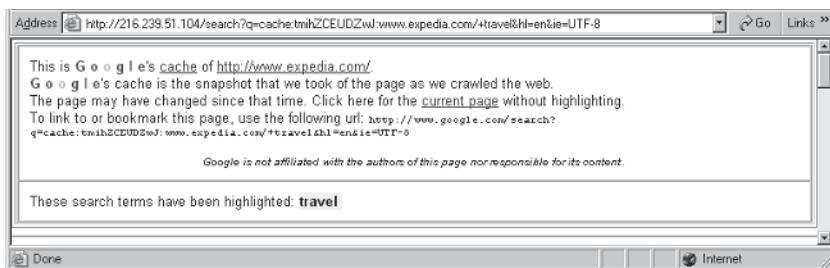


Figure 6.1: Google cached page information box

The information box includes two links to the actual web page; both the domain name and Current page links will take you to the genuine web page. I don't understand the duplicated effort. The information box also includes a link that you can bookmark. If you select this URL to bookmark, know that you are bookmarking the cached page — not the original web page.

Note that a cached page is also accessible by clicking on the Cached link on Google's main results page. See Figure 6.2. The link is highlighted, as are all links, at the bottom of the web page information block.

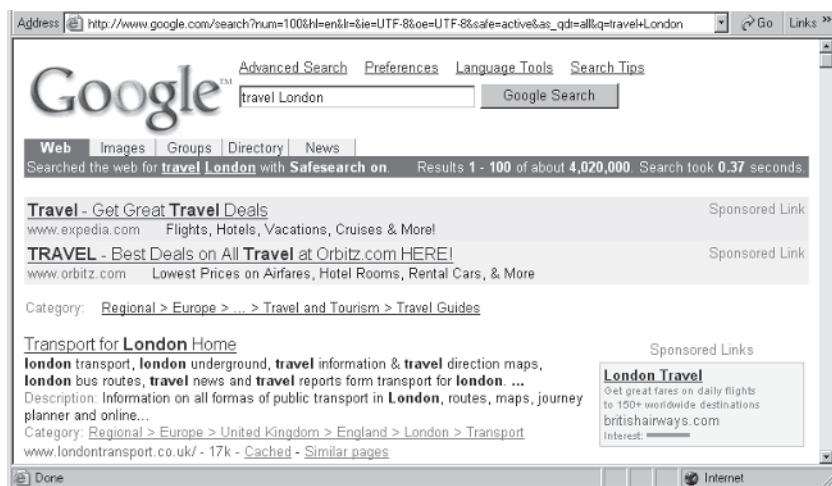


Figure 6.2: Google search results showing cached page link

If other words are included in the query, Google will highlight those words within the body of the cached document. As an example, [cache:www.google.com/help/index.html search] will show Google's cached version of the Google Help Central page with the word "search" highlighted. See Figure 6.3.



Address  Go Links

# Google™ Google Help Central

Home - All About Google

Google does not display pop-up advertising. Here's why.  
Find answers and discuss Google services in our [user support discussion forum](#).

---

<b>Search Help</b>	<ul style="list-style-type: none"><li>• <a href="#">Basics of Search</a></li><li>• <a href="#">Interpreting Results</a></li><li>• <a href="#">Customize Results</a></li></ul>	<ul style="list-style-type: none"><li>• <a href="#">Advanced Search Tips</a></li></ul>	
<b>Web Search Features</b>	<ul style="list-style-type: none"><li>• <a href="#">Cached Links</a></li><li>• <a href="#">Dictionary Definitions</a></li><li>• <a href="#">File Types</a></li><li>• <a href="#">I'm Feeling Lucky</a></li></ul>	<ul style="list-style-type: none"><li>• <a href="#">News Headlines</a></li><li>• <a href="#">PhoneBook</a></li><li>• <a href="#">Similar Pages</a></li><li>• <a href="#">Site Search</a></li><li>• <a href="#">Spell Checker</a></li></ul>	<ul style="list-style-type: none"><li>• <a href="#">Stock Quotes</a></li><li>• <a href="#">Street Maps</a></li><li>• <a href="#">Web Page Translation</a></li><li>• <a href="#">Who links to you?</a></li></ul>
<b>Google Services &amp; Tools</b>	<ul style="list-style-type: none"><li>• <a href="#">Google Answers</a></li><li>• <a href="#">Google Browser Buttons</a></li><li>• <a href="#">Google Catalogs</a></li><li>• <a href="#">Google Groups</a></li><li>• <a href="#">Google Image Search</a></li></ul>	<ul style="list-style-type: none"><li>• <a href="#">Google in your Language</a></li><li>• <a href="#">Google News</a></li><li>• <a href="#">Google Special Searches</a></li><li>• <a href="#">Google Toolbar</a></li><li>• <a href="#">Google Translate Tool</a></li></ul>	<ul style="list-style-type: none"><li>• <a href="#">Google University Search</a></li><li>• <a href="#">Google Web APIs</a></li><li>• <a href="#">Google Web Directory</a></li><li>• <a href="#">Google Web Search</a></li><li>• <a href="#">Google Wireless</a></li></ul>
<b>Contact Us</b>	<ul style="list-style-type: none"><li>• <a href="#">Google Search</a>: suggestions, comments, and questions about our web <a href="#">search</a></li><li>• <a href="#">Corporate Google</a>: inquiries regarding press and business services</li><li>• <a href="#">Google Groups</a>: feedback about our Usenet discussion <a href="#">search</a></li><li>• <a href="#">Google Web Directory</a>: feedback about our categorized web directory</li><li>• The Google Toolbar: feedback about our browser add-on <a href="#">search tool</a></li></ul>		
<b>Search Our Site</b>	<input type="text"/>	<input type="button" value="Search the Google site"/>	

---

©2003 Google - [Home](#) - [All About Google](#) - [We're Hiring](#) - [Site Map](#)

Figure 6.3: Google cached page with “search” highlighted

Note that the information box at the top of the cached page includes a message that the term “search” is highlighted within the body of the cached page. See Figure 6.4.

Figure 6.4: Google cached page information box with “search” highlighted

You can combine multiple words (up to ten), and Google will highlight each in the cached web page in a different color. See Figure 6.5. It illustrates highlighting more than one word. The query I used for the figure was [cache:[www.google.com/help/index.htm](http://www.google.com/help/index.htm) links definitions types lucky headlines pages search checker quotes].



maps Translation Spell]. Notice that the last two words, Translation and Spell, do not show up in the list of highlighted words in the information box, and they are not highlighted in the body of the cached document.

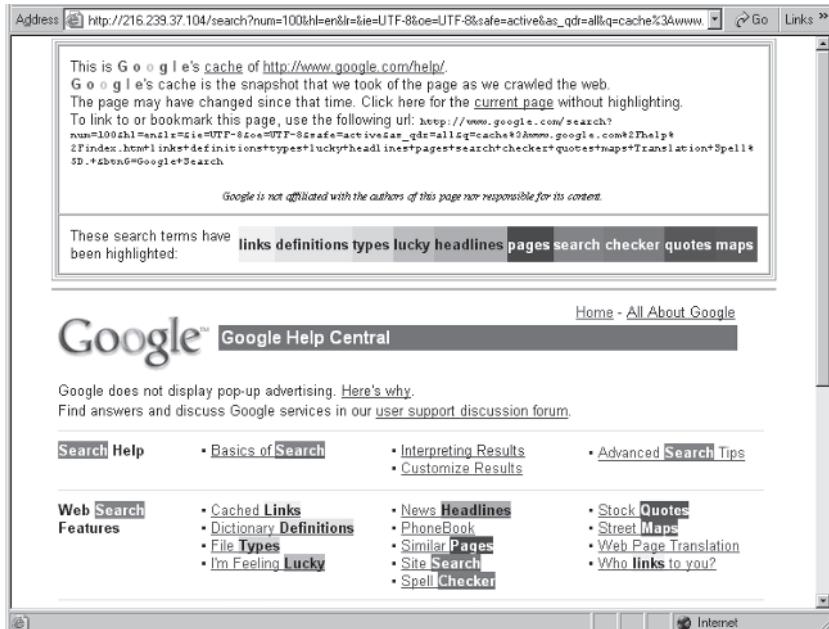


Figure 6.5: Google cached page with multiple search terms highlighted

When placing one or more terms after the domain name in the cache: query, keep these points in mind:

- The results are case sensitive.
- There must be a space between each word.
- Do not place more than ten words in the query after the domain name.
- It does no good to place a term before the word “cache” or anywhere within the domain name.

Points to remember about the cache: query:

- You can perform a cache query from any Google query/search box.



- cache: queries return potentially out-of-date web pages.
- If a web page is unavailable from its web server, try the Google cache.
- You can view a cached page from the regular search result links.

## link:

We briefly mentioned links in the previous chapter. You may recall that the Links advanced search feature allows you to find web pages that link to the web page you specify in the query box. The format for a link: query is just like that for cache:, [link:www.yahoo.com]. Let's see what we get if we perform a link: query on www.yahoo.com (Yahoo's home page).

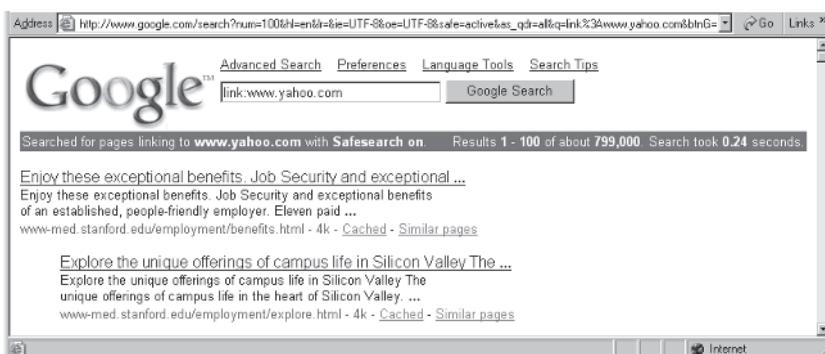


Figure 6.6: link: search results

The search results page of a link search appears exactly as any other search result page. See Figure 6.6. Note the links to a cached page and similar pages. We discussed cached pages above and similar pages in the previous chapter.

Here is a (big?) secret. Google does not say anything about placing words after the link: query, but I tried it to see what would happen. For the record, I tested [link:www.yahoo.com campus]. Two things (both good) happened. The search engine excluded any web pages without the word, and it highlighted the word in the search results page but not in the web page. If you try this, note that the words placed after the domain name must be separated by a space. The



search results that I observed after a few trials showed that the Google search engine treats all words after the domain name as if they are connected via the + (plus) symbol. That is, only links to web pages containing all of the words were returned in the search results page.

I imagine that you can use other operators as part of a link: query and modify your search accordingly. I will leave it up to you to try a few and see what you get.

Things to remember about a link: query:

- A link: query works in any Google query/search box.
- You can use search words in a link: query if you place them after the domain name, separated by spaces.
- A link: query tells you how many web pages (not necessarily web sites) are linked to the domain name used in the query. The number of “hits,” or search results, provides this information.
- If you have a business arrangement with web sites to link to your site, you can easily verify those links.

## **related:**

related: is similar to the Similar feature. It returns links to web pages similar to the domain name of the domain in the related: query. Recall that we discussed similar pages in the previous chapter. The difference is that the Similar feature is a link on a search results web page, whereas related: is a query. The format of a related: query is the same as a cache: or link: query. Let's see what we get when we do a related: search for Yahoo's home page, [related:www.yahoo.com]. See Figure 6.7. Looks like a familiar search result page with all of the usual information there. However, there is a new link to the Yahoo site that we have not seen in results before — a link to a web page where you can get a Yahoo stock quote.

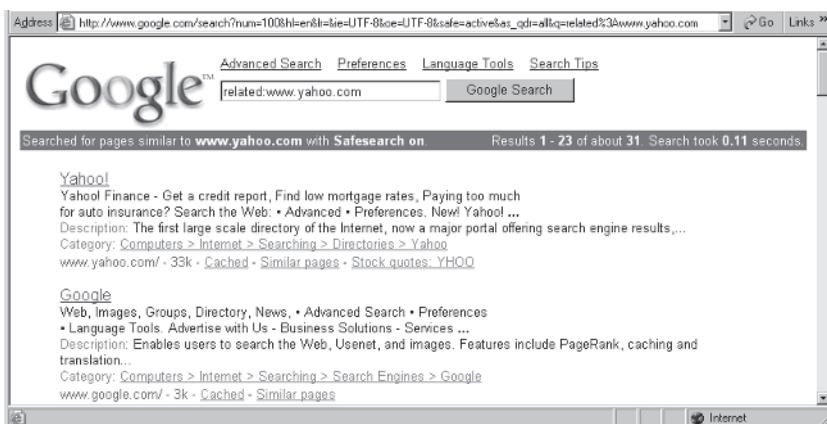


Figure 6.7: related: search results

What good is a related: query? Well, if you are looking for companies that manufacture and/or sell similar products, a related: query works amazing well. Let's see what we get when we search related:schwinn.com. Schwinn manufactures and sells bicycles. We received 31 web site links to companies that make bicycles. See Figure 6.8.

Try a related: query on other company domain names, including the company you work for.

Searched for pages similar to [www.schwinn.com](http://www.schwinn.com) with Safesearch on. Results 1 - 30 of about 31. Search took 0.55 seconds.

Figure 6.8: related:schwinn search results

Things to remember about a related: query:

- A related: query works in any Google query/search box.
- You cannot use any keywords after the domain name in a related: query.



## info:

An info: query returns links to the domain name included in the query along with the first 24 or so words in the web page's title or upper page information. See Figure 6.9. The format for an info: query is the same as a related: query, [info:www.yahoo.com].

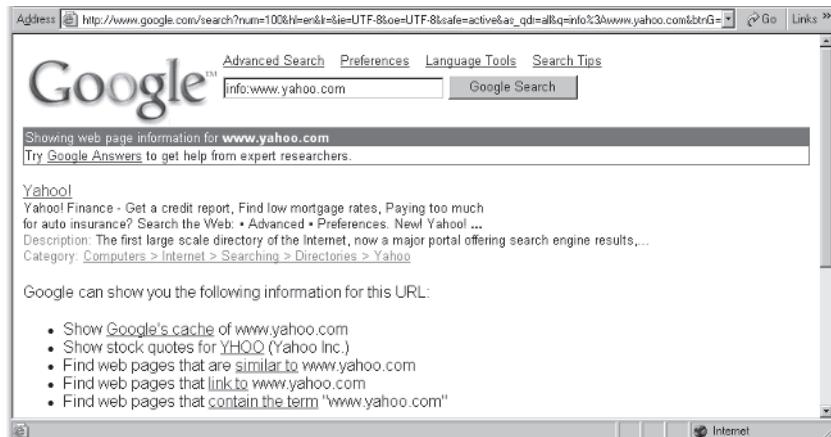


Figure 6.9: info: query

Just like a related: query, you can place keywords after the domain name in an info: query, and the Google search engine will exclude any links of web pages that do not include those keywords. Try an info: query on big corporations like Ford, GM, Lockheed, and your company.

Things to remember about an info: query:

- An info: query works in any Google query/search box.
- You can use up to ten keywords in an info: query if you place them after the domain name, separated by spaces.



## stock:

stock: treats all text after [stock:] as stock ticker symbols. The format for a stock: query is the same as a related: query, [info:www.yahoo.com]. For example, [stock:yhoo] will return a link where you can retrieve Yahoo's stock price. See Figure 6.10. Note that stock: also returns search results with links to pages that include both "stock" and "yhoo."

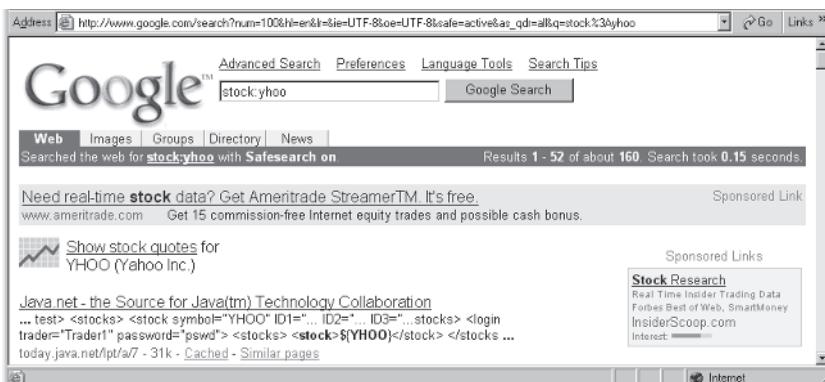


Figure 6.10: stock: query

You can include up to ten stock symbols in the query. The query will also work from any Google query box. You can query without the [stock:] and then click on the stock quote links contained in the search results page.

Things to remember about a stock: query:

- A stock: query works in any Google query/search box.
- You can use up to ten stock symbols in a stock: query if you place them after the query [stock:], separated by spaces.



## site:

The site: query will restrict search results to the specified web pages in the domain specified. The form of the site: query is [xxx site:TLD/domain name] or [site:TLD/domain name xxx], where xxx is one or more keywords. You can search up to ten keywords. The keywords after the query can appear anywhere within the document. Note that you can search on either a TLD (top-level domain, such as .com, .edu, .gov, etc.) or a domain name (such as www.my\_domain\_name.edu). For example, to search Google's Help page, our query would look like [help site:www.google.com]. The results are shown in Figure 6.11.

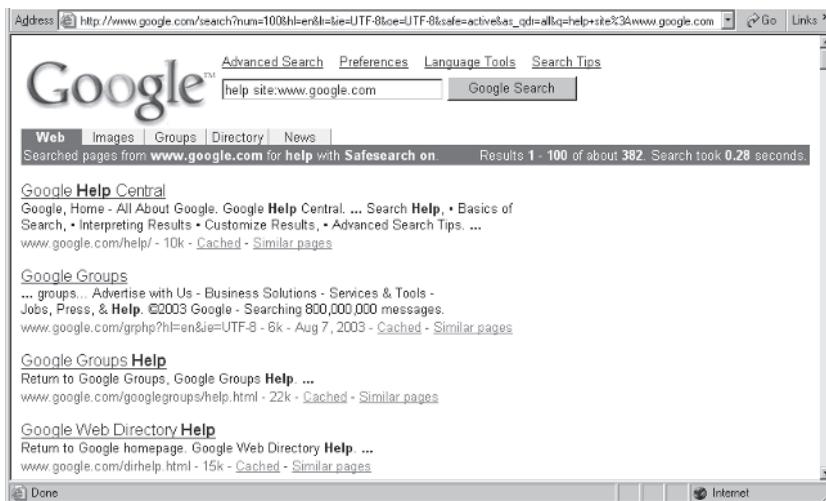


Figure 6.11: Google Help site: query

Note that the first search result in the search result page is Google Help Central. Since we searched Google's site for [help site:www.google.com] "help," it seems reasonable that we would acquire Google's Help Central as a search result. If we test this one more time by performing the same search on Yahoo's site, we should receive the same results. Let's search on [help site: www.yahoo.com].

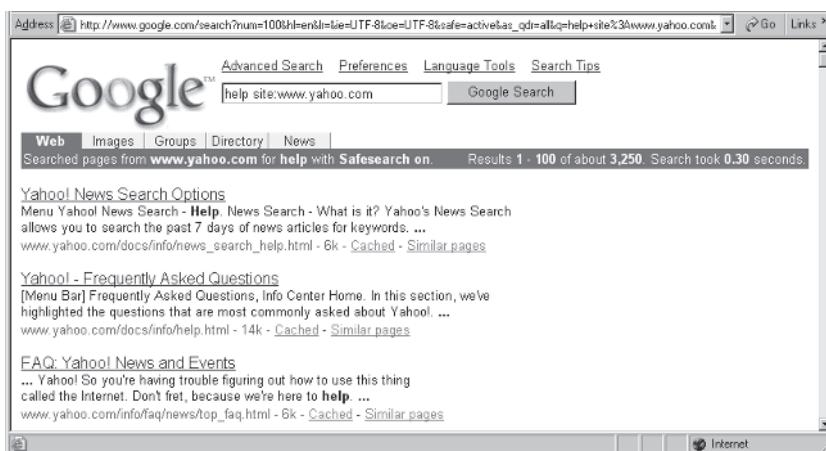


Figure 6.12: Searching for Yahoo Help site

Yahoo has a web page titled Help Central just like Google does. It stands the test of logic that if the site: query worked in a certain way for one web site, it would work exactly the same way for another. In this case, the site: query obviously does not return the same result for both Google and Yahoo. Methinks there is something fishy going on here.

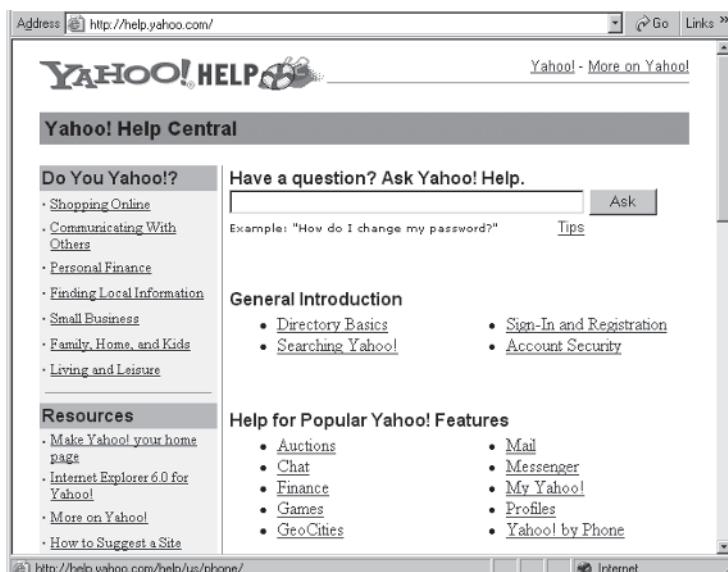


Figure 6.13: Yahoo Help site



Whatever is going on here, this example serves to illustrate that some Google products do not work consistently. Just keep that in the back of your mind as you ponder from time to time the seemingly odd results that you will see.

Remember in Chapter 1 that I said people search the Internet for career information? You can easily zero in on the best career web sites with a site: search. A site: search for engineering jobs would look like this: site:.com job engineer. It gives all sites pertaining to jobs for engineers. See Figure 6.14 for the search results.

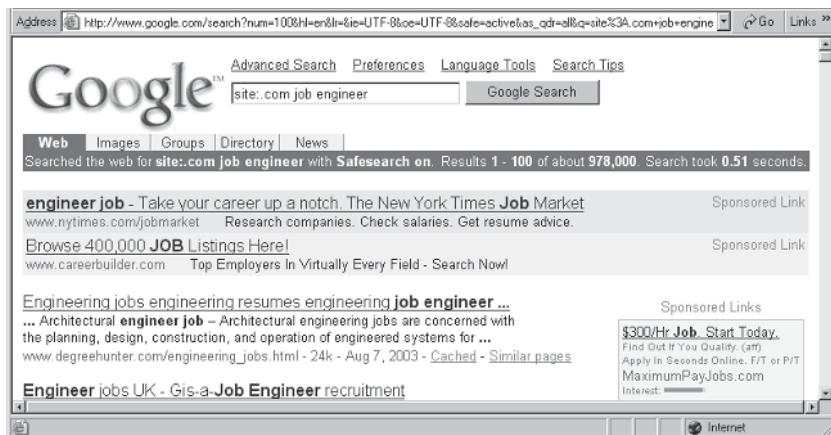


Figure 6.14: site: query for .com job engineer

Try site:.com job programmer java -coffee. This site: search will list links to web pages pertaining to Java programming jobs and will exclude any site with the word coffee (coffee = java). Try other professions, including your own. You can be more specific by adding additional keywords describing the profession. A search for a Unix software engineering position would look like site:.com job “Unix software engineer” (with the quotes).



#### Note:

It does not seem to make any difference if the keywords come before “site” or after the domain. That is, “job engineer site:.com” works exactly the same as “site:.com job engineer”.



Things to remember about a site: query:

- A site: query works in any Google query/search box.
- You can use up to ten keywords in a site: query if you place them after the domain name, separated by spaces.
- The site: query functionality is also present in the Advanced Search page. See Figure 6.15.

The screenshot shows the Google Advanced Search interface. At the top, there is a dropdown menu labeled "Domain". Below it, a search bar contains the text "Only" followed by a dropdown arrow, with the placeholder "return results from the site or domain" next to it. To the right of the search bar is a text input field containing "e.g. google.com, .org" and a link labeled "More info".

Figure 6.15: Site: query functionality in Advanced Search page

- Note that you can use site: to search any TLD. TLDs are .edu, .com, .gov, .org, etc., or any web site.

## allintitle:

The allintitle: query will restrict search results to links to those web pages containing all of the keywords in the title of the page. The form of the allintitle: query is [allintitle:keyword xxx], where xxx is one or more keywords, up to a maximum of ten. All of the keywords appear in the title since the operator is “all in the title.” For example, to search for resumes of engineers other than civil engineers, our query would look like [allintitle:resume engineer -civil -post]. We exclude civil (-civil) in the query since we are not interested in civil engineers, at least in this example. Also, we want to exclude those commercial sites seeking resumes to post, so we use “-post”. The search results are shown in Figure 6.16. You can include up to ten words in the query. All words in the query must appear in the title for links to those pages to appear in the search results.

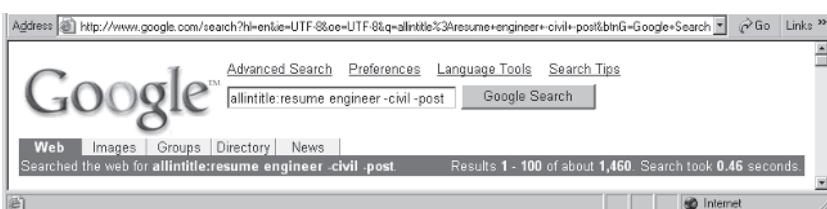


Figure 6.16: allintitle:resume engineer -civil -post search results



Note the number of web pages (1,460) containing the keywords “resume” and “engineer” and excluding “civil” and “post” shown in Figure 6.16. If you are seeking employees, isn’t this a wonderful way to find prospective candidates?

Things to remember about an allintitle: query:

- An allintitle: query works in any Google query/search box.
- You can include up to ten keywords in an allintitle: query.
- The allintitle: query functionality is also present in the Advanced Search page. See Figure 6.17.



Figure 6.17: *allintitle: query functionality in Advanced Search page*

## **intitle:**

The intitle: query restricts search results to links of those web pages containing the first word after the intitle: query in the title of the page. The form of the intitle: query is [intitle:keyword xxx], where xxx is one or more keywords. You can search up to nine additional keywords for a maximum of ten. The second and additional words after the query are keywords that can appear anywhere within the document, including the title. For example, let’s look for the latest research on liver cancer. The form of the search is [intitle:cancer “latest research” liver]. The results are shown in Figure 6.18.

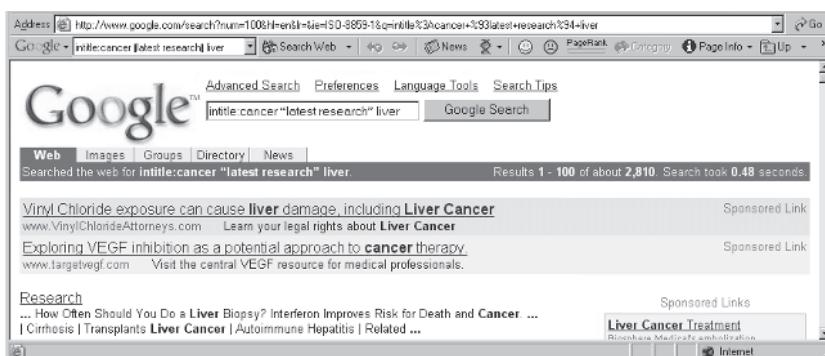


Figure 6.18: intitle:cancer “latest research” liver search results

The power of an intitle: search is graphically demonstrated in Figure 6.18. We just captured the links to 2,810 web pages describing the latest research material for liver cancer.

Another interesting intitle: search is [intitle:linux “user groups” member members membership]. See Figure 6.19 for the search results.

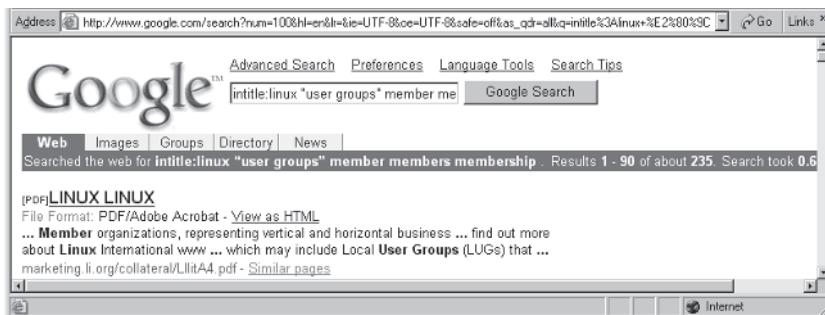


Figure 6.19: intitle:linux “user groups” member members membership search results

Things to remember about an intitle: query:

- The intitle: query works in any Google query/search box.
- You can include up to ten keywords in an intitle: query. Only the first keyword after the query must appear in the title, unless you include the first word with others surrounded by



quotes, while the others may appear anywhere within the web page.

- Placing intitle: in front of every keyword is the same as searching with allintitle:.
- The intitle: query functionality is also present in the Advanced Search page. See Figure 6.17.

## allinurl:

The allinurl: query will restrict search results to links of those web pages containing all of the keywords in the URL of the page. The form of the allinurl: query is [allinurl:keyword]. For example, to search for Google help in the URL of web pages, our query would look like [allinurl:google help]. The results are shown in Figure 6.20. You can include up to ten words in the query. All words in the query must appear in the URL for links to those pages to appear in the search results. Note that in Figure 6.20 the search results shown each include “Google” and “help” in their URL (domain name).

The screenshot shows a Microsoft Internet Explorer window with the following details:

- Address Bar:** Http://www.google.com/search?num=100&hl=en&lr=&ie=UTF-8&oe=UTF-8&safe=off&as\_qd=allq+allinurl%3AGoogle+Help+&cgt=Go+Links>
- Google Logo:** The classic Google logo is at the top left.
- Search Bar:** The query "allinurl:Google Help" is entered in the search bar.
- Buttons:** "Google Search" button is visible.
- Navigation Links:** "Advanced Search", "Preferences", "Language Tools", "Search Tips".
- Toolbar:** "Web", "Images", "Groups", "Directory", "News".
- Search Results:**
  - Google Help Central:** [Google, Home - All About Google. Google Help Central. Google does not display pop-up advertising. Here's why. Find answers and discuss ...](http://www.google.com/help/)
  - Google Help:** [Return to Google homepage. Advanced Search ... Additionally, Google supports several advanced operators which are query words that have special meaning to Google. ...](http://www.google.com/help/refinerearch.html)
  - Google Web Search Features:** [Google Web Search Features. Home All About Google Help Central Google Features Services & Tools Our Technology Find on this site: Google Web Search Features ...](http://news.google.ca/help/features.html)
  - Google Toolbar Help Index:** [Google Toolbar Help Index. Frequently Asked Questions See common questions about the Google Toolbar. Toolbar Buttons Each button ...](http://toolbar.google.com/help/html)
- Sponsored Links:** "Google T-shirts", "Show the world how you find it", "Google Pens, Lava Lamps, Bean Bags", "www.GoogleStore.com", "Interest: [redacted]".
- Message:** "See your message here..."
- Bottom Status Bar:** Shows "Internet" and other standard browser icons.

Figure 6.20: allinurl:Google Help search results



An allinurl: query does not operate on any punctuation, so characters usually found with URLs, such as the slash (/), will not contribute to the search results.

Things to remember about an allinurl: query:

- An allinurl: query works in any Google query/search box.
- You can include up to ten keywords in an allinurl: query.
- The allinurl: query functionality is also present in the Advanced Search page. See Figure 6.21.

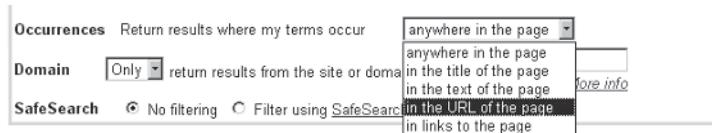


Figure 6.21: allinurl: query functionality in Advanced Search page

## inurl:

The inurl: query will restrict search results to links to those web pages containing the first word after the inurl: query in the URL of the page. The form of the inurl: query is [inurl:keyword]. You can search up to nine additional keywords in the inurl: query. The second and additional words after the query are keywords that can appear anywhere within the document, including the URL. For example, to search for Google help (again) in the URL of web pages, our query would look like [inurl:google help]. The results are shown in Figure 6.22.



The screenshot shows a Google search results page. The address bar indicates the query: `http://www.google.com/search?num=100&hl=en&lr=&ie=UTF-8&oe=UTF-8&safe=off&as_qdr=all&q=inurl%3AGoogle+Help&btnG=1`. The search term `inurl:Google Help` is entered in the search bar, and the button `Google Search` is highlighted. The results page displays 100 out of approximately 1,300,000 results, completed in 0.47 seconds. The first result is **Google Help Central**, which links to `http://www.google.com/help/`. This page includes sections like **About Google**, **Google Web Search Features**, and **Google help - Constructing and refining searches in Google**. A sidebar on the right titled **Sponsored Links** features a box for **Google T-shirts** with a link to `www.GoogleStore.com`. The browser interface at the bottom shows standard window controls and the word `Internet`.

Figure 6.22: `inurl:Google Help` search results

Another `inurl:` search example, using travel as the subject, is shown in Figure 6.23. We received 39,300 results, which seems extraordinary given we are searching the URL for travel.

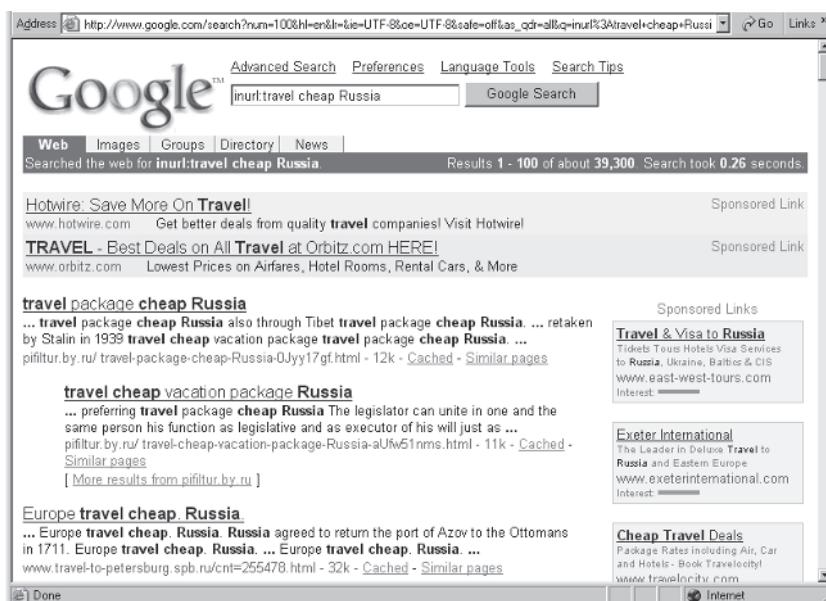


Figure 6.23: *inurl:travel cheap Russia* search results

Let's exchange the position of "travel" and "Russia" and see if we can reduce the number of search results.

After exchanging "Russia" and "travel," our search results are encouraging. We now have only 3,380 results to choose from (keeping in mind Google only displays the first 1,000). See Figure 6.24.



Figure 6.24: *inurl:Russia cheap travel* search results

So, let's try one more time and see if we can get even fewer results. We will enclose "cheap travel" in quotes. Now we're cookin'! We have 112 high-quality search results. See Figure 6.25.



Figure 6.25: *inurl:Russia "cheap travel"* search results



Still too many? Then add more keywords to narrow the scope of the search. Maybe we are interested in traveling to a certain part of Russia. How about a visit to Moscow?



Figure 6.26: *inurl:Russia "cheap travel" Moscow* search results

Figure 6.26 shows that we received 85 search results when we narrowed the scope of our search to Moscow. We can handle 85 top-quality search results, can't we? I think so.

Things to remember about an *inurl:* query:

- An *inurl:* query works in any Google query/search box.
- You can include up to ten keywords in an *inurl:* query. Only the first keyword after the query must appear in the URL, while the others may appear anywhere within the web page.
- Placing *inurl:* in front of every keyword is the same as searching with *allinurl:*.

In some of the example searches that I used in this chapter, I excluded certain words to illustrate how you can narrow the search field and bring the quality of the search results within reasonable limits. Of course, I just picked these examples because they were illustrative of the point I was making. In real life, you may or may not know for any particular search which words to include and which ones to exclude.

Here is the way to approach the issue. When you search, examine the first 20 or so search results. If most of them are not what you are looking for, then look for words common to most or all of the results in which you are not interested. Select those words as keywords and exclude them in the next search.

As an example, say we wanted to find the resumes of Java programmers. We could do a search on [allintitle:programmer java]. The results we get back could include links to web pages whose subject matter concerned coffee, since java is a synonym for coffee. A more focused search would be [allintitle:programmer java -coffee]. But,



you might not think about coffee until after you have searched on [allintitle:programmer java]. If you received a lot of search results with links to pages about coffee, you might wonder how you can eliminate them from the results. Bingo! Search again with the “-coffee.” Some experimentation may be necessary, but you can quickly sort out the path to search nirvana with a little effort and by thinking about your search goals and examining your search results. With some experience, you will realize great satisfaction with your ability to find what you seek.

## Summary

---

In this chapter we examined Google’s advanced search operators. We discovered that these operators are very useful for confining our searches to specific domains, domain names, and certain portions of web pages. By clearly defining the search goal, identifying the proper keywords (both those we want to include and those we wish to exclude), and then deftly selecting our search operator, we can mine the web for information and hit the mother lode!

# Chapter 7

## Advanced Image Search

### Introduction

---

In Chapter 2, we discussed basic image searching. In this chapter we delve deeper into the features that Google offers in the Advanced Image Search page. These features permit you to base your image search upon various criteria that exclude nonconforming images. The ability to exclude images allows you to bring your search objectives into specific relief, thereby reducing the amount of time required to find a particular image meeting your needs. You can access Advanced Image Search at [http://www.google.com/advanced\\_image\\_search?hl=en](http://www.google.com/advanced_image_search?hl=en) or by selecting Images on the Google home page and then left-clicking the Advanced Image Search link. (Bookmark it!) See Figure 7.1.

Images (and web pages, too) are copyrighted material. While Google provides a means for us to search the web and find, display, and save images, Google does not have the legal authority to grant permission to use an image in any manner. If you want to use an image, you must contact the owner of the image and receive his or her written permission.



I must reiterate that if you search the web for images, you will be exposed to material that is defined as adult or mature material. Regardless of the precautions you take, if you perform any number of image searches, you will (sooner or later) come across material that you may find objectionable. For those of you who take no precautions, it will be sooner rather than later. In the interest of fair play, I have included Google's warning about mature/adult material here:

**Warning:** The results you see with this feature may contain mature content. Google considers a number of factors when determining whether an image is relevant to your search request. Because these methods are not entirely foolproof, it's possible some inappropriate pictures may be included among the images you see. (The mature content filter is only available from an English interface.)

A note of caution. If you have a slow computer or a computer that does not have sufficient memory and/or you are using a 56KB dial-up modem, I would recommend that you forget about serious image searching, unless you have plenty of time on your hands and you are a patient person.

What is sufficient memory? That depends upon the speed of your computer and the speed of your (cable/DSL?) modem. Here is a way to test if your equipment is new enough for speeding down the image interstate highway. Go to the Advanced Image Search web page and enter the two keywords "flower" and "pot" in the top query box. Left-click on the Google Search button and see what happens. If you run out of patience before the results page loads, then you should press **Ctrl+Alt+Delete** simultaneously. This key sequence will bring up the Close Program dialog. In the dialog, you should see something like "Google Search: "flower pot" – your browser's name." Click on End Task, and in a moment you will regain control of your computer. (I wish it were that easy with my teenage boys!) If you must perform this task more than once per day because you became impatient, then it is time to get a new computer or a new cable/DSL modem, as appropriate!



Advanced Image Search results are displayed 20 per results page. They are arranged four per row, and there are five rows. See Figure 7.3. The maximum number of search results is the usual “about” 1,000. Unlike the Advanced Search page where you can set the number of results per results page to as many as 100, there is no ability to change the number of results per page in Advanced Image Search. Also, the Advanced Image Search query boxes (just like all other query boxes) are not case sensitive.

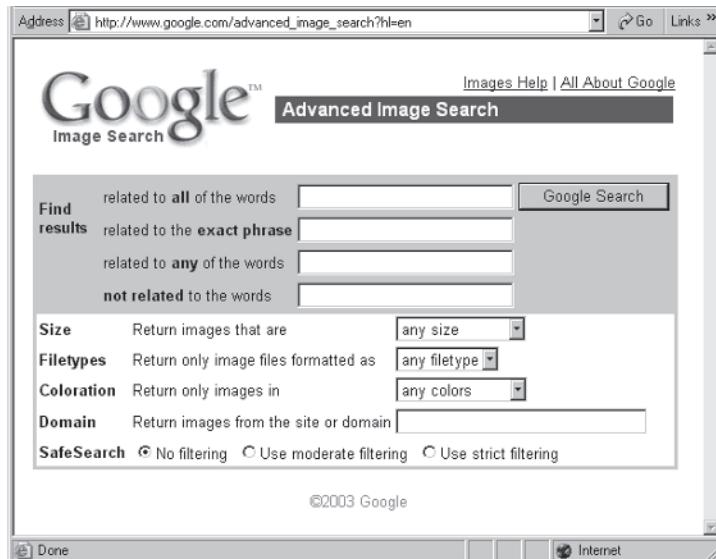


Figure 7.1: Google Advanced Image Search page

Advanced Image Search has some features similar to Google’s Advanced Search, while other features are different. The Advanced Image Search features are:

- **Find Results:** Find Results includes the image search box and three additional boxes that allow you to focus your search using specific criteria.
- **Size:** Size allows you to choose a particular image size. You have seven sizes to choose from.
- **Filetypes:** Filetypes gives you control over the types of image files to search for.



- **Coloration:** Coloration permits you to choose four different image color schemes.
- **Domain:** Domain lets you choose a specific domain to search.
- **SafeSearch:** SafeSearch provides some control over the adult content.

Find Results is similar to Advanced Search Find Results. Find Results includes four query boxes that allow you to specify keywords related to your search objectives. All of the operators that we have discussed in this book can be used in conjunction with keywords to focus your search.

The maximum number of keywords per query box is the usual ten. A detailed description of the Advanced Image Search features follows in the next section.

## Advanced Image Search

---

Here we begin our exploration of the Advanced Image Search application. Note that there is a help link in the top-right corner of the web page. Next to the Images Help link is an All About Google link. You should left-click on both of these in turn, so you are familiar with the material in the web pages of the two links. After you have visited those two web pages, look at the Advanced Image Search query boxes and options in closer detail.

### Find Results

Find Results is the heart of the Advanced Image Search features. Find Results consists of four query boxes, each with a special function to govern or control some aspect of the search engine. The Advanced Image Search Find Results works just like the Advanced Search Find Results feature, except in the case of the Advanced Image Search version, the keywords and search results refer to



images and image titles, not documents. Find Results is shown in Figure 7.2.

Figure 7.2: Advanced Image Search Find Results



### Note:

A word about the use of the word “title”: When I am referring to the text associated with an image’s label, I will refer to the image’s name. I use the word “title” to refer to the title of a web page.

The method utilized to search for images is simple. You enter in the keywords describing what image you seek in the Related to **all** of the words query box, place keywords in the remaining query boxes as appropriate, select the remaining advanced options of interest, and then left-click on the Google Search button. See Figure 7.2.

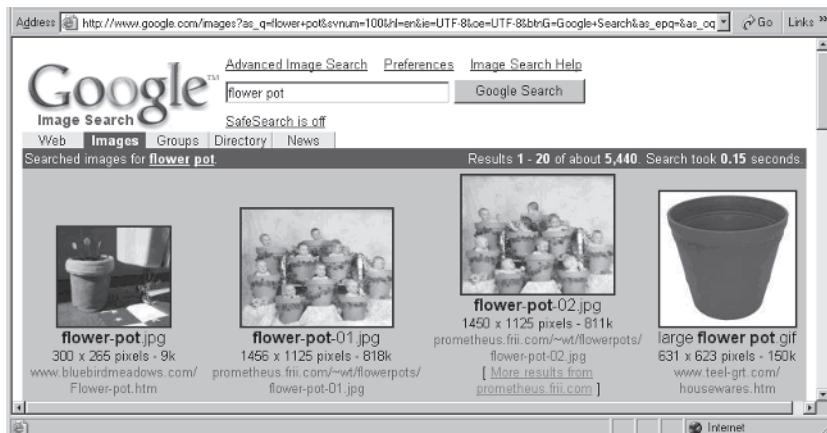


Figure 7.3: Google Advanced Image Search results page for flower pot



Before we actually search for an image, let's discuss the Find Results query boxes. Then we use what we have learned to illustrate other Advanced Image Search options. There are four query boxes shown in Figure 7.2.

When an Image Search results page is displayed (such as the one shown in Figure 7.3), the images you see are called *thumbnails*. A true thumbnail is smaller than the original image. You can view the full-size image by left-clicking on it. Besides displaying a full-size image, Google also displays the web page where the image is located below the full-size image. Understand that the web page you are viewing is "inside" the Google page. If you want to view the actual web page, Google provides a link to the page below the full-size image.

## Related to All of the Words

The Related to **all** of the words box applies the plus operator function between the keywords that you list in the box. It is the primary query box. If we enter the keywords "flower pot" without the quotes, we see the results in Figure 7.4. We should not be surprised to see that the first images are of flower pots!



Figure 7.4: Google Advanced Image Search results page for flower pot



If we examine the search results in closer detail, we discover that there are some odd ducks (images) included in the results. For the first 94 or so results, both the keywords “flower” and “pot” appear in the image name. Then we have an image whose name contains the word “flower” but not “pot.” In Figure 7.4, observe the image on the first row, third from the left. It is named flower.gif. All of the previous 94 images included “flower” and “pot” in their name. Why not this one? Before I answer this question, let’s see if Google uses PageRank to rank image search results.

Examining the web page where this image comes from reveals that the web page itself has the keywords “flower pot” in the body of the web page. So that explains why this image is included in the results. The search engine not only looks at image names, but it also looks at the parts of the web page where an image is found.

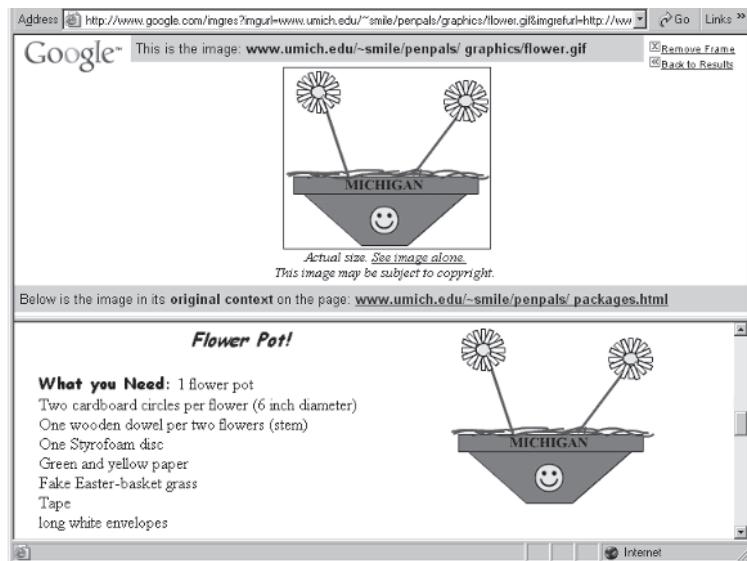


Figure 7.5: Google Advanced Image Search results page for flower pot



Google claims it employs PageRank to rank search results without any distinction made between image and web page searches. Do we have a tool to determine if PageRank is involved in image results ranking? Of course. Recall that we have examined the Links operator. Let's use it to see how several of our image search results measure against each other. To perform the check, I copied the domain name of search result 95 into the Advanced Search query box and performed a link search. The results I received were:

`http://www.umich.edu/~smile/penpals/packages.html link  
results = 2 links`

Next, I went back to the first image and performed the same link search. We would expect (if Google used PageRank to rank image search results) that the first image would have at least two links. Right? Well, the result I received this time was unexpected. The number of links for the first image is:

`www.bluebirdmeadows.com/Flower-pot.htm link results = 1  
link`

The web page where the first image resides has one less link than the 95th ranked image web page. So, it appears that PageRank is not used to rank image search results. Maybe keyword occurrence is the ranking factor. I visited each web page and counted the number of keyword occurrences. The results are:

First search result:

Number of occurrences of the keywords in the web page:

`flower = 4, pot = 4`

Ninety-fifth search result:

Number of occurrences of the keywords in the web page:

`flower = 5, pot = 5`

Well, obviously keyword occurrence is not the criteria used to rank the results, since the number one result has two total keywords less than #95. But maybe a test sample of two results is not



enough. With that in mind, I looked at search result #41. Search result #41 had neither flower nor pot in its name, and the number of links were:

[www.bluemud.com/dept.asp?dept\\_if=127&toc\\_id=127](http://www.bluemud.com/dept.asp?dept_if=127&toc_id=127) link  
results =212 links

Number of occurrences of the keywords in the web page:

flower = 4, pot = 4

If the results were ranked by the number of links, then clearly #41 with its 212 links should be before #1 and #95. If the number of occurrences was the criteria, then it should be ranked closer to #1. What gives? I have no clue. Repeated tests did not reveal any ranking scheme that I could fathom. So how does Google determine image search results ranking? From other limited experiments, it appears to me there is no scheme. I would hazard a guess that they are ranked based upon the date and time (a typical computer sorting algorithm) that the Google spider found the image and indexed it in the Google database.

What does this mean for us, the users of this search service? I think it means that if we wish to maximize our search results so we minimize the amount of time we spend searching, we must make use of the additional search options available to us in the Advanced Image Search page. The narrower we focus our search, the less time we will spend searching for the mother lode. Onward to those additional options!

Points to remember:

- The query box is not case sensitive.
- There can be a maximum of ten words in the query box.
- You can use other operators on keywords, such as quotes, in this query box.



## Related to the Exact Phrase

The search engine operates on keywords in the Related to the **exact phrase** query box as if the words had quotes around them. For example, the keywords “in” and “bloom” would appear to the search engine as “in bloom.” You can get the same search result if you placed the keywords in the Related to **all** of the words query box in this manner: [flower “in bloom”]. So, why the need for another query box? It is easier to see the query box and remember what its purpose is (especially since there is a little explanatory text to go with it) than it is to remember [flower “in bloom”]. Will you remember the “short form” six months after you finish this book?

This query box operates on the words even if the Related to **all** of the words query box is empty. To illustrate, I searched on the keyword “flower” in the top query box (Related to **all** of the words) and “in bloom” (without the quotes) in the Related to the **exact phrase** box. The search results are shown in Figure 7.6. Notice how the search words are displayed in the search results query box (flower “in bloom”). Somewhere, either in the image name or the web page title or within the body of the web page, you will find the phrase “in bloom.”

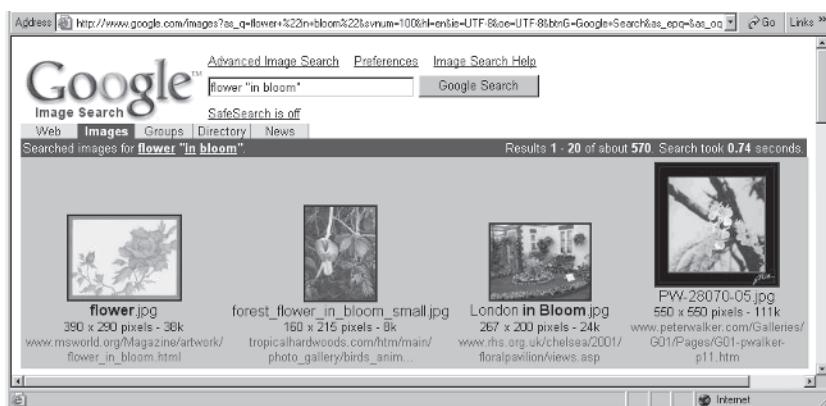


Figure 7.6: Google Advanced Image Search results page for flower “in bloom”



Let's see what happens if we search with "in bloom" in the query box without the associated "flower" in the Related to **all** of the words query box. See Figure 7.7. The search results are images of flowers that are in bloom, but the results are not the same results that we received when "flower" was included in the search.

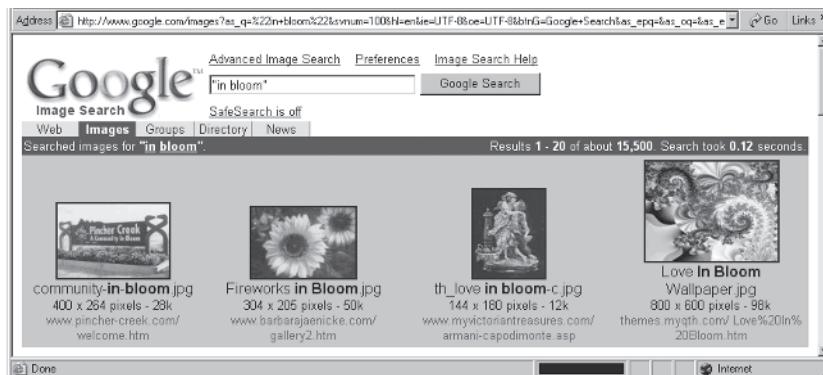


Figure 7.7: Google Advanced Image Search results page for "in bloom"

### Points to remember:

- The query box is not case sensitive.
- There can be a maximum of ten words in the query box.
- Keywords in this query box appear to the search engine as if they are surrounded by quotes in a normal query box.
- If the other query boxes are empty, the search engine will still operate upon the words in this query box.



## Related to Any of the Words

The search engine treats keywords in the Related to **any** of the words box as if there is an “OR” between each one. For example, the keywords “in” and “bloom” would be treated not as “in bloom” but rather as “in” OR “bloom.” I searched on “flower” in the top query box and “in bloom” in the Related to **any** of the words query box. See Figure 7.8. The results that I received are shown in Figure 7.9.

The screenshot shows the Google Advanced Image Search interface. In the 'Find results' section, there are four input fields:

- related to **all** of the words: flower
- related to the **exact phrase**: [empty]
- related to **any** of the words: in bloom
- not related to the words: [empty]

A 'Google Search' button is located to the right of the input fields.

Figure 7.8: Google Advanced Image Search results page for flower in OR bloom

Observe the search terms in the results page query box in Figure 7.9. Google put the “OR” between “in” and “bloom.” That is about the sum of this Advanced Image Search feature. As usual, the query box is not case sensitive and you are limited to ten keywords.

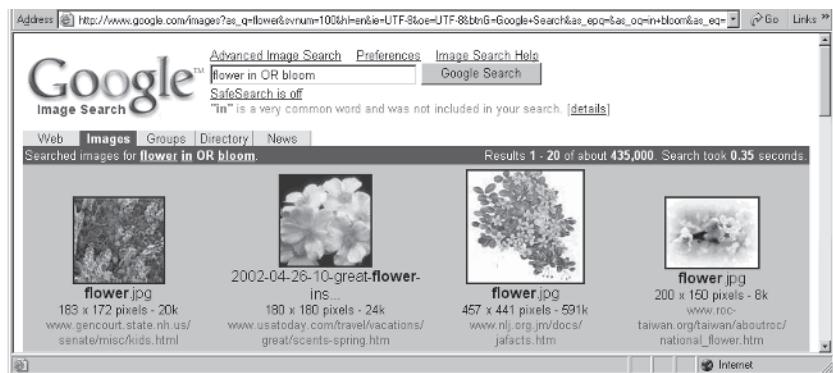


Figure 7.9: Google Advanced Image Search results page for flower in OR bloom

Points to remember:

- The query box is not case sensitive.
- There is a maximum of ten words in the query box.



- Keywords in this query box appear to the search engine as if they are separated by the OR operator.
- If the other query boxes are empty, the search engine will still operate upon the words in this query box.

## Not Related to the Words

Keywords in the **Not related** to the words query box tell the search engine to ignore images titled or otherwise associated in some manner with the keyword(s). The search engine views the keywords in this query box as if they are preceded by a minus (-) operator. You may recall from Chapter 4 that the minus operator performs the Boolean function NOT.

To check this feature out, I searched on “flower” with “in bloom” in the **Not related** to the words query box. See Figure 7.10.

The screenshot shows the Google Advanced Image Search interface. On the left, there's a sidebar with the heading "Find results" and four options: "related to all of the words", "related to the exact phrase", "related to any of the words", and "not related to the words". The "not related to the words" option is selected, indicated by a blue border around its input field. To the right of the sidebar, there are two input fields: one for "flower" and another for "in bloom". Both fields have a blue border, indicating they are active. Below these fields is a "Google Search" button.

Figure 7.10: Google Advanced Image Search results page for flower NOT in bloom

The search results are shown in Figure 7.11. I opened a few of the web pages and I did not find the words “in” or “bloom” anywhere within the web page. Observe the search results query box. Google placed the minus sign (-) in front of the two query words that were in the **Not related** to the words query box.

Points to remember:

- The query box is not case sensitive.
- There can be a maximum of ten words in the query box.
- Keywords in this query box appear to the search engine as if they are separated by the minus (-) operator.
- If the other query boxes are empty, the search engine will still operate upon the words in this query box.



Figure 7.11: Google Advanced Image Search results page for flower NOT in bloom

## Size

The Size feature allows you to select any of six different image sizes. Or you can select all of them. Left-click on the Size dropdown menu and then move the mouse icon over the size of interest to highlight it. Then left-click to select it. See Figure 7.12.

<b>Size</b>	Return images that are	any size
<b>Filetypes</b>	Return only image files formatted as	any size
<b>Coloration</b>	Return only images in	icon-sized
<b>Domain</b>	Return images from the site or domain	small
<b>SafeSearch</b>	<input type="radio"/> No filtering <input type="radio"/> Use moderate filtering <input checked="" type="radio"/> Use strict	medium

Figure 7.12: Google Advanced Image Search Size feature

The various sizes of images are:

### Icon size



Figure 7.13: Google Advanced Image Search icon size



### Small



Figure 7.14: Google Advanced Image Search size small

### Medium



Figure 7.15: Google Advanced Image Search size medium

### Large



Figure 7.16: Google Advanced Image Search size large

### Very large



Figure 7.17: Google Advanced Image Search size very large

### Wallpaper size



Figure 7.18: Google Advanced Image Search wallpaper size



If you experience slow load times when trying to search images, try reducing the image size. The difference in time between a wallpaper size image and an icon image can be lengthy on a slow computer and/or slow modem.

## Filetypes

Filetypes gives you the option of searching only for specific image file types or for all file types. The file types are the most common used today — jpg, gif, and png. To make a selection, left-click on the Filetypes drop-down menu, place the mouse cursor over the item of choice, and then left-click.

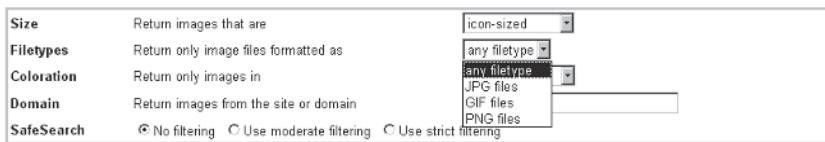


Figure 7.19: Google Advanced Image Search Filetypes feature

That just about covers it. Simple, eh? If it is so simple, then why the file type choices? Long ago, when dinosaurs roamed the planet (circa 1995), image file formats were proprietary formats, and user programs usually could only deal with one type of image format. So, if you had this or that software program and it was a single format program, you would only be interested in the one format supported by your software. Perhaps there are many people still using old software today that can process only one format. Image software today can process all image formats. I am up to date, hardware and software wise. I can handle any image file format. Whoopee!



## Coloration

Coloration gives you the option of selecting any one of three image color formats (or all of them). The three formats are black and white, grayscale, and full color. To make a selection, left-click on the Coloration drop-down menu, place the mouse cursor over the item of choice, and then left-click.

The screenshot shows the 'Coloration' dropdown menu open. It contains four options: 'any colors' (selected), 'any colors', 'black and white', 'grayscale', and 'full color'. Other dropdown menus show 'icon-sized' and 'any filetype'.

Size	Return images that are icon-sized
Filetypes	Return only image files formatted as any filetype
Coloration	Return only images in any colors (selected)
Domain	Return images from the site or domain
SafeSearch	<input type="radio"/> No filtering <input checked="" type="radio"/> Use moderate filtering <input type="radio"/> Use strict filtering

Figure 7.20: Google Advanced Image Search Coloration feature

The Coloration choices are pretty clear. If you do not know what any of the three selections are, then select that color choice and search for images. Why the three choices? Speed and load times. It takes a lot less time to load a black and white image than a grayscale image, and a lot less time to load a grayscale image than a full-color image. If you experience very slow load times when searching Any colors or Full color, try changing to Grayscale or Black and white.

## Domain

The Domain box is just like the Domain box for Advanced Search for web pages. See Chapter 5.

The screenshot shows a 'Domain' input field with the placeholder text 'Return images from the site or domain'.

Figure 7.21: Google Advanced Image Search Domain feature

Entering a domain in the box will restrict the search to that domain only.



## SafeSearch

Advanced Image Search SafeSearch is identical to Advanced Search SafeSearch for web pages. See Chapter 5.



Figure 7.22: Google Advanced Image Search SafeSearch feature

For the adventurous, searching the web can be fun. But you just never know what you may be exposed to.

## Summary

---

Searching for images is a very enjoyable pastime if you have the equipment and Internet connection to do it. Search on keywords used in your favorite sports, hobbies, and work. You will find a great assortment of beautiful and pleasing images to behold. However, do not forget that the images are copyrighted and cannot be used in any manner without the copyright owner's permission. Also, remember that the laws of the United States specify that whom-ever holds the negative (or original artwork in the case of graphics) holds the copyright.

You can use complex search operators in the Advanced Image Search query boxes, just as you could use them in Advanced Search for web pages. You can use all of the operators. For example, the plus sign (+), the minus sign (-), double quotes (""), link:, cache:, etc., work equally well in the Advanced Image Search query boxes.

You are now armed with sufficient search skills to find just about anything on the Internet that you desire without any undue frustration. You could say that your (search the web) education is complete. However, the next chapter has a few tricks and treats waiting for you. So, let's go see what there is to offer an expert web searcher.

---

# Chapter 8

## The Google Toolbar

### Introduction

---

In this chapter, we examine the Google Toolbar and its associated features and functions. The Google Toolbar is a browser-installable toolbar convenient for quickly searching the web. When installed, the toolbar is integrated with the web browser. See Figure 8.1.



Figure 8.1: Google Toolbar

The first publicly available version of the Google Toolbar was released in late 2000. It has undergone a number of revisions, and at the time of publication was in version 2.0. We first examine the 1.1.x version of the Google Toolbar and then discuss the newest features in version 2.0.



## System Requirements

---

Currently, your system must meet the following system requirements to install the toolbar. Note that the only browser the toolbar is compatible with is Internet Explorer. Sorry, Netscape fans! The minimum system requirements to take advantage of the Google Toolbar are:

- Microsoft Windows 95/98/ME/NT/2000/XP operating system
- Microsoft Internet Explorer version 5 or later web browser (the pop-up blocker available in version 2.0 requires IE 5.5)

## Google Toolbar Privacy Policy

---

Before we install the Google Toolbar, we should take the time to review Google's privacy policy with respect to the installation and use of the toolbar. The Toolbar Privacy Policy web page is located at <http://toolbar.google.com/privacy.html>. (Bookmark it, as you will find virtually no reference to it anywhere else.)

There are two important issues to note regarding the privacy policy. The first is that Google checks the version of your toolbar from time to time and automatically updates it if there is a newer version available. If a firewall is installed on your computer, the toolbar updates may be blocked. Additionally, depending on your security settings, you may see a dialog box requesting confirmation that you want to update the toolbar.

Programs that automatically update your computer assume that you want to get updated, but updates can cause compatibility problems. Suddenly, bad things start happening with your computer, such as freeze-ups, and you don't know why. That can cost money, if you are not technical. I like control over my computer, so I know what is going on and who is involved. I decide if I want to get updated or



not. Then, if there is a compatibility issue, I know it, and I know which piece of software is causing the problem.

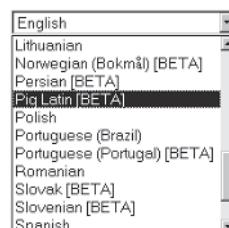
The second important issue is the transfer of information to Google regarding the web page that you are viewing. This is necessary for Google to provide you with certain information. This issue is only of concern if you have advanced toolbar features enabled. We discuss what advanced toolbar features must be enabled and how to enable them in the appropriate sections. For now, just be aware that your search habits can be transmitted to Google if you enable the advanced features. The default condition of the advanced features when you install the toolbar is disabled. Now let's install the toolbar.

## Toolbar Installation

The Google Toolbar is free. It costs you nothing but a little time to install it. The following instructions will help you get the toolbar installed in your browser.

To use the toolbar, you must install it in your browser. Journey to the toolbar home page at <http://toolbar.google.com/> to install the toolbar. Left-click on the toolbar button  at the bottom of the page. Next, you are presented with a page where you are asked if you agree to the terms of use. If you do, left-click on the  button. Your toolbar will be installed automatically. Now, you are ready to use the handy toolbar to search the web, unless you want to change the language that the toolbar is displayed in.

The toolbar is available in 37 languages. Select the language you prefer from the toolbar page at <http://toolbar.google.com/>. Left-click on the drop-down box down arrow and select the language you prefer by left-clicking on it. In the example shown, I chose Pig Latin. Note that





some languages may have [BETA] beside the language. The [BETA] sign warns you that the browser interface of that language may not be thoroughly checked out and may cause problems.

If you do not want to view the toolbar all of the time, you do not have to uninstall it. You can turn the toolbar on/off as appropriate. To turn the toolbar on or off, left-click on the browser's View | Toolbars menu selection. See Figure 8.2. Google will appear somewhere in the list of toolbars, if it is installed. When you left-click on Google, you turn the toolbar off, if it is on; if the toolbar is off, you turn it on. A check mark adjacent to Google signifies that the toolbar is in the on, or display, mode.

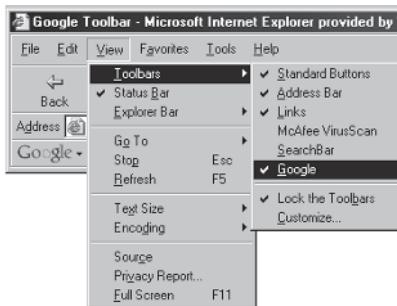


Figure 8.2: Displaying the Google Toolbar

Now that we have the toolbar installed (in a language we understand), let's check out some of its features.

## Toolbar Help

---

You will find Toolbar Help at [http://toolbar.google.com/button\\_help.html](http://toolbar.google.com/button_help.html). (Bookmark it!)



## Drag-and-Drop Functionality

A really neat feature of the Google Toolbar is its drag-and-drop functionality. You can drag and drop any text or URL from the current page you are viewing onto the Google Toolbar in order to perform a search. It is simple. Just highlight the desired text, drag it to the toolbar, and then drop it anywhere on the toolbar. The Google Toolbar will then perform a search on the search terms or the URL.

## Right-click Functionality

Another really neat toolbar feature for searching the web is its right-click functionality. Highlight any text on the current page, right-click the mouse, and then select Google Search from the pop-up menu.

## Toolbar Features

The toolbar includes these features:

- **Google Toolbar Menu:** Provides selection of additional toolbar features
- **Google Search:** Gateway to Google's search technology from your browser
- **Search Site:** Limit the search range to only the pages of the site that you're currently visiting
- **PageRank:** View Google's ranking of the current page
- **Page Info:** View additional information about the current page, including similar pages, pages that link back to the current



page, and a cached snapshot of the page, and translate the page to English

- **Up:** Changes the view to the web page next “up” in the site’s hierarchy
- **Highlight:** Highlights the search terms wherever they appear on the current page; individual query word is shown in its own color

To use any of these toolbar features, just left-click on the one of interest. Some of the features are disabled until you have met the appropriate conditions for them to be enabled. As you size the browser window and/or add or remove toolbar buttons, you may or may not see the » toolbar continuation button on the far-right side of the toolbar. If you see this button (after installing the toolbar, I had to close the browser window and reopen it to see the continuation button), you can left-click on it, and more enabled toolbar buttons will be visible. Of course, if a toolbar button has not been enabled for use (as opposed to a button that is enabled but is “grayed out” because the functionality is not enabled at the moment) on the toolbar, it will not appear on the toolbar. Let’s now look at each feature in depth.

## Google Toolbar Menu

Google provides a menu of additional toolbar features. You can access the toolbar menu by left-clicking on the down arrow adjacent to Google on the toolbar itself . See Figure 8.12 later in this chapter to view the toolbar menu. We examine these menu selections in some detail in later sections.

## Google Search

Google Search is the gateway to Google’s search technology from the convenience of your browser. It consists of a query box that can contain more than 700 characters. See Figure 8.3. To key words into the search box, the blinking cursor must be in the box. If there is no blinking cursor in the box, place the mouse icon in the box and



left-click once. After you enter your keywords in the search box, press the Enter key on the keyboard to initiate the search.



Figure 8.3: Toolbar search box

The search box contains a history box that you can view and select items from by left-clicking on the down arrow located on the right side of the box.

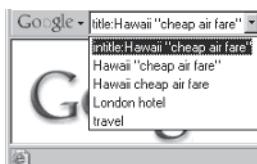


Figure 8.4: Search box with history

After you left-click on the down arrow, move the mouse cursor over the selection of interest and left-click once. The search is initiated as soon as you click.

Keyboard shortcuts:

- Press Alt+G to move the keyboard focus to the search box.
- Hold the Shift key down to display the search results in a new window.

## Search Site

You can limit the search to only the pages of the site that you're currently visiting. You must have a page from a web site open for this toolbar button to be active; otherwise, it is “grayed out.”



Figure 8.5: Search Site feature



## PageRank

If you left-click on PageRank, Google's ranking of the current web page will be displayed.

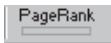


Figure 8.6: PageRank feature

PageRank is disabled until you visit Google's Privacy web page at `res://c:\windows\googletoolbar.dll/prefetch.html`. (Don't bother to bookmark it. Copy and paste the hyperlink just as it is — do not place "http://" at the beginning of the link.) Left-click on the Enable All Advanced Features button at the bottom of the page. Then left-click on the OK just below the Enable... button. PageRank must be enabled because whenever you use this feature, the web page URL that you are viewing is sent to Google so Google's servers can identify the web page and match it with its page rank. Then, the page rank is returned to your computer for display.

Google provides an explanation regarding the reasons that PageRank must send information regarding your search to Google's servers on the Privacy Information page. Google also explains that no personal information is collected in this process. The privacy policy is located at <http://toolbar.google.com/privacy.html>. I strongly suggest that you read it before you install the toolbar.

The manner used to display the page ranking is shown in Figures 8.7 and 8.8. The ranking bar is solid green for the site ranked highest (Figure 8.7). As the site ranking decreases, the bar displays less green and more white (Figure 8.8).



Figure 8.7: Google PageRank



Figure 8.8: Yahoo PageRank

**Note:**

If you select all of the toolbar options, the PageRank button will not be displayed on the toolbar.

## Page Info

Page Info provides a drop-down menu to view additional information about the current page including similar pages, pages that link back to the current page, a cached snapshot of the page, and a link to translate the page to English. See Figure 8.9. We have discussed these features in some detail in previous chapters. They are the same.

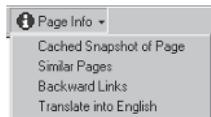


Figure 8.9: Page Info feature

To view the menu, just left-click on the down arrow adjacent to Page Info. Move the mouse cursor over the menu selection of interest and then left-click.

## Up

Up changes the current page to the web page that is next “up” in the hierarchy of the currently viewed web page.



Figure 8.10: Up feature

## Highlight

Highlight causes all of the search terms to be highlighted wherever they appear on the current page. See Figure 8.11. Each individual search word is shown in a different color. For example, the search term “flower” is highlighted like this:

Flower



Figure 8.11: Highlight feature

We have finished our examination of the toolbar features. Next, we look at the toolbar menu.

## Toolbar Menu

---

You can customize the toolbar to meet your search needs. Once you have the toolbar installed, left-click on the down arrow adjacent to the word Google. A drop-down menu will appear.



Figure 8.12: Toolbar menu

The toolbar menu includes these items:

- **Google Home Page:** Loads Google's home page in the browser (<http://www.google.com/webhp?sourceid=navclient>)
- **Google Images:** Loads the Google Image Search page in the browser (<http://www.google.com/imghp>)



- **Google Groups:** Loads the Google Groups page in the browser (<http://www.google.com/grphp>)
- **Google Web Directory:** Loads the Google Web Directory page in the browser (<http://www.google.com/dirhp>)
- **Google News:** Loads the Google News page in the browser (<http://news.google.com/>)
- **Google Answers:** Loads the Google Answers page in the browser (<http://answers.google.com/answers/main>)
- **Advanced Search Page:** Loads Google's Advanced Search page in the browser ([http://www.google.com/advanced\\_search](http://www.google.com/advanced_search))
- **Search Preferences Page:** Loads Google's Preferences page in the browser (<http://www.google.com/preferences>)
- **Language Tools Page:** Loads Google's Language Tools page in the browser ([http://www.google.com/language\\_tools](http://www.google.com/language_tools))
- **Toolbar Options:** Allows you to customize the buttons in your toolbar (res://c:\windows\googletoolbar.dll/opts.html — Copy and paste link as is. Do not add “[http://](#)” to the beginning of the link.)
- **Clear Search History:** Clears the toolbar search history
- **Help:** Loads the Google Toolbar Help page in the browser (<http://toolbar.google.com/help.html>)
- **Privacy Information:** Loads the Google Toolbar Privacy Policy page in the browser (res://c:\windows\googletoolbar.dll/prefetch.html — Do not add “[http://](#).”)
- **About Google Toolbar:** Loads Google's About Toolbar dialog
- **Contact Us:** Loads Google's Toolbar Support page in the browser (<http://toolbar.google.com/feedback.html>)
- **Uninstall:** Loads the Google Toolbar Uninstall page in the browser (res://c:\windows\googletoolbar.dll/uninstall.html — Do not add “[http://](#).”)



## Google Home Page

A left-click on Google Home Page loads Google's home page in the browser. The home page is located at <http://www.google.com/webhp?sourceid=navclient> (should already be bookmarked).

## Google Images

A left-click on Google Images loads the Google Image Search page in the browser. The Image Search page is located at <http://www.google.com/imghp> (should already be bookmarked).

## Google Groups

A left-click on Google Groups loads the Google Groups page in the browser. Google Groups is located at <http://www.google.com/grphp> (should already be bookmarked).

## Google Web Directory

A left-click on Google Web Directory loads the Google Web Directory page in the browser. Google Web Directory is located at <http://www.google.com/dirhp> (should already be bookmarked).

## Google News

A left-click on Google News loads the Google News page in the browser. Google News is located at <http://news.google.com/> (should already be bookmarked).

## Google Answers

A left-click on Google Answers loads the Google Answers page in the browser. Google Answers is located at <http://answers.google.com/answers/main> (should already be bookmarked).



## Advanced Search Page

A left-click on Advanced Search Page loads the Google Advanced Search page in the browser. The Google Advanced Search page is located at [http://www.google.com/advanced\\_search](http://www.google.com/advanced_search) (should already be bookmarked).

## Search Preferences Page

A left-click on Search Preferences Page loads Google's Preferences page in the browser (<http://www.google.com/preferences>).

## Language Tools Page

A left-click on Language Tools Page loads the Language Tools page in the browser. Language Tools is located at [http://www.google.com/language\\_tools](http://www.google.com/language_tools) (should already be bookmarked).

## Toolbar Options

A left-click on Toolbar Options allows you to customize the buttons in your toolbar. The Toolbar Options web page is located at <res://c:\windows\googletoolbar.dll\opts.html>. (Bookmark it!) It is worth a trip to this web page to check out and install a few of the options. Add options by left-clicking in the check box and placing a check mark there. You can remove options by left-clicking on the check mark. After you have made your selections, save the options by left-clicking on OK at the bottom of the page. Notice that at the bottom of the page, just above OK, there are two buttons that allow you to choose if you want to install the advanced features. We discussed the advanced features earlier in the book. Make the appropriate choice by left-clicking on the appropriate button before left-clicking on OK. We review the options in the following sections.



## General Options

There are three options to choose from in this category. The first option selects the Google site from which to base your searches. The second option allows a new browser window to open with every search. The third option adds menu items to the browser's right-click menu.

Users in the United States will want to select [www.google.com](http://www.google.com) most of the time. Why? Speed! [www.google.com](http://www.google.com) is located in the United States. Those other Google search sites are located around the world. To select the Google search site of your choice, left-click on the drop-down menu and then scroll down until you find the site you prefer. See Figure 8.13.



Figure 8.13: Google search location

You can elect to open a new browser window for each search you perform. See Figure 8.14. The plus side about this choice is the ability to keep the original search page and any other pages open while perusing the current page. The bad thing about it is the amount of memory that each new page consumes. If you have an older machine or a machine with less than 512MB of RAM memory, I would suggest leaving this option off.



Figure 8.14: New browser window and right-click menu options



You can add Google stuff to the browser's right-click menu, as shown in Figure 8.14. I find this option very useful and keep it enabled at all times. I recommend you do also. See Figure 8.15 for a view of the Google items in my right-click menu. You may or may not have the same Google features, depending upon what you have enabled. The Google menu items that I have are:

- Cached Snapshot of Page
- Similar Pages
- Backward Links
- Translate Page

Each of these features were discussed in previous chapters, so I will not describe them here.

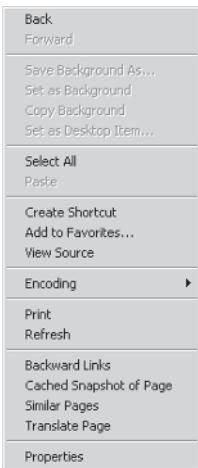


Figure 8.15: Browser right-click menu

The changes to the browser's menu will not take place until you close all browser windows and then start the browser again.

## Search Box

The search box option includes four choices. The first choice lets you select various search box sizes. The second choice enables the search box to keep a history of the query words for which you searched. The third choice allows you to keep the search history



between browser sessions. The fourth choice turns on automatic searching when you click on a search word from the history list.

The search box size choice allows you to specify if the search box will be:



When left-clicking on the buttons to select the width that you prefer, notice that the search box in the browser changes after you make a selection. The change is immediate. Obviously, as you increase the length of the search box, the amount of space available on the toolbar to display other buttons decreases.

The search box second choice enables or disables a drop-down search history list. This list is convenient if you search the same query words repeatedly and you want to keep the search words between browser sessions. Otherwise, it can be a nuisance. See the “Google Search” section above for information regarding the use of this feature. Enable this option by left-clicking in the check box. A check mark signifies that it is enabled.



Figure 8.16: Search box with history

The fourth search box option is a fast way to initiate a search and save a mouse click. As soon as you select a keyword from the list, Google initiates a search. No additional clicking is required. Enable this option by left-clicking in the check box. A check mark signifies that it is enabled.



## Buttons

You can choose three options for displaying toolbar button icons with or without text labels. The choices are All text, Selective text only, and No text. All text will place a text label beside icons that have a text descriptor, Selective text only will place a text label beside some icons, and No text will remove all icon text labels. See Figure 8.17. These changes are immediate, so you can left-click on the three different selections, observe the changes in your toolbar, and decide which choice is best for you.

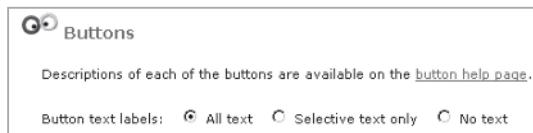


Figure 8.17: Button text labels

## Search Buttons

Search buttons enable several Google search features to be present in the toolbar. We have already discussed these features in previous chapters.



Figure 8.18: Enabling Search buttons

- **Search Web button:** Search the web
- **Search Site button:** Search only the current web page
- **I'm Feeling Lucky search button:** Let Google decide for you
- **Search Images button:** Search the web using Google Image Search



- **Search Groups button:** Search Google Groups
- **Search Directory button:** Search Google's Web Directory

We have already discussed these features in previous chapters. I recommend adding as many of these search features to your toolbar as you can stand. It sure makes searching the web a lot easier. Enable these options by left-clicking in the check box of your choice. A check mark signifies that it is enabled.

## Advanced Features Are ENABLED

The two advanced features are the PageRank  and Category  buttons. We already discussed advanced features in another section. This is just another place where you can choose to enable or disable them. Enable these options by left-clicking in the check box of your choice. A check mark signifies that it is enabled.

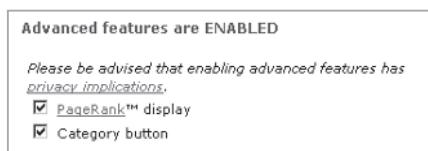


Figure 8.19: Enabling Advanced Features buttons

## Page Information

There are two Page Information choices. See Figure 8.20. One allows you to enable or disable the Page Info menu. The other enables or disables voting buttons.

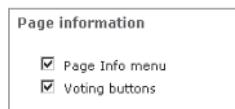


Figure 8.20: Page Information buttons



When you left-click on the Page Info button after it is installed on the toolbar, the drop-down menu displays four options, as shown in Figure 8.21. We have previously discussed these options.



Figure 8.21: Page Info options

Voting buttons allow you to vote for or against a site. The voting buttons are applicable to the web page that you are currently viewing. If you like a site, left-click on the yellow icon, and if you dislike a site, left-click on the blue icon. Enable these options by left-clicking in the check box of your choice. A check mark signifies that it is enabled.

## Navigation

You have two button choices in Navigation. The first installs a Google News button , and the second installs an Up button . The first takes you to the Google News web page, and the second ratchets you up the hierarchy of web pages. Both of these features are discussed elsewhere. Enable these options by left-clicking in the check box of your choice. A check mark signifies that it is enabled.



Figure 8.22: Navigation buttons

## Finding Words within a Page

You have two button choices in this section. The first, , highlights search words, each word with a different color. The second, , installs buttons corresponding to the search



words that you use. I searched on “intitle:Hawaii” so I was awarded an intitle:Hawaii toolbar button for future one-click searching on “intitle:Hawaii.” This feature could become cumbersome if you have too many search terms. Enable these options by left-clicking in the check box of your choice. A check mark signifies that it is enabled.



Figure 8.23: Finding words within a page buttons

## Google Compute Is ENABLED

This choice enables or disables the Google Compute button . We discussed Google Compute in Chapter 2. Enable this option by left-clicking in the check box. A check mark signifies that it is enabled.

## Experimental Features

For the adventurous toolbar user, three experimental features are available. The three features are Combined Search button, Browser Control, and Navigation. The Combined Search feature allows you to compact the toolbar somewhat by adding one button with a menu of Google web sites. We have already discussed these sites and their functions. Note that the Combined Search button is identical to the regular search button, except the Combined Search button has a drop-down arrow adjacent to Search Web.



Figure 8.24 Combined Search button and menu



After you enable the combined button, you may want to disable or remove any duplicate buttons. You can also choose to keep the last search performed with Combined Search as the default.

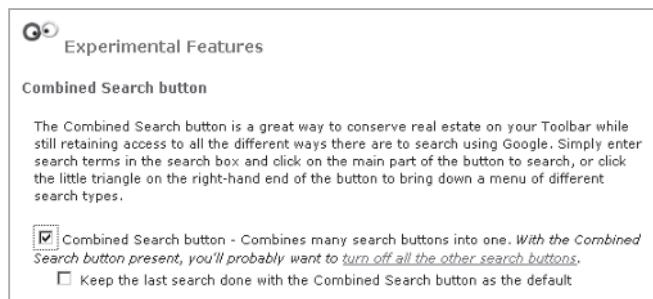


Figure 8.25: Enabling the Combined Search button

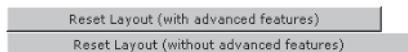
Browser Control is not a button on your menu but rather a line of software code that tells your browser to ignore pop-up advertisements (a worthwhile and very valuable feature that works!). I have not had a pop-up appear since I added it to my web browser.

The third choice, , enables or disables back and forward buttons. I could not get these to work. These buttons do not duplicate the Back button on your browser. These buttons jump back and forth in your search results. This is useful, as it keeps you from going back to the search results page to get the next or a previous result.

Enable these options by left-clicking in the check box of your choice. A check mark signifies that it is enabled.

## Default Layout

This option allows you to return the toolbar to the default condition. All changes that you have made are erased, and the toolbar will appear as it did when you first installed it. You can choose to reset the toolbar with or without the advanced features (PageRank and Category) buttons enabled. Enable this option by left-clicking in the check box of your choice. A check mark signifies that it is enabled. Then, left-click on one of the Reset Layout bars as appropriate.



Finally, left-click on OK at the bottom of the page to save the choices that you have made.

## Clear Search History

You may recall that the search history is a list of previously searched-on keywords that are exposed when you left-click on the query box down arrow.



Figure 8.26: Toolbar search history

To clear the search history list, left-click on the Google down arrow  in the toolbar to expose the toolbar menu. A left-click on Clear Search History clears the toolbar search history.

## Help

A left-click on Help loads the Google Toolbar Help page in the browser. Toolbar help is located at <http://toolbar.google.com/help.html>.  
(Bookmark it.)

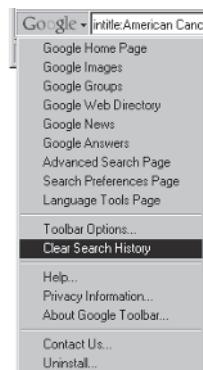


Figure 8.27:  
Clearing the  
search history



## Privacy Information

A left-click on Privacy Information loads the Google Toolbar Privacy Note page in the browser (res://c:\windows\googletoolbar.dll/prefetch.html).

It is worth a visit to the page to enable or disable the two advanced features to your liking. See Figure 8.28. If you want the advanced features enabled, left-click on the box next to each, placing a check mark in the box. To disable, left-click on the box and remove the check mark. Next, left-click on the OK button below the advanced features.



### Caution:

Google cautions on the Privacy Note page that if you enable the advanced features, information regarding the sites you visit will be sent to Google.

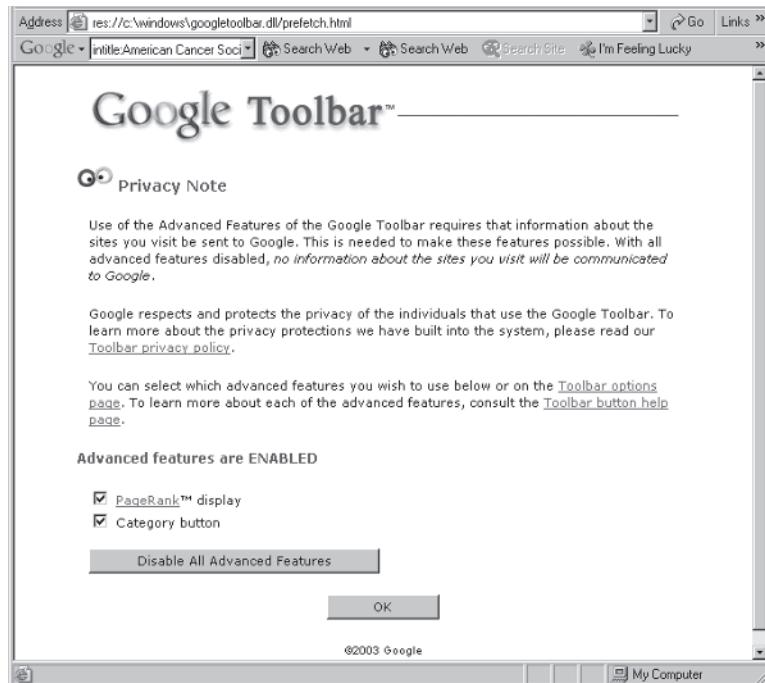


Figure 8.28: Privacy Note page



## About Google Toolbar

This loads Google's About Toolbar dialog.

## Contact Us

Left-clicking on Contact Us loads Google's Toolbar Support page in the browser. The Support page is located at <http://toolbar.google.com/feedback.html>. (Bookmark it? Probably not.)

## Uninstall

Left-clicking on Uninstall loads the Google Toolbar Uninstall page in the browser. The Uninstall page is located at `res://c:\windows\googletoolbar.dll/uninstall.html`. (Don't bother to bookmark it.) Uninstall will uninstall the toolbar.

## Toolbar Version 2.0

---

As of fall 2003, the standard Google Toolbar is version 2.0.102. You can find it at <http://toolbar.google.com>. To install the latest version, left-click on the Download Google Toolbar button, select Run this program from its current location, and left-click OK. In the next page that appears, left-click Yes to download the installer. A terms of use page will appear, to which you must agree to (by left-clicking Agree) to continue the installation. Follow the remaining instructions that appear.

If you already have the Google Toolbar installed on your computer, Google may have automatically updated it to the most current version. However, if a firewall is installed on your computer, the toolbar updates may be blocked. Additionally, depending on your security settings, you may see a dialog box requesting confirmation that you want to update the toolbar. To see what version your toolbar is, select About Google Toolbar.



In addition to the features already discussed in this chapter, the 2.0.x version of the toolbar has the following new options. They are:

- **Pop-up Blocker:** Blocks those annoying pop-ups. We discussed this earlier in the “Experimental Features” section. It is the same as Browser Control, only in a new version of the toolbar.
- **Word Find:** Finds your query words wherever they appear on the page
- **AutoFill:** Automatically fills in forms with the click of a mouse button
- **BlogThis:** Creates a weblog post pointing to the current page

Go ahead and install version 2.0. Nothing bad will happen.

Perhaps the most significant difference in the 2.0.x version of the toolbar is the addition of Google Links and an Options menu item. The Google Links menu item is just a rearrangement of the various web sites found in the previous toolbar menu. See Figure 8.29. The Options item, however, is much more interesting. See Figure 8.30.

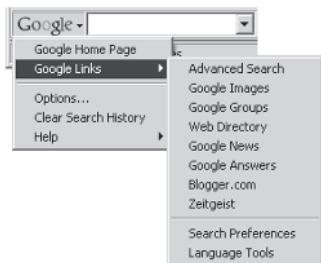


Figure 8.29: Google Links menu

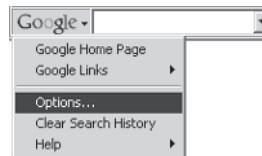


Figure 8.30: Options menu

Left-click on Options, and you can see the Toolbar Options dialog as shown in Figure 8.31. The Toolbar Options dialog has three tabs: Options, More, and AutoFill.



Figure 8.31: Toolbar Options dialog

The Options tab displays the basic toolbar options. You can enable each option by moving the cursor over to the appropriate check box and left-clicking in the box. A check mark in the box signifies that the option is enabled. To disable an option, left-click in the check box to remove the check mark. Note the Options button check box at the bottom of the dialog. When you check that box, you place an icon on the toolbar that gives you direct access to this dialog. To make your changes, left-click on Apply and then on OK at the bottom of the dialog.

Left-click on the More tab to see the options in this dialog. See Figure 8.32. Just as in the previous dialog, left-click to place/remove a check mark in the check boxes to enable/disable any options that you wish to appear on the toolbar. As before, to make your changes, left-click on Apply and then on OK at the bottom of the dialog.

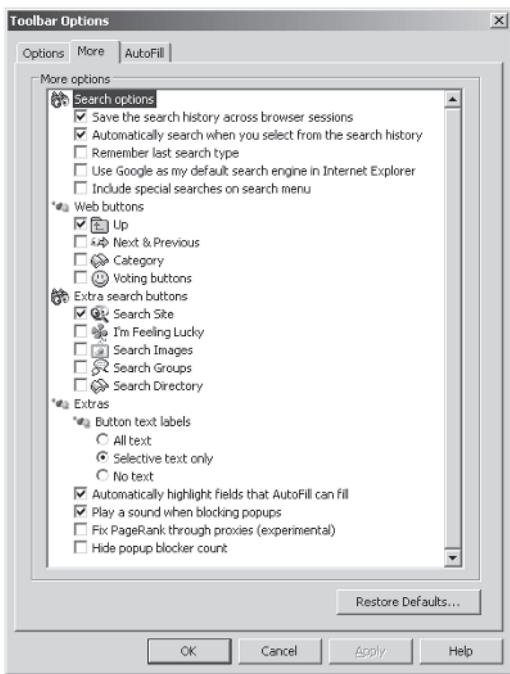


Figure 8.32: More options dialog

Now, left-click on the AutoFill tab. As you can see in Figure 8.33, you are requested to fill out various text boxes with personal information. Move the cursor from box to box using either the mouse or the Tab key. If you love the convenience of AutoFill, then you will find this feature very useful. However, if you are concerned about privacy and people gaining access to your personal information, avoid AutoFill like the plague! Once you fill in the information, hackers can access the data.

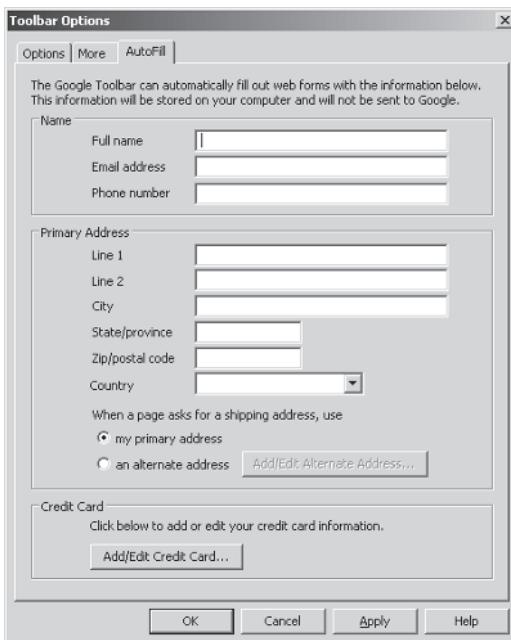


Figure 8.33: AutoFill dialog

If you decide to use AutoFill and want it to store your credit card info, left-click on the Add/Edit Credit Card button and then complete all the text boxes in the Credit Card Information dialog.

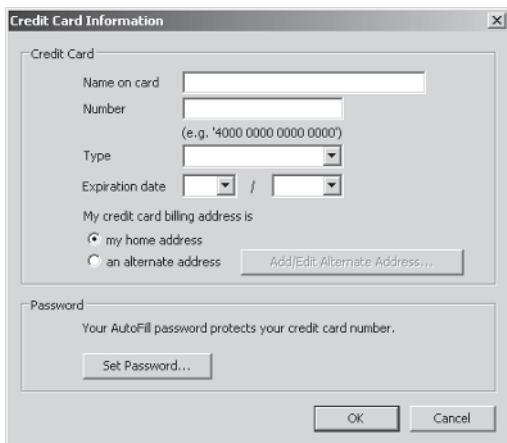


Figure 8.34: AutoFill Credit Card Information dialog



You can set a password that limits access to your credit card information via a browser, but it will not prevent a hacker from gaining access to your credit card information. Until there are better controls preventing any unauthorized access to our machines, I am not in favor of keeping this kind of information anywhere on a computer. But you may feel differently. In that case, left-click on Set Password and complete the two text boxes. You can have up to 15 characters and any combination of letters, numerals, and text formatting characters, such as the asterisk (\*). Use the mouse and left-click inside each box to move the cursor, or use the Tab key. Left-click on OK (or Cancel) when finished.

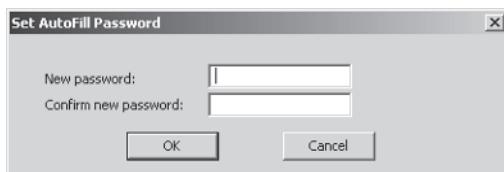


Figure 8.35: Set AutoFill Password dialog

Now that you have decided to use AutoFill, you can take advantage of the Alternate Address feature to add a second address for shipping or billing purposes. Left-click on the An alternate address option. Then left-click on Add/Edit Alternate Address. Fill in the required information in the text boxes, moving from box to box using the mouse or Tab key. See Figure 8.36. Left-click on OK (or Cancel) when you are finished.

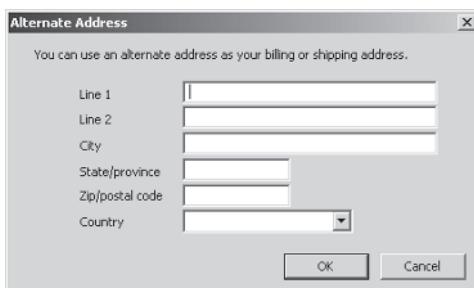


Figure 8.36: Alternate Address dialog



Now, left-click on OK (or Cancel or Apply) and then OK to exit the Options dialog.

Google is always experimenting with the toolbar features, so who knows what you will see by the time this book gets in your hands. Hopefully, there will be enough similarity that you can experiment and work through any new options or deviations to the current ones. Be brave and explore. Not much can happen, other than crashing your machine. So always save your work before going off on exploration journeys!

## Summary

---

In this chapter we discussed the Google Toolbar with its many features, drop-down menus, pop-up menus, options, and tools. The toolbar is certainly useful to have. The choices given by the toolbar to customize the search, and especially the pop-up blocker, make searching the web a much more pleasurable experience. Install the toolbar today and increase your search capability while reducing your search frustration immediately.

# Chapter 9

## Other Google Features

In this chapter, we examine various features, including Google Help Central, Google Site Map, All About Google, General FAQ, configuring preferences, and setting Google to be your default home page and default search engine. Additionally, we discuss web APIs, webmaster information, submitting your URL, advertising with Google, special tricks and treats, and finally your security.

### Google Help Central

Google Help Central is located at <http://www.google.com/help/index.html>. (Yes, bookmark it!) See Figure 9.1. Help Central is your search “fire station.” Go to Help Central and click on the topic of interest to get advice on how to perform the topic function. Be sure that you have bookmarked it, as you will find it invaluable.

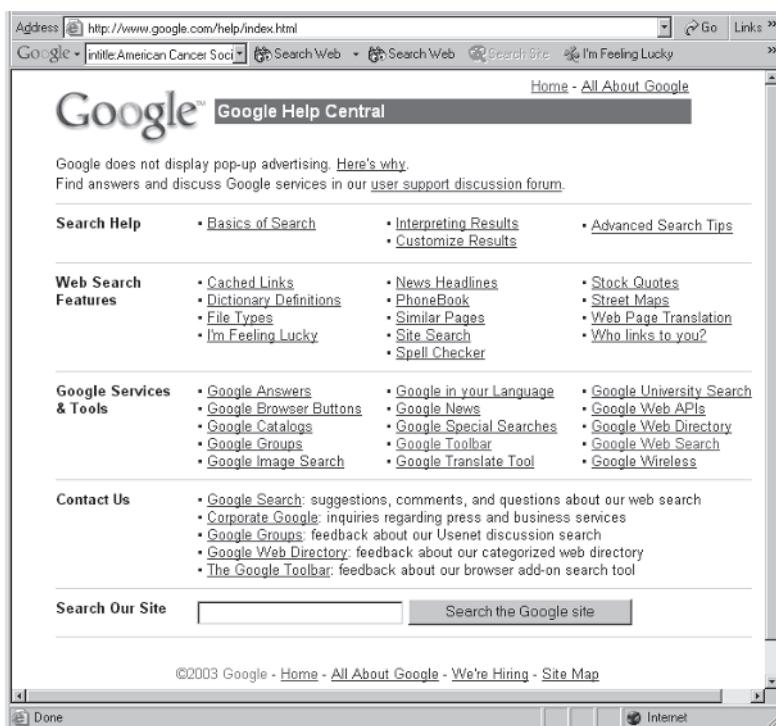


Figure 9.1: Google Help Central

## Google Site Map

---

Google's excellent Site Map is located at <http://www.google.com/sitemap.html>. (Bookmark it!) Use the Site Map in conjunction with Help Central to navigate your way around Google.



Our Search	Our Company	More Google	For Site Owners
<b>Search Tips</b> <ul style="list-style-type: none"> <li><a href="#">Overview</a></li> <li><a href="#">Basics of Search</a></li> <li><a href="#">Interpret Results</a></li> <li><a href="#">Advanced Search Tips</a></li> <li><a href="#">Customize Google</a></li> <li><a href="#">Special Features</a></li> <li><a href="#">FAQ</a></li> </ul> <b>Services &amp; Tools</b> <ul style="list-style-type: none"> <li><a href="#">Overview</a></li> <li><a href="#">Google Answers</a></li> <li><a href="#">Google Catalogs</a></li> <li><a href="#">Google Groups</a></li> <li><a href="#">Google Image Search</a></li> <li><a href="#">Google News</a></li> <li><a href="#">Special Searches</a></li> <li>- <a href="#">University Search</a></li> <li>- <a href="#">U.S. Government</a></li> <li>- <a href="#">Linux</a></li> <li>- <a href="#">BSD</a></li> <li>- <a href="#">Apple Macintosh</a></li> <li><a href="#">Google University Search</a></li> <li><a href="#">Google Directory</a></li> <li><a href="#">Wireless</a></li> <li><a href="#">Google Browser Buttons</a></li> <li><a href="#">Google in your Language</a></li> <li><a href="#">Google Toolbar</a></li> </ul>	<b>Press Center</b> <ul style="list-style-type: none"> <li><a href="#">Press Kit</a></li> <li><a href="#">Press Releases</a></li> <li><a href="#">In the News</a></li> <li><a href="#">Image Gallery</a></li> <li><a href="#">Reviewer's Guide</a></li> <li><a href="#">Google Zeitgeist</a></li> </ul> <b>Our Company</b> <ul style="list-style-type: none"> <li><a href="#">Core Overview</a> <ul style="list-style-type: none"> <li>- <a href="#">Milestones</a></li> <li>- <a href="#">Fun Facts</a></li> </ul> </li> <li><a href="#">Fact Sheet</a></li> <li><a href="#">Management</a></li> <li><a href="#">Investor Info</a></li> <li><a href="#">Awards</a></li> </ul> <b>Our Technology</b> <ul style="list-style-type: none"> <li><a href="#">Tech Overview</a></li> <li><a href="#">Features</a></li> <li><a href="#">Search Tips</a></li> <li><a href="#">Life of a Query</a></li> <li><a href="#">Tech Highlights</a></li> </ul> <b>Our Business</b> <ul style="list-style-type: none"> <li><a href="#">Business Overview</a></li> <li><a href="#">Success Stories</a></li> <li><a href="#">Customers</a></li> </ul> <b>Corporate Info</b> <ul style="list-style-type: none"> <li><a href="#">Profile</a></li> <li><a href="#">What's a Google?</a></li> </ul>	<b>Picture Us</b> <ul style="list-style-type: none"> <li><a href="#">Inside Google</a></li> </ul> <b>Logos</b> <ul style="list-style-type: none"> <li><a href="#">Official Logos</a></li> <li><a href="#">Holiday Logos</a></li> <li><a href="#">Fan Logos</a></li> <li><a href="#">Permissions</a></li> <li>- <a href="#">Guidelines</a></li> <li>- <a href="#">Request Form</a></li> <li>- <a href="#">Brand Terms</a></li> </ul> <b>Wear Us</b> <ul style="list-style-type: none"> <li><a href="#">Google Store</a></li> </ul> <b>Talk to Us</b> <ul style="list-style-type: none"> <li><a href="#">Contact Us</a></li> <li><a href="#">Share Success</a></li> <li><a href="#">Google Newsletter</a></li> </ul> <b>Advertise with Us</b> <ul style="list-style-type: none"> <li><a href="#">Advertising Programs</a></li> </ul> <b>Premium Sponsorships</b> <ul style="list-style-type: none"> <li><a href="#">Getting Started</a></li> <li><a href="#">Overview</a></li> <li><a href="#">Take the tour</a></li> <li><a href="#">Benefits</a></li> <li><a href="#">Preview tool</a></li> <li><a href="#">FAQ</a></li> <li><a href="#">Terms &amp; conditions</a></li> </ul> <b>Account Resources</b> <ul style="list-style-type: none"> <li><a href="#">Campaign login</a></li> <li><a href="#">Performance tips</a></li> <li><a href="#">Client Services</a></li> </ul> <b>Stats &amp; Studies</b> <ul style="list-style-type: none"> <li><a href="#">Demographics</a></li> <li><a href="#">Case studies</a></li> </ul> <b>Contact Sales</b> <ul style="list-style-type: none"> <li><a href="#">Worldwide offices</a></li> <li><a href="#">Inquiry form</a></li> </ul> <b>Google AdWords</b> <ul style="list-style-type: none"> <li><a href="#">Introduction</a></li> <li><a href="#">Program Overview</a></li> <li><a href="#">Compare Programs</a></li> <li><a href="#">Common Terms</a></li> <li><a href="#">Tips</a></li> <li><a href="#">FAQ</a></li> </ul>	

Figure 9.2: Google Site Map

We have covered the Our Search section of Google in great depth in this book. The other three sections of Google (Our Company, More Google, and For Site Owners) are not covered in this book. Those sections are, after all, of little or no use to people trying to improve their search skills, the topic of this book. Do take note of the Contact Us link in the Talk to Us subsection of More Google.



# All About Google

---

All About Google is found at <http://www.google.com/about.html>. (Don't bookmark it unless you are writing a book about Google.) Note the Search our site search box in the top-right corner.

The screenshot shows a web browser window with the URL <http://www.google.com/intl/en/allabout.html> in the address bar. The page itself is titled "All About Google". It features several sections of links:

- Top News**: Google Expands Advertising Monetization Program for Websites
- Your Stories**: Luck is where you find it... people share their Google stories
- What's up in the lab?**: Google Viewer and more. Try our latest demos in Google Labs
- We're Hiring**: See open positions
- The Google Toolbar**: Free download for your IE browser.
- Our Search**: Help and How to Search, Basics, Advanced Features, FAQs...
- Google Web Search Features**: Translation, I'm Feeling Lucky, Cached...
- Google Services & Tools**: Toolbar, Google Web APIs, Buttons...
- Our Technology**: PageRank, Benefits of Google...
- Jobs at Google**: Openings, Perks, Culture...
- Corporate Info**: Executives, Company Profile, Address...
- For Site Owners**: Advertise with Us, Premium Sponsorships, AdWords...
- Business Solutions**: Google Search Appliance, WebSearch, Wireless...
- Webmaster Info**: Guidelines, Ranking questions...
- Submitting your Site**: How Google adds URLs...
- Our Company**: Press Center, Articles, Releases, Zeitgeist...
- More Google**: Contact Us, FAQs, Feedback, Newsletter, Permissions...
- The Google Store**: Pens, Shirts, Lava lamps...
- Logos and Photos**: Logos, Googlers at work and play...

At the bottom of the page, there is a footer with the text "©2003 Google Privacy Policy - Terms of Service".

Figure 9.3: All About Google



# General FAQ

Google's Frequently Asked Questions (FAQ) site is located at <http://www.google.com/help/faq.html>. (Bookmark it!) The FAQ web page is shown in Figure 9.4. Note the section titled International Questions. There are several questions and answers that many Americans will find useful. Most of us collide with foreign words, phrases, and web pages from time to time. So if you find yourself in a quandary with a foreign language web page, remember that this section of FAQs may help you.

The screenshot shows a web browser window displaying the Google General FAQ. The address bar shows the URL <http://www.google.com/help/faq.html>. The page title is "Google General FAQ". On the left sidebar, there are links to Home, All About Google, Help Central, Search Help (with sub-links for Basics of Search, Advanced Search, Interpret Results, and Customize), Google Features, General FAQ, and Contact Us. Below these links is a search bar labeled "Find on this site:" with a "Search" button. The main content area is titled "Frequently Asked Questions" and contains a link to the "user support discussion forum". Under the "Search Questions" heading, there are two sections: "Search Questions" and "International Questions". The "Search Questions" section lists six numbered questions, and the "International Questions" section lists five numbered questions. At the bottom of the page, there is a link to "Where can I send my question?"

Figure 9.4: Google FAQ

Also, note the links on the left-hand side of the web page. Help Central is just a left-click away.



# Saving Preferences

---

We covered preferences in a previous chapter. However, we assumed that your browser was configured properly for saving your preferences. If you find that your preferences are not being saved, you may need to perform the following procedure. How can you tell if your preferences were saved? Simple; after clicking on Save Preferences in the Preferences page, close your browser and then reopen it. Go to the Preferences page, and see if Google remembered your preferences.

To save your preferences, you must have cookies enabled. A help page that provides instructions on how to set cookies for Internet Explorer versions 4.0 and 5.0 is located at [www.google.com/cookies.html](http://www.google.com/cookies.html). However, there are no instructions for Internet Explorer 6.0. To fill that void, Internet Explorer 6.0 instructions are provided here.

To enable cookies, follow the instructions below for the browser version Microsoft Internet Explorer 6.x. You can set your browser to accept cookies from all sites or you can set the browser to accept cookies only from sites that you specify. I recommend that you set your browser to accept cookies from sites that you specify. See the following sections.

## Accepting Cookies

This option sets the browser to accept cookies from anyone who wants to feed it one. It is not the recommended option, but it is the only choice that allows you to set it and then forget about it. To begin:

1. Select Internet Options from the Tools menu.
2. Left-click on the Privacy tab.
3. Left-click the Advanced button. See Figure 9.5.
4. Select Override automatic cookie handling. See Figure 9.6.



5. Select Accept for both First-party Cookies and Third-party Cookies.
6. Left-click OK to close the Advanced dialog.
7. Left-click OK to close the Internet Options dialog.



Figure 9.5: Setting cookies



Figure 9.6: Setting cookies



## Accepting Cookies from Specific Sites

This option lets you specify a web site from which you will accept cookies. It is the recommended way to accept cookies from trusted sites.

1. Select Internet Options from the Tools menu.
2. Left-click on the Privacy tab.
3. Left-click the Edit button.
4. Enter `http://www.google.com/cookies.html` in the Address of Web site text box. See Figure 9.7.
5. Left-click on the Allow button. You should see the http address you entered in the text box appear in the Managed Web sites text box, with Always Allow under Setting.
6. Left-click OK to close the Per Site Privacy Actions dialog.
7. Left-click OK to close the Internet Options dialog.

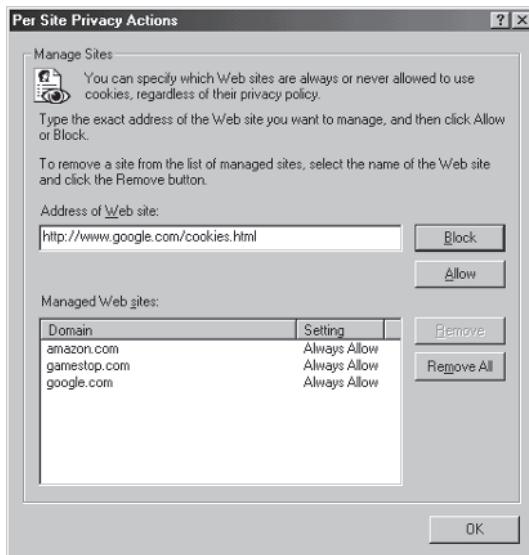


Figure 9.7: Setting cookies



# Setting Google as Your Default Home Page

This section explains how you can make Google your default home page. The default home page is the page that is displayed when you start your browser. The instructions for changing your default home page to Google can be found at <http://www.google.com/options/defaults.html>. (Bookmark it? Nope.) The instructions are included here and cover the following browsers:

## Netscape 4.0 to 6.0

1. Left-click on the Edit | Preferences menu selection.
2. Left-click on Navigator.
3. In the Navigator Starts with section, select Home page.
4. In the Home page section, type <http://www.google.com/> in the text box.
5. Left-click OK.

## Internet Explorer

1. Left-click on the Tools | Internet Options menu selection.
2. Left-click on the General tab.
3. In the Home page section, type <http://www.google.com/> in the text box.
4. Left-click OK.

## America Online

1. Make sure that you are viewing Google's home page ([www.google.com](http://www.google.com)).
2. Copy Google's URL ([www.google.com](http://www.google.com)) to the clipboard.



3. Go to the Members drop-down menu and select Preferences.
4. Left-click on the WWW icon.
5. Left-click on the Home Page box at the bottom of the pop-up window and then press your Ctrl and V keys (at the same time) to paste in the address. If the address does not appear in the box, simply type in what you tried to paste from above.
6. Left-click the OK button.

## Setting Google as Your Default Search Engine

---

Making Google your default search engine can be problematic. It is recommended that only experienced computer users attempt the following changes and only then after backing up the system registry. To make Google your default search engine, follow the instructions given here.

### Netscape Communicator

For Netscape 4.0 to 4.7, it is possible to search by entering ? <search terms> in the URL box instead of a URL. You can set the search engine used in this case by directly editing the prefs file. This can be dangerous! See the Unofficial Netscape FAQ (<http://www.ufaq.org/commonly/userprefs.html>) for more details.

Before you go any further, quit any running copies of Netscape.

### Windows Operating System

To edit the prefs file, use Notepad to load ...\\Netscape\\Users\\*{username}*\\prefs.js. (“...” is probably the Programs folder.)



## Macintosh Operating System

To edit the prefs file, use SimpleText or the equivalent to load ...:preferences:netscape f:netscape preferences. (“...” is your system folder.)

## UNIX Operating System

Edit the file .netscape/preferences.js in the home directory.

Once the preferences file has been loaded, add the following line to the end of the file:

```
user_pref("network.search.url",
"http://www.google.com/keyword/");
```

Press Enter at the end of the line. The end of the line is the semi-colon! Note that there is no line feed (do not press Enter here!) at the end of “user\_pref("network.search.url",”.

## Internet Explorer

### Windows Operating System

To make Google the default search engine for Internet Explorer versions 4.0 to 6.0, make the following registry changes. Reboot the system after making the registry changes.



#### Note:

You must have Administrator privileges on your computer in order to edit the registry.

To have the search results displayed in the search channel window:

```
[HKEY_CURRENT_USER\Software\Microsoft\Internet
Explorer\Main]
"Use Search Asst"="no"
"Search Page"="http://www.google.com"
"Search Bar"="http://www.google.com/ie"
```



```
[HKEY_CURRENT_USER\Software\Microsoft\Internet  
Explorer\SearchURL]  
""="http://www.google.com/keyword/%s"  
"provider"="gogl"  
  
[HKEY_LOCAL_MACHINE\SOFTWARE\  
Microsoft\Internet Explorer\Search]  
"SearchAssistant"="http://www.google.com/ie"
```

The Search button will bring up a Google search box on the left side of the browser.

The Search The Web menu option from the Go menu will take you to Google's home page.

For IE4 only, you can type ? <search terms> in the URL box instead of a URL, and it will perform the search.

To have the search results displayed in the browser's main window:

```
[HKEY_CURRENT_USER\Software\Microsoft\Internet  
Explorer\Main]  
"Use Search Asst"="no"  
"Search Page"="http://www.google.com"  
"Search Bar"="http://www.google.com/ie_rsearch.html"  
  
[HKEY_CURRENT_USER\Software\Microsoft\Internet  
Explorer\SearchURL]  
""="http://www.google.com/keyword/%s"  
"provider"="gogl"  
  
[HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\  
Internet Explorer\Search]  
"SearchAssistant"="http://www.google.com/ie_rsearch.html"
```

If you decide that you want to revert to the original defaults, you can do one of two things: You can enter the following registry keys or you can restore your registry. To set the search engine back to the original Internet Explorer default:

```
[HKEY_CURRENT_USER\Software\Microsoft\Internet  
Explorer\Main]
```



```
"Use Search Asst"="yes"  
"Search Page"="http://home.microsoft.com/access/allinone.asp"  
"Search Bar"="http://home.microsoft.com/search/lobby/  
search.asp"  
  
[HKEY_CURRENT_USER\Software\Microsoft\Internet  
Explorer\SearchURL]  
""="http://home.microsoft.com/access/autosearch.asp?p=%s"  
"provider"="  
  
[HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\  
Internet Explorer\Search]  
"SearchAssistant"="http://ie.search.msn.com/{SUB_RFC1766}  
/srchassst/srchassst.htm"
```

## Macintosh Operating System



### Note:

This does not work with Mac OS X.

Setting up IE 5 to launch Google in the sidebar is covered at <http://www.visakopu.net/ie5google/>. Good luck!

## Mozilla/Netscape 6 Search Options

For Mozilla and Netscape 6 search options, visit <http://www.google.com/mozilla/google-search.html>. You can set Google to be the default search engine, set up search keywords, and put Google in the sidebar search panel.

## IE QuickSearch

You can add Google to the list of search engines by adding a new entry (select New) and setting up a shortcut gg and a custom URL <http://www.google.com/search?q=%s>. Now you can perform a search by entering “*gg keywords*” in the address bar. For example, “gg flower pots” will search Google for “flower pots” (without the quotes).

**Note:**

The QuickSearch utility for Internet Explorer must be installed.

## Macintosh OS X Service

You can register Google as a service available to all applications under the application menu. Go to <http://gu.st/proj/SearchGoogle.service/> (the web site in the Google FAQ [http://www.apple.com/downloads/macosx/internet\\_utilities/searchgoogleservice.html](http://www.apple.com/downloads/macosx/internet_utilities/searchgoogleservice.html) is no longer working) and install the Google service. After installation, you can select text in any application and press Shift+Cmd+G to launch a Google search for that text.

## Additional Google Web Search Features

---

Google contains a few additional search features that were not covered in other chapters. We examine them in the following sections. The features are:

- **Calculator:** Calculates mathematical expressions entered in the search box
- **Dictionary Definitions:** Dictionary definitions for one or more of the search words
- **I'm Feeling Lucky:** Go immediately to the web page Google ranks number one
- **PhoneBook:** U.S. street addresses and phone book
- **Spell Checker:** Checks the spelling of keywords including alternate spellings
- **Street Maps:** Find/display street maps of U.S. cities



## Calculator

Google has added a math capability to the search box. When mathematical expressions are entered in the search box and either Enter is pressed or the Search button is clicked, the expression is evaluated and the answer is displayed on a results page. Google's calculator solves math problems ranging from basic arithmetic to more complicated math and can manage units of measure including conversions while recognizing physical constants. Standard mathematical notation is also recognized. Expressions in parentheses are evaluated before expressions outside the parentheses. Calculator recognizes standard mathematical operators. Table 9.1 lists some of the math operators that are valid in Calculator:

**Table 9.1: Math operators**

Operator	Function	Example
+	addition	2+4
-	subtraction	8-3
*	multiplication	6*9
/	division	24/8
^	exponentiation (raise to a power of)	2 ^ 5
%	modulo (finds the remainder after division)	9%8
choose	X choose Y determines the number of ways of choosing a set of Y elements from a set of X elements	12 choose 3
nth root of	calculates the nth root of a number	6th root of 64
% of	X % of Y computes X percent of Y	25% of 250
sqrt	square root	sqrt(25)
sin, cos, etc.	trigonometric functions (numbers are assumed to be radians)	cos(pi/2) tan(45 degrees)
ln	logarithm base e	ln(13)
log	logarithm base 10	log(100)
!	factorial	17!
in	unit conversion	10 miles in kilometers



Unit conversions are managed by including the operator “in” — without the quotes — at the end of the expression followed by the unit of measure you want the result to be in. So, to convert miles to kilometers, use “x miles in kilometers” where x is some whole number. For example, “10 miles in kilometers” will yield 16.09344 kilometers.

The “in” operator is used to specify what units you want used to express the answer. Put the word “in” followed by the name of a unit at the end of your expression. This works well for unit conversions such as: 5 kilometers in miles.

Calculator recognizes units of measure in their long form (kilometers) and their short form (km). Calculator also recognizes physical constants, such as c for the speed of light (in a vacuum). When searched upon, pi and c returned these search results:

pi = 3.14159265

c = the speed of light = 299 792 458 m / s

The constants seem to be limited to those you can enter with the typical keyboard, that is, ASCII characters 0 to 127. When I tried to use constants identified on the keyboard by ASCII characters 128 to 255, I did not get anything related to the constants. So, Greek symbols did not seem to be working at the time of publication.

Calculator recognizes hexadecimal, octal, and binary numbers. To use this feature, prefix hexadecimal numbers with 0x, octal numbers with 0o, and binary numbers with 0b. For example: 0x8f + 0b10110101.

I tried various mathematical operations using negative numbers and received the correct answer every time. Next, I wanted to see if I could spoof Calculator by giving it a tougher assignment. I asked it to tell me the ( $\sqrt{-1}$ ). It gave me the correct response, ( $\sqrt{-1} = i$ ). I wanted to see how it handled undefined operations, so I entered (15/0) and received 2,680,000 web page hits that had “15,” “/,” and “0” somewhere within the page. So, Calculator does not handle undefined math relationships as a math expression but treats such as just a plain ordinary search on the terms entered.



Well, at least it did not freeze my machine as such math expressions were prone to do 20 and even five years ago.

I did not take the time to see if Calculator can convert from rectangular coordinates to polar coordinates. If you check this out, please let me know the results you achieved.

This is another one of those fun things to play around with. Now, let's hope Google adds integral and differential calculus to Calculator's capability and a graphing function to see what we have. Then it's "goodbye" Mathsoft and Mathcad! And to think I paid \$600 for an HP calculator to do the same thing in 1991. What was I thinking?

## Dictionary Definitions

Dictionary Definitions is exactly what it sounds like. This feature provides definitions of words that you supply. I illustrate Dictionary Definitions with a search example. To view Google's definition of your query word(s), perform the usual search. In this example, we search for "flower pot." See Figure 9.8. In the blue bar below the Google search text box, you can see the search words "flower" and "pot" underlined. If Google has a word definition in its dictionary for a search word, it will underline the word in that blue bar. To view the definition, click on the underlined word of interest. If we click on "flower," we see Google's definition. See Figure 9.9. Note that the definition page also includes a link to a thesaurus.

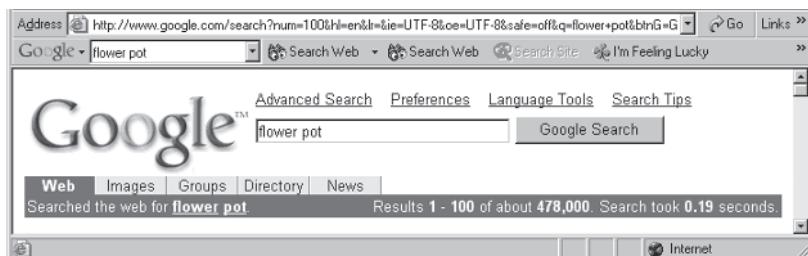


Figure 9.8: Dictionary Definitions page

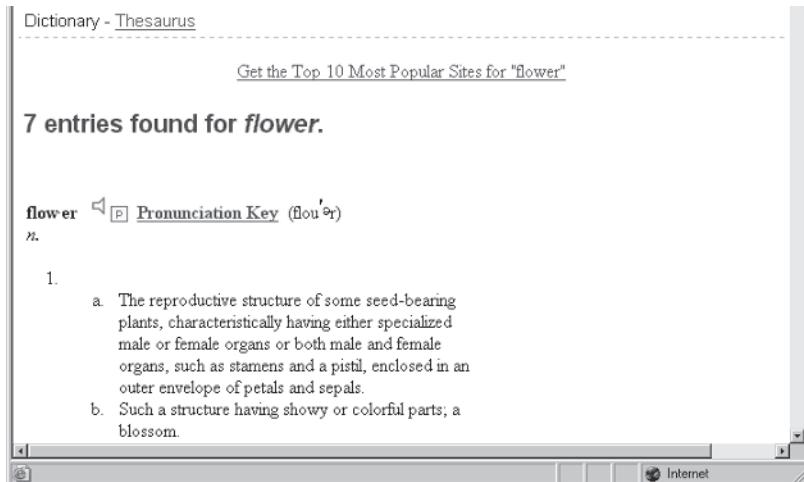


Figure 9.9: Flower definition

The Pronunciation Key link shown in Figure 9.9 pops up a dictionary.com window with a pronunciation key from *The American Heritage Dictionary*.

The Get the Top 10 Most Popular Sites for “flower” link will take you to a web page listing ten web sites.

Google infers in some of its material that the dictionary supplier may change. In the meantime, this is a feature that can be very useful. If you see a word in a document that is unfamiliar, just copy it into the search box and search on it. If the blue bar on the search results page at the top of the page has the word underlined, then click on the underlined word and read the definition. It sure beats getting up and walking over to the bookcase and pulling the old dictionary off the shelf.

## I'm Feeling Lucky

Let me say that I do not like this feature. I'm Feeling Lucky makes it too easy to ignore the vast ocean of information waiting at your fingertips to be navigated. Left-clicking on I'm Feeling Lucky takes you immediately to the web page that Google ranks number one.



Well, that shuts out a lot of exciting, interesting, and useful web sites. Just because a web site paid \$1,000,000 to get 10,000 other web sites to link to it, giving it a much higher page rank than Mom and Pop's home page, does not make it the web site of choice. Didn't know that, eh? Yes, companies will pay other web sites to link to their web site so they can get a very high search ranking. Ever wonder why you see many of those banner ads and pop-up ads? Each one of those is a link that PageRank will count in its popularity contest. By now you may realize that I am not a fan of PageRank. Anyway, if you do not like being manipulated, I suggest you do not use this feature.

## PhoneBook

Want to find a business or a person? Easy. Just key the appropriate information in the search box, and if they are publicly listed, Google will provide a phone number and address. As an example, I keyed in my name, "Michael Busby," (without the quotes) and the city "Plano" (again, without the quotes) and pressed Enter. If a publicly listed telephone number and address are available, PhoneBook will place the information at the top of the search results page. The results for "Michael Busby Plano" are shown in Figure 9.10. My phone number and address have been grayed out, since I don't want any irate readers egging my house. Notice that the Google PhoneBook feature returned two phone numbers and addresses, one for a "Michael Busby" and the other for a "Michael G. Busby." Ahh, you will never know if they are one and the same, as no one knows my middle initial. However, the point is, PhoneBook returned all those named "Michael Busby" that it found in Plano.

If your search words contain certain keywords, PhoneBook is activated and a check for public listings is performed. No more directory assistance at \$1.00 (or more) per pop! In a year's time, you will save in directory assistance charges what you spent on this book. A one-year return on an investment is awesome! Try it.



Figure 9.10: Google PhoneBook

I mentioned earlier that you need to use certain words in your search query to activate the PhoneBook. The keywords are:

**For a U.S. business:**

- Key the business name in the Google search box and the business city and state in the Google search box.

Or

- Key the business name in the Google search box and the city zip code.

Or

- Key in the business phone number (include the area code) to find the business address.

**For a U.S. resident, key in any one of the following options:**

- First name (or first initial), last name, city (state is optional)
- First name (or first initial), last name, state
- First name (or first initial), last name, area code
- First name (or first initial), last name, zip code
- Phone number, including area code



- Last name, city, state
- Last name, zip code

If the query results in both a business and a personal or residential listing, both will be displayed. Also, PhoneBook does not support the use of wildcard characters, such as the asterisk (\*).

Note in Figure 9.10 that I did not place quotes around my name. You may recall that quotes around query words require the search engine to return search results with links to web pages containing exactly the same sequence of words. Without quotes, the search engine returns search result links to web pages containing those words anywhere within the web page, not necessarily in the same sequence or even located within shouting distance of each other. Yet, PhoneBook acts as if quotes are around my name. That is because the PhoneBook database is a directory listing, very much the same as a paper phone book lying near your phone, not a web page.

Notice that there are two links under “Michael Busby” in Figure 9.10. Both of these, Yahoo Maps and MapQuest, locate and provide driving directions to my home in Plano. Just left-click on the link to see the map display.

If you left-click on the  telephone icon, you will see an informational page (Google Web Search Features at <http://www.google.com/help/features.html#wp>) about the Google PhoneBook feature. The page includes a link to a form that you can use to request that your personal information be removed from PhoneBook. We discuss removing a listing from PhoneBook later in this section.

In the next three figures, we are going to search on my name again. I point out an interesting twist to PhoneBook. Figure 9.11 illustrates where we start out when performing most Google searches, the Google home page. After entering my name and the state I live in (Texas), I left-click on the Google Search button.



Figure 9.11: Searching for M Busby Texas

The resulting search results include not only my address and phone number, but 14,000 other search results. See Figure 9.12. Does this mean that there are 14,000 people named “M Busby” in Texas? No. It means that there are 14,000 total web pages that are part of the search results.

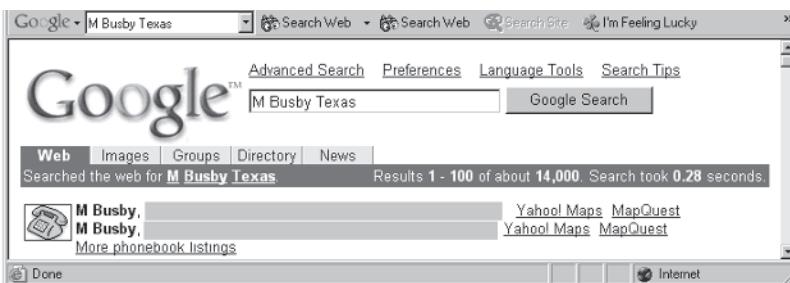


Figure 9.12: Searching for M Busby Texas search results

There is a mysterious More phonebook listings link below my name. When we click on it, a page of “M Busby” phone numbers is displayed, “1 - 30 of about 107.” See Figure 9.13. The names, phone numbers, and addresses are not shown for obvious reasons. The really neat thing about searching in this manner is the ability to locate someone in a particular state if you do not know the



community in which he or she resides. You can search the state and get addresses and telephone numbers. Armed with this information, you can narrow the search by phoning and mailing. What legitimate purpose could such a search effort have? Well, I found a long-lost relative searching the Internet in this manner. The power to search an entire state for someone with just a couple of key-strokes is awesome.

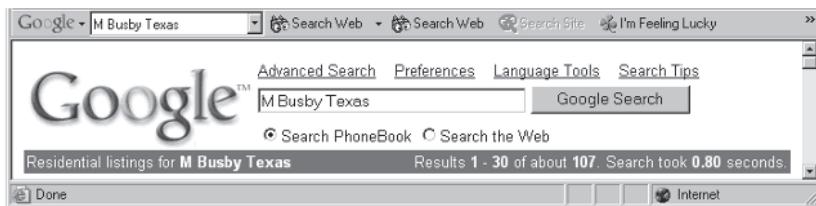


Figure 9.13: Searching for M Busby Texas search results

Notice in Figure 9.13 the Search PhoneBook option below the query box. PhoneBook search results pages are the only web pages where I have been able to locate this feature. We can resolve the domain name to a PhoneBook “home page” that we can bookmark and conveniently search future names and addresses without having to perform a general search on the web, as we were required to do to get to this page.

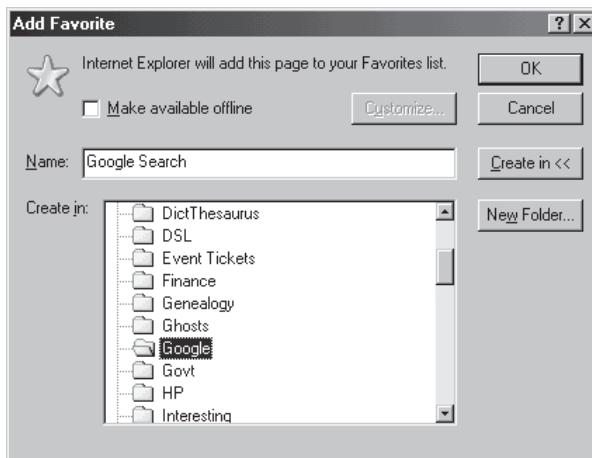


Figure 9.14: Saving Google Search (PhoneBook) bookmark



By lopping off the portion of the domain name that is specific to the search on my name, I get a PhoneBook “home page.” The domain name is [http://www.google.com/search?sa=X&oi=fwp&pb=f&q=.](http://www.google.com/search?sa=X&oi=fwp&pb=f&q=)

Bookmark it and name it PhoneBook! Initially, the Add Favorite dialog (IE 6.0) box will prompt you to save the bookmark as “Google Search.” Well, we already have a bookmark named Google Search. Besides, this is a search of PhoneBook, so it is wise to rename the bookmark “Google PhoneBook.” To rename the bookmark, delete “Search” and then key in “PhoneBook.” Next, left-click on OK.

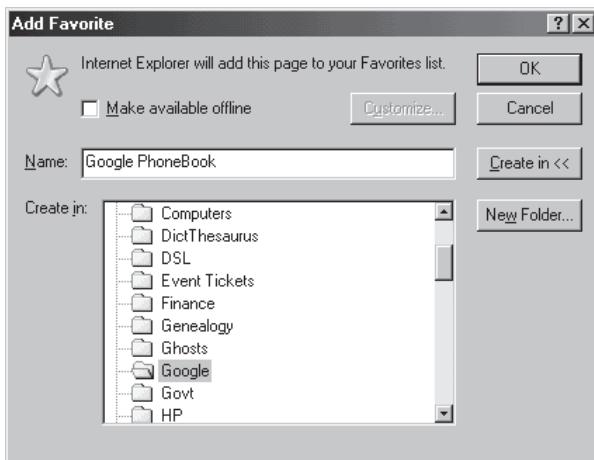


Figure 9.15: Renaming Google Search to Google PhoneBook

Did you search on your name? Did you discover that you are easily identifiable on the web? If you are disturbed that your name, phone number, and address appear when searched upon, you may request that your information be removed by visiting <http://www.google.com/help/pbremoval.html>. The removal form is simple. See Figure 9.16. The Google PhoneBook removal page provides links to six other reverse phone number lookup services. Of course, removing it from Google and the six shown services does not get it out of the other several hundred or so web phone books. Only if you have an unlisted number can you have some degree of privacy. The people who make a living gleaning personal information and selling it can get your data from other sources, such as schools and colleges. So,



even if you opt for an unlisted number, you may still find your phone number and address present on the web.

The screenshot shows a Google search results page for "M Bushy Texas". The first result is a link to a removal service. The page contains the following text and form fields:

**IMPORTANT NOTE:** Removing your phonebook listing will not remove your personal information from other pages on the web or from other reverse phone listing lookup services, such as:

- [Anywho](#)
- [Whitelpages.com](#)
- [Switchboard.com](#)
- [Reverse Phone Directory](#)
- [Phonenumber.com](#)
- [Smartpages.com](#)

For a comprehensive list of reverse phone lookup services, try a [Google search](#).

Please enter all information exactly as it appears in your phonebook entry:

Enter your name  
[Text input field]

Enter your city and state (e.g. Mountain View, California)  
[Text input field]

Enter the phone number to be removed  
( [Text input field] ) [Text input field] - [Text input field]

Enter your email address (e.g. you@domain.com)  
[Text input field]

Reason for removal

My phone number is incorrect

Privacy concerns

Other

Figure 9.16: Removing your PhoneBook listing

## Spell Checker

Google's Spell Checker checks the spelling of all your search words when you click to begin the search. Google suggests an alternate spelling if there is any. The software behind the spell checker returns search results using the alternate spelling if it improves your chances of getting more hits. Then, it asks you on the results page if you really meant to use the alternate spelling. So much for the old spinster aunt correcting my spelling (now a machine has to do it).



## Street Maps

Street Maps returns a web page with a link provided by a third-party map supplier. To use Street Maps, enter a U.S. street address, including the city and state or the zip code, in the Google search box. Simple as that.

## Special Tricks and Treats

Google does not recognize the asterisk (\*) as a wildcard in lieu of a character. That is, “aut\*” will not return “auto” or “autumn.” However, Google does recognize the asterisk as a wildcard in lieu of a word. That is, “\* cancer” will return search results listing the links to web pages describing every type of cancer.

Google \* cancer

Advanced Search Preferences Language Tools Search Tips

\* cancer Google Search

Web Images Groups Directory News

Searched the web for \* cancer Results 1 - 100 of about 17,300,000. Search took 0.21 seconds.

News: [Betel a major cancer risk - study](#) - Independent Online - Aug 9, 2003  
[Bruno Has Prostate Cancer, But Says Prognosis Is Good](#) - New York Times - Aug 8, 2003  
[US team finds hints of how, why cancer spreads](#) - Hindustan Times - Aug 8, 2003  
[Try Google News: Search news for \\* cancer or browse the latest headlines](#)

**Cancer Diagnosis?**  
 17 secrets of survival you must know within 1 week of diagnosis.  
[www.survivors-wisdom.com](#)  
 Interest: [progress bar]

[See your message here...](#)

**American Cancer Society Homepage**  
 The American Cancer Society is dedicated to eliminating cancer as a major health problem by preventing cancer, saving lives, and diminishing suffering through ...  
 Description: Dedicated to helping persons who face cancer. Supports research, patient services, early detection, ...  
 Category: Health > Conditions and Diseases > ... > American Cancer Society  
[www.cancer.org/- 36k](#) - [Cached](#) - [Similar pages](#)

**ACS : Error**  
 ... avoid seeing this message again. Help | About ACS | Legal & Privacy  
 Information Copyright 2002 © American Cancer Society, Inc.  
[www.cancer.org/cancerinfo/main\\_cont.asp?st=ds&ct=5](#) - 14k - [Cached](#) - [Similar pages](#)  
 [ More results from [www.cancer.org](#) ]

**Cancer.gov**  
 text only skip navigation. National Cancer Institute, Cancer.gov, ... home, cancer information clinical trials statistics research programs research funding about NCI ...  
 Description: The US National Institute of Health's cancer research wing.  
 Category: Health > Conditions and Diseases > ... > Institutional  
[www.nci.nih.gov/- 30k](#) - [Cached](#) - [Similar pages](#)

Figure 9.17: \* cancer search results



## Double Words

An interesting but not necessarily very useful quirk of Google's search engine is the differing results that you get when searching on double words. What is a double word search? Simply, a double word search is entering the same word twice in the query box. As an example, if we search on the single keyword "cat," the first search result is "Caterpillar-Heavy Equipment & Engine Manufacturer." If we search on the keywords "cat cat," (without the quotes) the first search result is "HotAIR – Feline Reactions to Bearded Men." Interestingly, regardless of the method of search (one "cat" or two) the search results appear in the same category, namely "Recreation > Pets > Cats > You and Your Cat." Scanning through the remainder of the results for both single and double word searches reveals that all of the results pertain to felines. But the order of each web page differs between the two searches.

## Right-Click

If you have the Google Toolbar installed, when you right-click in your browser, the resulting menu includes several Google search operators:

- Cached Snapshot of Page
- Similar Pages
- Backward Links
- Translate Page

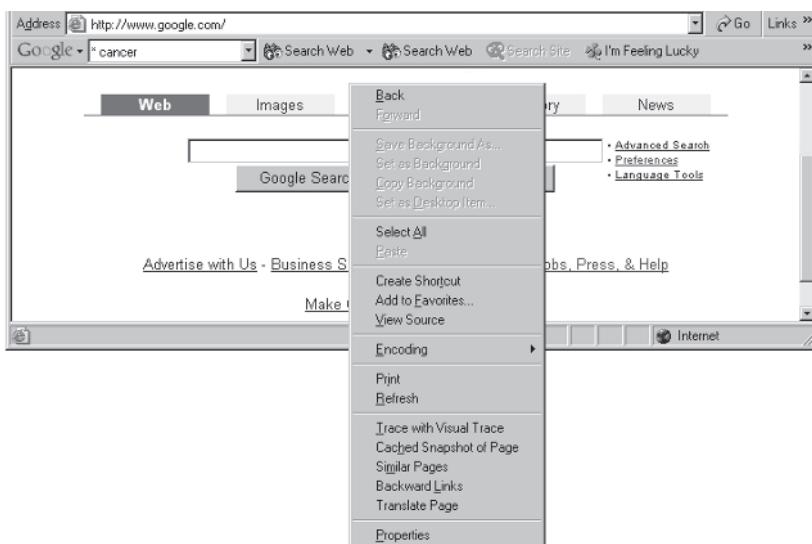


Figure 9.18: Browser menu

We have discussed each of these Google features in previous chapters. Check this menu out from time to time. Google may add additional menu items.

## Miscellaneous Google Information

---

### Web APIs

Web APIs are software modules or functions used by web developers to add some Google functionality into their web page/site. You need to be an experienced web developer to even determine if you need to use an API, much less determine how they work. Suffice it to say, they are available and useful if you are a web developer.



## Webmaster's Information

If you are a webmaster, Google has a web page designed to help you with certain functions, such as "How do I get my site listed on Google?" Go to <http://www.google.com/webmasters/index.html> to determine if you are maximizing your web site's potential.

## Submit Your URL

You are the proud owner of a web page telling the world about your trip across the deserts of the Najd last summer, and you want to drive as many people to the web page as possible. First, you need to submit your URL to Google. To do that, go to <http://www.google.com/addurl.html>. Good luck!

A friend of mine from Hope, Arkansas (Hi, Cindi! Didn't know I was going to mention you in a book, eh? Tell your buds at the Hope hospital you really do know a bona fide author.), recently mentioned to me that there was a vast, unseen ocean of web pages out there in cyberland that no one could find. Well, since I was writing this book and I take just a little professional pride in being able to find whatever I need on the Internet, I boldly asked her to challenge me, and I would find whatever was there, seen or unseen. By golly, she stumped me! I assumed (we know all about assume, don't we?) Google had indexed every web page there is on the web. But Google has not indexed every web page. And if Google has not indexed the page, we ain't gonna find it (using Google)! So, Google's spider needs to take shorter lunch breaks and work harder!

Cindi pointed out a person's home page (a longtime friend of hers) that has been on the web for two years. Google's spider has never visited the page and the web page owner has never submitted the URL to Open Directory Project. So, even though I eventually knew (Cindi told me) the URL and title, I still could not find the page using Google. The important message here is, if you own a web page, you must submit your URL to Google (Open Directory Project) to ensure your page can be found. There is no cost to have



your URL listed. It is free! Cindi said her friend thought he must pay to have his site listed with Google since he gets (spam) e-mails saying a service will list his home page with 400 search engines for \$19.95. Most search engine directories do not charge a fee for site owners to list their sites. You can do it for free. But, if you want me (or any other “for hire” out-of-work, formerly productive engineers) to list your site with 400 search engines *for you* (using my automated software), then I am going to charge something for the service.

## Advertising Programs

To advertise with Google, visit <http://www.google.com/ads/index.html>. The page is self-explanatory.

# Organizing Your Google Bookmarks

---

This discussion is applicable to Internet Explorer, version 6. If you use another browser, the general information is applicable, but the specific steps to reorganize your bookmarks may be different. If this is the case, you should reference your browser’s Help menu. For anyone who does not know what a browser is, I will explain. A browser is the guy who walks into your kitchen, opens the refrigerator door, and then opens the various containers to see what is inside. Okay, I’m just joking. However, the metaphor is very applicable to computer browsers. A computer browser is the computer or software program that you open (the fridge door) and use to view web sites (the food containers).

By now, you should have numerous Google sites bookmarked. If you are like me, you just click on Favorites without giving any thought to their organization. Now, you probably have a dozen or more bookmarks that appear at the bottom of the browser’s Favorites menu. See Figure 9.19. As we add each bookmark, it is placed at the bottom of the menu, lengthening the menu until it becomes



unwieldy. If we have bookmarks from other sites (I do but did not show them in the figure), the Favorites menu can easily become very cumbersome to use. Besides lengthening the menu, the bookmarks do not appear alphabetically, and some are prefixed with “Google,” while others are not. I like consistency. Either all of the bookmarks should be prefixed with “Google” or none should be. I like for my bookmarks to be arranged alphabetically. How do we place these bookmarks into a folder and change them to be consistent and in alphabetical order? That is the subject of this section.

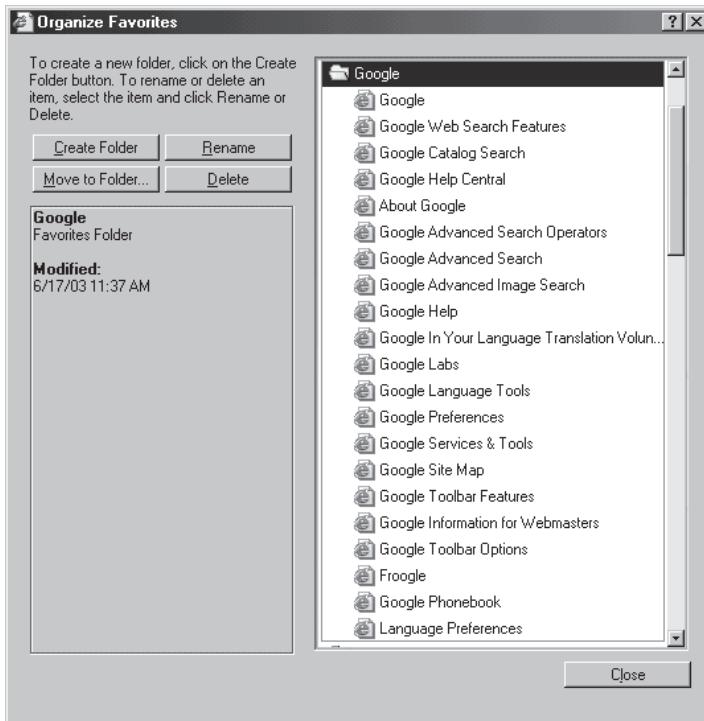


Figure 9.19: Google bookmarks

We begin by left-clicking on the browser’s Favorites menu item and then selecting Organize Favorites. See Figure 9.20.



Figure 9.20: Organize Favorites menu option



After selecting Favorites | Organize Favorites, the Organize Favorites dialog box will appear. See Figure 9.21.

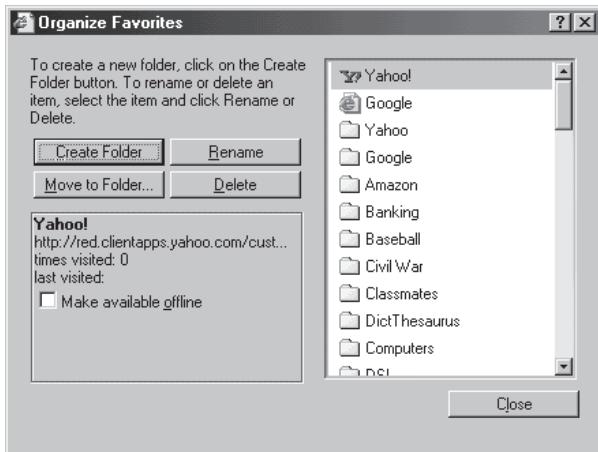


Figure 9.21: Organize Favorites dialog

Our objective is to create a folder named Google and then move all of the Google bookmarks into it. Next we organize the bookmarks alphabetically. Finally, we rename some of the bookmarks. Let's get started!

1. Click on Create Folder. The result is a new folder at the bottom of the bookmarks menu displayed on the right side of the Organize Favorites dialog box. See Figure 9.22. The new folder is highlighted in a text box, and the cursor is blinking. This signifies that the text in the box can be changed. Who wants a folder named “New Folder”? Not me! Let’s change it.

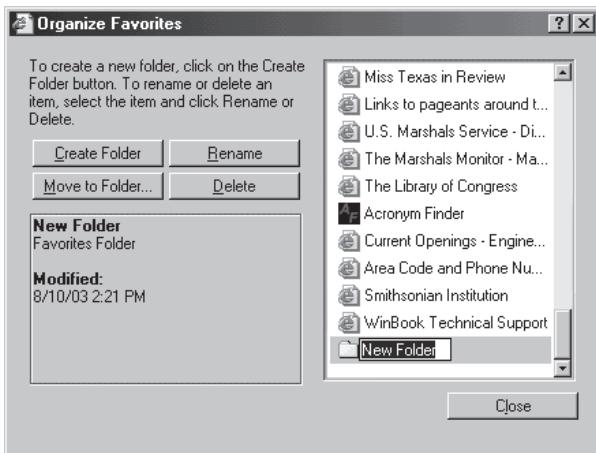


Figure 9.22: Organize Favorites dialog

2. Press the Backspace key. Notice the “New Folder” text disappears, leaving a blinking cursor in the box. See Figure 9.23.



#### Note:

If you pressed a key that did not result in the text disappearing, then you have fixed the name of the folder as “New Folder.” Now you will have to left-click on the Rename button and start this step over.

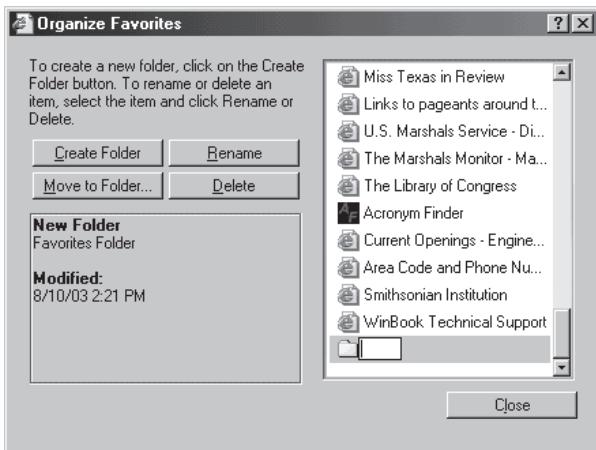


Figure 9.23: Renaming the new folder



### 3. Key in Google.

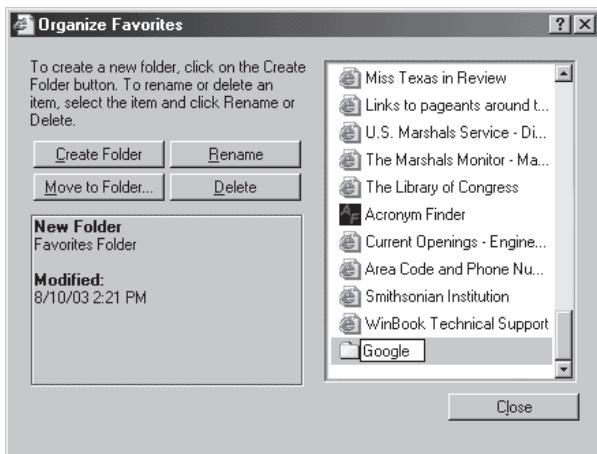


Figure 9.24: Renaming the new folder Google

4. Press the Enter key. This will fix the name of the folder. Notice that you now have a folder named Google at the bottom of your Favorites menu.

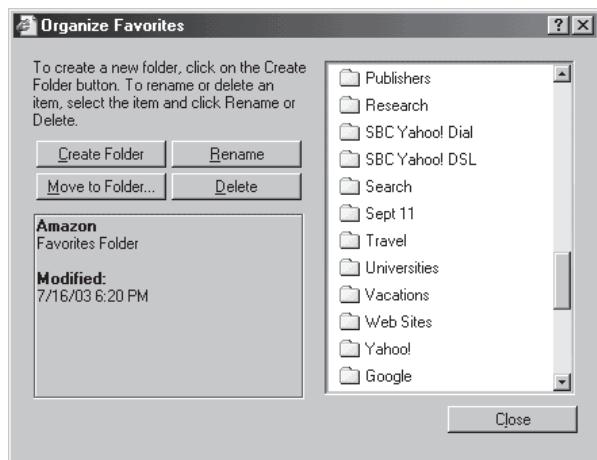


Figure 9.25: The new Google folder



Let's now add the Google bookmarks to the Google folder.

5. Left-click on the bookmark just above the Google folder. Continuing to hold down the mouse button, move the bookmark over the folder icon, as shown in Figure 9.26. Once you have moved the bookmark over the folder, release the mouse button. Voilà — the bookmark should now be inside the Google folder. To verify that it is, double left-click on the Google file icon. The file will open, and there the bookmark is. See Figure 9.27. Now left-click on the Google folder icon to close the folder.

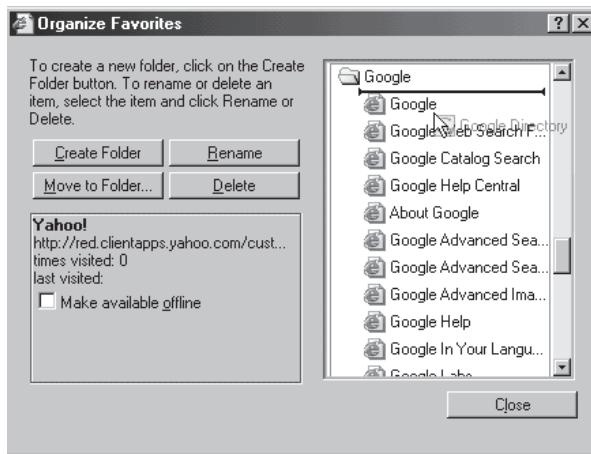


Figure 9.26: Moving a bookmark to the Google folder

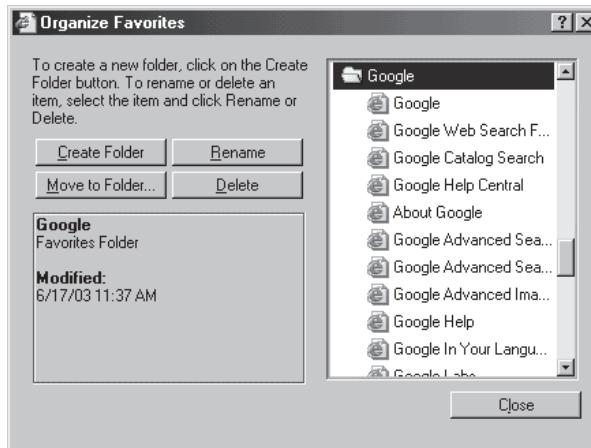


Figure 9.27: After moving a bookmark to the Google folder



6. Repeat the above procedure for every Google bookmark. Do not be concerned about the alphabetical arrangement of the bookmarks right now. We arrange them alphabetically in a moment.
7. Do you have them all in the Google folder? Okay, let's move the Google folder up the folder ladder to a place that suits us. Since I use Google often throughout the day, I like to have my Google folder not in alphabetical order as my other folders are but at the very top of all folders. Then I can access it quickly. So, let's move it up. Left-click on the folder itself, not on the folder name, and while holding the mouse button down, drag the folder to the top of the bookmarks menu; or if it suits you, drop it off at its alphabetical place in your menu by releasing the left mouse button when you have reached the place in the menu that you prefer. The images in this book show it placed at the top of the menu. Notice that when you move the folder you see a black line just above the mouse icon, as shown in Figure 9.28. The line shows you where the folder will reside when you release the mouse button.

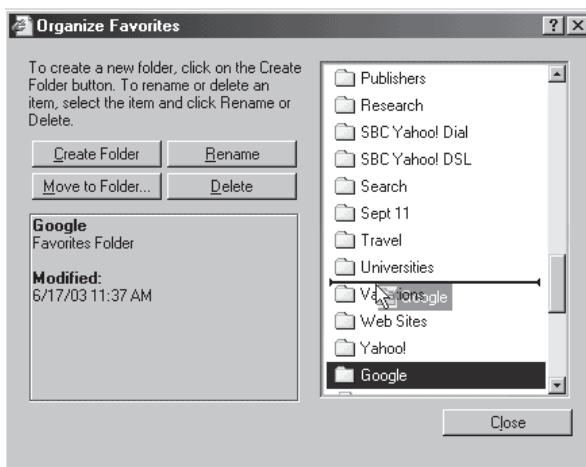


Figure 9.28: After moving a bookmark to the Google folder



8. If you pause over another folder, it will open, and if you release the left mouse button while the file icon is inside the folder, can you guess where your Google folder will disappear to? Right — inside the folder you paused over. But just grab the Google icon again, keep moving it up until you get to the position you desire, then release the mouse button.

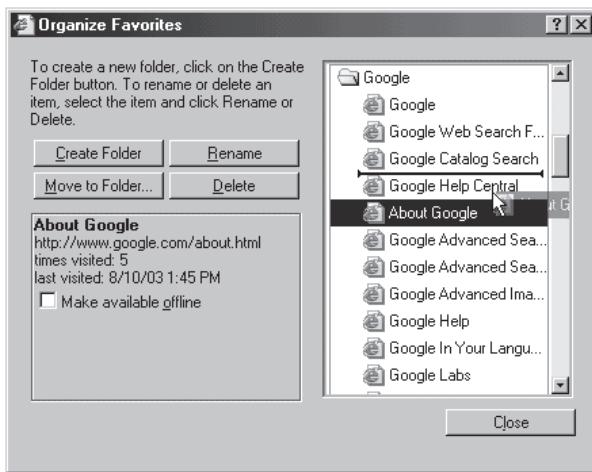


Figure 9.29: Alphabetizing Google bookmarks

9. We are now ready to move the bookmarks within the Google folder and arrange them alphabetically. Left-click on the top-most bookmark (assuming it is not already in alphabetical order) and drag it down to the approximate position where it will be when you are finished. Continue moving the bookmarks up or down until you have them all in alphabetical order. Next, let's rename a couple of bookmarks.
10. You may notice that most of the bookmarks are prefixed with the word "Google." I like consistency, so let's rename the one or two that do not include "Google" to include it. Left-click on the bookmark of interest to highlight it. Then left-click on the Rename button on the left side of the Organize Favorites dialog. Press the left arrow key to move the cursor over to the far left side. Then key in the word Google and a space. Left-click on the bookmark again to fix the name change. We are done! You



are now organized and ready to search the web. I hope you have as much fun as I do.

## Your Security

---

There has been much written about hackers and Internet security. The primary focus of such articles and discussions has been the issue of identity theft. There is much material written on that subject, so I do not discuss it here. Besides, what does identity theft have to do with Google anyway? Well... I want to bring up an Internet security issue that is getting little or no press attention — your personal physical security. What does that have to do with Google? Bear with me and you will see.

Since 9/11/2001, the American government has enacted new laws and regulations regarding its ability to monitor, collect, and use for, among other things, prosecutorial purposes what many of us would consider personal and private information. Included in this knee-jerk reaction to 9/11 is the federal government's ability to examine our lives and require libraries to hand over the book-checkout records of patrons. The government can also order, without judicial review (i.e., no search warrant), libraries and other organizations, such as web hosting services, to turn over records, even computers, for purposes that can only be described as fishing expeditions for supposed terrorists. Why should this concern law-abiding you? Ask any law-abiding citizen who has been falsely accused by the government of a crime.

Most people do not realize it, but everything you do on the Internet is recorded. Everything. Period. Every web page you visit, every image you look at, every e-mail you send or receive. Absolutely everything is recorded. By whom? By every Internet service provider and every web hosting service. The records of every transaction conducted by every computer connected to the Internet is recorded in something called the Management Information Database of the computer. This information about your Internet habits is



used by service providers and web hosting companies for various reasons, usually associated with keeping their service slim and trim.

Because the traffic passing through an Internet service provider/web host can be voluminous, each one has its own rules regarding how long the data exists and therefore how long it is recoverable. But the information is there. The government knows it is there. The government can easily tap into it via the mandatory software loads now required in electronic switches, gateways, and routers. In other words, if you are concerned about privacy issues and specifically your privacy, you better not get on the Internet from home. If you get on from a third-party site, such as a library or an Internet café, you better use an alias.

There is another security issue involving searching the Internet. You can innocently click on a link and suddenly be whisked away to a child porn site. Or a click can download a child porn image onto your computer. The government has convicted a person for as little as one child porn image on a computer. If you should inadvertently wind up with something on your computer that you do not want on it, delete the offending file or image. Next, empty the Recycle Bin. Then run ScanDisk to ensure that your file system is intact and fix any errors as suggested by ScanDisk. Finally, run Disk Defragmenter. ScanDisk and Disk Defragmenter are found in Start | Programs | Accessories | System Tools on Windows operating systems. Performing this process removes any residual evidence of your inadvertent download of a file or image.

A doctor in Turkey hired some programmers to write a program for him that he loads up on various newsgroups. This malicious software, disguised as an image, searches your computer for images and sends a report back to the good doctor telling him what images are on your computer. The good doc performs this public service ostensibly to catch people engaged in the trafficking of child pornography. Unsuspecting people from every corner of the globe fall under the doc's watchful eye. What the good doc can do, the government can do and do even better; you will never know until they place the handcuffs on you. So, to protect yourself, use a firewall



and virus protection software and do not open anything — files or images — without all of your protective mechanisms in place.

By the way, if the doc can read your images, he can also read any personal information on your computer, such as passwords and credit card numbers. We should support legislation in this country making it illegal for anyone — private, company, or government — to enter into your computer without your *express* permission. In the meantime, go to Zone Lab's web page (<http://www.zonelabs.com/store/content/home.jsp>) and download and install the company's Internet firewall, ZoneAlarm (from <http://www.zonelabs.com/store/content/home.jsp>). Click on the Download button. It is free (at the time of publication) and it is very effective.

## The Search Challenge

Do you recall my search challenge from Chapter 1? You may remember that I mentioned that there was only one time I have failed to achieve search goals when searching the Internet. I have never been able to determine the names of the three Rockwell International employees assassinated in Iran in 1976. If you feel up to a challenge, and you want to flex those new search skills, see if you can find their names.

In the next edition of this book, I tout the first five people who e-mail me a link containing the employees' names. I must be able to verify their names. If you participate in this challenge and you are one of the first five to e-mail me the correct information, I will ask for a short biography, including your search interests and how this book helped you with your search needs. Individuals at Rockwell International are ineligible to participate, as are the employees of the search engine companies. We want the challenge to be fair to the majority of people. I am the sole judge of the contest, and I am the sole arbiter of any dispute. My decision will be final, and there is no appeal to anyone else. You may e-mail me at [Michael\\_Busby@yahoo.com](mailto:Michael_Busby@yahoo.com). Good searching!



## Summary

In this chapter we discussed various topics of both general and special interest. We covered some neat Google features and some special tricks and treats. Everything we have learned will not necessarily remain long in our memory, but we will remember where the information is — in this book. So keep it as a handy reference, which I am certain you will refer to repeatedly over time. Otherwise, it makes a good doorstop. I strongly suspect, now that you are armed with powerful search weapons, you will find searching the web as entertaining as I do. Good luck information hunting!

We now possess all of the tools we need to find our “golden fleece” and we have almost completed our quest. But there remains one more “unexplored land” to visit before we can return to our homeland victorious with the spoils of questing — our knowledge of searching the Internet. Onward to Google Groups!

This page intentionally left blank.

# Chapter 10

## Searching Newsgroups

### Google Groups

---

We previously mentioned that Google Groups is Google's incarnation of the former Deja News Usenet forum, also called newsgroups. In this chapter we examine Google Groups features that include not only searching the huge Usenet archive but also posting to Usenet and removing your posts. The Google Groups home page is shown in Figure 10.1. The Google Groups site URL is <http://www.google.com/grphp>. (Bookmark it.) The Google Groups FAQ page is <http://www.google.com/groups/help.html>. (Don't bookmark it.)



Figure 10.1: Google Groups home page

Newsgroups were originally the domain of nerds and techies. In the early days (circa 1981), the primary purpose of the Usenet was to share a common technical database, and many of the postings were related to the solution of problems that ultimately yielded today's Internet and World Wide Web. As the technology evolved, Usenet groups became less focused on technical issues, broadening the subjects to include political and social commentary. As newsgroups evolved, the primary users became less technically oriented, such as engineers and scientists, and more general profession oriented. Newsgroups still seem to be the domain of the post-university crowd.

Many discussion groups are set up to discuss industry and job specific issues and share images, music, multimedia content, and videos. Self-help groups abound, as do groups promoting specific political ideologies and social agendas. Regardless of your sentiments, if you want to be entertained for a very small investment,



you can find a group that can have you rolling on the floor. You can also find groups that will have your blood boiling. Newsgroups are the epitome of freedom of information. More music and video files and other copyrighted material, such as electronic books, have been exchanged over Usenet than Napster could ever possibly accommodate. I cannot determine if Usenet is a well-kept secret or the entertainment and publishing industry has turned a blind eye to the wholesale exchange of pirated material.

Google Groups consists of the entire Usenet archive that dates to 1981. Using Google's search engine, you can search the enormous discussion groups with great speed and efficiency. The archive contains over 700 million posts. This immense database of viewpoints and human endeavor will undoubtedly be of great value to historians, psychologists, sociologists, and cultural anthropologists in the coming millennium. For now, it is of value to anyone with a question in search of an answer.

Google Groups is managed by individual Internet service providers. When a post is made to a group, the life of the post is determined not by Usenet or Google, but by the Internet service provider upon whose machine the post is maintained. One Internet service provider may keep posts for only a day or so, while other service providers may keep posts for a week. A query to your service provider will unveil the length that the service provider will keep postings.

One aspect of Google Groups and the Google search engine ranking algorithm is worth mentioning. Recall that PageRank ranks pages by the number of web pages that it determines a page is linked to. The thinking goes that the greater the number of pages linked to the page of interest, the greater the popularity of the page. Here is a bit of useful information for web page developers: PageRank also includes in its link popularity algorithm any link mentioned in a discussion group.

You can access the discussion groups via a "reader." Most e-mailers, such as Outlook Express, now have the ability to read newsgroups, as do some browsers. I have used Forte's Free Agent



for about ten years. The newsreader is free and can be used as a POP e-mail client. If Free Agent is used as an e-mail client, you can get rid of all those nasty viruses written for Microsoft's products. You can find Free Agent at <http://www.forteinc.com/agent/index.php>. I recommend Free Agent for one primary reason: Free Agent does not impose any censorship or restrictions on what you may view or read on Usenet.

If you are accessing the Internet with a newsreader and a slow modem, you will find that it takes some amount of time to download the 55,000+ groups. When you download the groups, you are downloading the "headers" (subjects) of the messages posted to the groups. To view the actual material, you must download the header file. These files can be very lengthy. However, with Google Groups, you do not download all 55,000+ groups. You only access the specific group of interest. If you are working with a slow modem, this is the best way to access Usenet.

---

**Note:**

Not all newsgroups are present on all ISP servers. My ISP only carries about 39,000 newsgroups.

It also takes a lot of time to download images, multimedia, electronic books, songs, and videos. If you are keen to get the latest music album free, it would be wise for you to invest in a DSL or cable modem. I recommend a DSL modem.

I want to make sure that this point is clear: You can access Usenet groups from numerous e-mailers, browsers, and newsreaders. However, you can only access Google groups via Google Groups. In other words, Google Groups is another Usenet newsreader that can be accessed through a web browser, such as Explorer or Netscape.

As an example of using Google Groups to access a Usenet group, go to Google Groups and left-click on alt (see Figure 10.1). You will see the web page illustrated in Figure 10.2. Now left-click on alt.abuse-recovery. The web page you see will be similar to the one shown in Figure 10.3.



The screenshot shows a Google search interface for newsgroups. The address bar shows <http://groups.google.com/groups?group=alt&hl=en>. The search term 'cancer' is entered in the search field. The search results are for the 'alt' group, with the heading 'Group: alt'. A search bar at the top right contains 'alt\_0 .. aeifie' with a 'Go' button. Below the search bar is a navigation menu with links for 'Advanced Groups Search', 'Preferences', and 'Groups Help'. There are three radio button options: 'Search only in alt.\*' (unchecked), 'Search all groups' (checked), and 'Search the Web' (unchecked). The main results table has columns for 'Activity' and 'Group'. The first few results are:

Activity	Group	Activity	Group
	<a href="#">alt.0.* (1 group)</a>		<a href="#">alt.abide</a>
	<a href="#">alt.Qd</a>		<a href="#">alt.ablecommerce</a>
	<a href="#">alt.1.* (1 group)</a>		<a href="#">alt.abortion.* (1 group)</a>
	<a href="#">alt.12hr</a>		<a href="#">alt.abp.* (1 group)</a>
	<a href="#">alt.12step.* (3 groups)</a>		<a href="#">alt.abraham-lincoln</a>
	<a href="#">alt.12steps.* (5 groups)</a>		<a href="#">alt.abuse.* (4 groups)</a>
	<a href="#">alt.1997</a>		<a href="#">alt.abuse-recovery</a>
	<a href="#">alt.1d</a>		<a href="#">alt.accounting</a>

At the bottom right of the results area is a link 'Next 50 Groups >>'. The bottom of the window shows standard browser controls.

Figure 10.2: *alt* group

Google Groups combines posted files to individual groups into thread subjects. See Figure 10.3. The header you see on the Google Groups page is the latest posting of a group of related postings, ergo the use of the term “thread.” Power users will find Google Groups too regimented for everyday use. However, casual users will find Google Groups organized in an easy-to-use fashion.

The screenshot shows a Google search interface for newsgroups, similar to Figure 10.2. The address bar shows <http://groups.google.com/groups?hl=en&hl=&ie=UTF-8&safe=off&group=alt.abuse-recovery>. The search term 'cancer' is entered in the search field. The search results are for the 'alt.abuse-recovery' group, with the heading 'Group: alt.abuse-recovery'. A search bar at the top right contains 'alt.abuse-recovery' with a 'Go' button. Below the search bar is a navigation menu with links for 'Advanced Groups Search', 'Preferences', and 'Groups Help'. There are three radio button options: 'Search only in alt.abuse-recovery' (checked), 'Search all groups' (unchecked), and 'Search the Web' (unchecked). The main results table has columns for 'Date', 'Thread Subject', and 'Most Recent Poster'. The results are:

Date	Thread Subject	Most Recent Poster
Aug 7, 2003	<a href="#">trance like visions</a> (1 article)	markman
Aug 4, 2003	<a href="#">I Remember PaPa</a> (5 articles)	GoldenMan
Aug 1, 2003	<a href="#">That's cooking</a> (1 article)	Phil (the Extreme One)

At the top right of the results area is a link 'Post a new message to alt.abuse-recovery' and below it 'Next 100 threads >>'. The bottom of the window shows standard browser controls.

Figure 10.3: *alt.abuse-recovery*



The mechanics of “exact phrase” searching on Google Groups is somewhat different from using Google to search the web. Google Groups does not return results for common words, such as “the” and “and.” Therefore, to include those words in the search results, you must include a “+” before the word. As an example, if I wanted to find an exact match for “The Star Spangled Banner” (without the quotes) I would need to enter “+The Star Spangled Banner” in the search text box (without the quotes surrounding the text).

Google Groups gives you one great feature that no other news-reader or e-mailer has: It allows you to search the entire 55,000+ newsgroups either all together or individually. For someone seeking an answer to a compelling question, this feature is powerful and very useful.

The screenshot shows the "Advanced Groups Search" interface on the Google Groups website. The URL in the address bar is `http://groups.google.com/advanced_group_search?dq=&num=100&hl=en&lr=&ie=UTF-8&group=alt.abuse-reco`. The search term "cancer" is entered in the main search input field. The search results are set to 100 messages and sorted by relevance. The search button is labeled "Google Search".

**Find messages:**

- with **all** of the words:
- with **the exact phrase**:
- Sort by relevance
- with **at least one** of the words:
- without the words:

**Newsgroup:** Return only messages from the **newsgroup**  (Example: `rec.games.misc, comp.os.*, *linux*`)

**Subject:** Return only messages where the **subject** contains

**Author:** Return only messages where the **author** is

**Message ID:** Find the message with **message ID**  (Example: `moderated-rg-faq-1-983174581@swcp.com`)

**Language:** Return messages written in  any language

**Message Dates:**  
Return messages posted:  anytime   
 Return messages posted between  May  1981  and  Aug  2003

**SafeSearch:**  No filtering  Filter using SafeSearch

©2003 Google

Figure 10.4: Google Groups advanced search page



Google's Advanced Groups Search feature, illustrated in Figure 10.4, utilizes a special notation to restrict a search to a particular newsgroup. To restrict a search, use the \* symbol after the name of the newsgroup to which you want to restrict the search. For example, to restrict the search to alt.abuse-recovery and all of its subgroups, enter "alt.abuse-recovery\*" (without the quotes) into the Newsgroup search text box.

The screenshot shows a search interface for Google Groups. On the left, there is a label 'Newsgroup' followed by the text 'Return only messages from the newsgroup'. To the right is a search input field containing the query 'alt.abuse-recovery\*'. Below the input field, there is a note '(Example: rec.games.misc, comp.os.\*, \*linux\*)'.

Figure 10.5: Special search notation

Limiting your search scope is just as crucial when searching groups as it is when searching the web. A Google Groups search can return thousands of “hits” that are of little or no value to your ultimate search goal.

## Posting to Google Groups

You can use Google Groups to post to Usenet groups. There are two types of Usenet groups. One type is called “unmoderated,” and the other type is called “moderated.” The postings to unmoderated groups are uncensored, while “moderators” censor the postings to moderated groups. That is, a posting to an unmoderated group goes directly to that Usenet group, while a posting to a moderated group goes to a human who reads the post and does one of three things: sends a post on to the Usenet group without modification for immediate posting, edits the post before sending it on to the group, or sends the post back to the originator with some appropriate comment(s).

You can post to more than one group at a time. Such postings are called *crossposts*. Crossposts that include a mix of moderated and unmoderated groups will go to the groups that are moderated. It is up to the moderator(s) to determine if the posting will be sent on to



the unmoderated groups. So, the best policy is to not mix moderated and unmoderated postings.

You can usually distinguish moderated groups from unmoderated groups by the title of the newsgroup. Unmoderated groups usually do not have any modifier distinguishing them as unmoderated.

There are exceptions, such as alt.dads-rights.unmoderated. However, moderated newsgroups always have “moderated” as part of their title. An example of such a newsgroup might be alt.feminazis.moderated. A note of caution for the unwary and easily offended: Unmoderated newsgroups are the ultimate in freedom of speech. Personal attacks on people who post are common.

Let's use an example to illustrate how to post to a newsgroup via Google Groups.

To post to a Google group:

1. Go to <http://www.google.com/>.
2. Left-click on Groups.
3. Left-click on alt.
4. Left-click on alt.activism.
5. Left-click on Post a new message to alt.activism.

See Figure 10.6. Note the existence of numerous subgroups in alt.activism. You can click on one of the subgroups and either read or post messages to it.

Activity	Group	Activity	Group	Hide Groups
—	<a href="#">alt.activism.</a>	—	<a href="#">alt.activism.microsoft</a> * (1 group)	
—	<a href="#">alt.activism.children</a> * (4 groups)	—	<a href="#">alt.activism.noise</a> * (1 group)	
—	<a href="#">alt.activism.community</a>	—	<a href="#">alt.activism.noise-pollution</a>	
—	<a href="#">alt.activism.d</a>	—	<a href="#">alt.activism.peacefire</a>	
—	<a href="#">alt.activism.death-penalty</a>	—	<a href="#">alt.activism.student</a>	
—	<a href="#">alt.activism.drug-war</a> * (1 group)	—	<a href="#">alt.activism.youth-rights</a>	
—	<a href="#">alt.activism.latino-youth</a>			

[Post a new message to alt.activism](#)

Threads 1-100 of about 292,000 in alt.activism

Date	Thread Subject	Most Recent Poster
------	----------------	--------------------

[Next 100 threads >>](#)

Figure 10.6: Posting to Google Groups



Next, you will be asked to sign in to Google Groups, if you have not already done so. You will be directed to an accounts page, as shown in Figure 10.7. If you already have a Google account, go ahead and sign in. If you do not have a Google account, left-click the Sign up for your account now link. Figures 10.8 through 10.12 show the registration process.

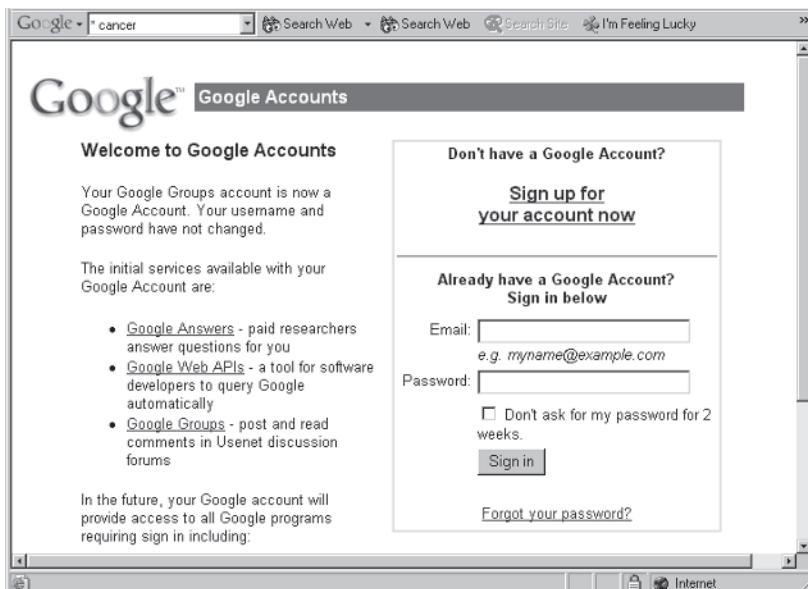


Figure 10.7: Google Accounts page



Figure 10.8: Signing up for a Google account

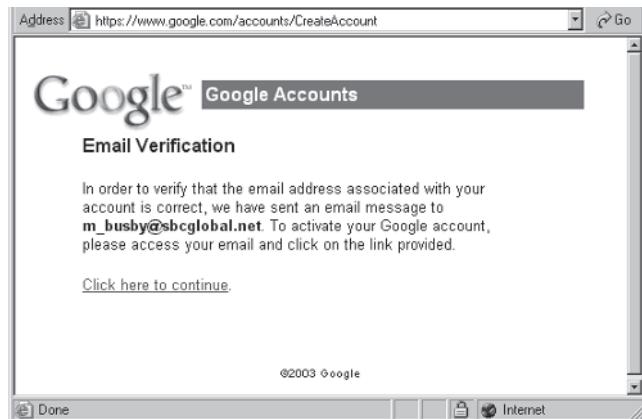


Figure 10.9: Signing up for a Google account

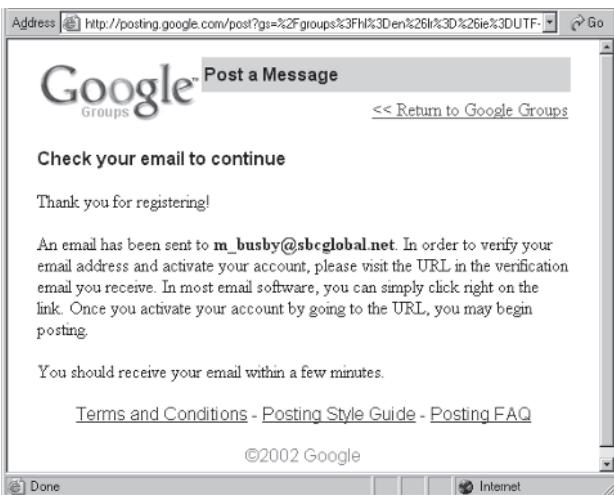


Figure 10.10: Signing up for a Google account

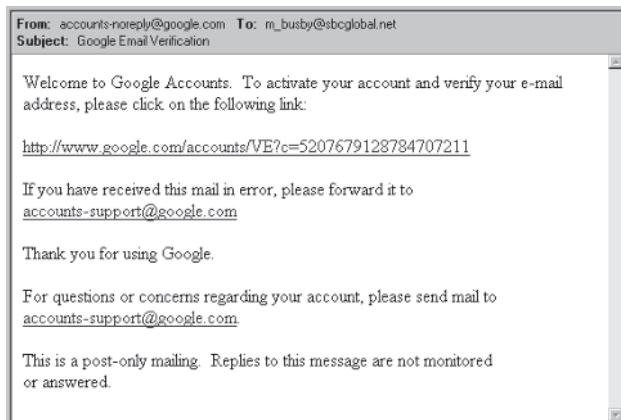


Figure 10.11: Signing up for a Google account

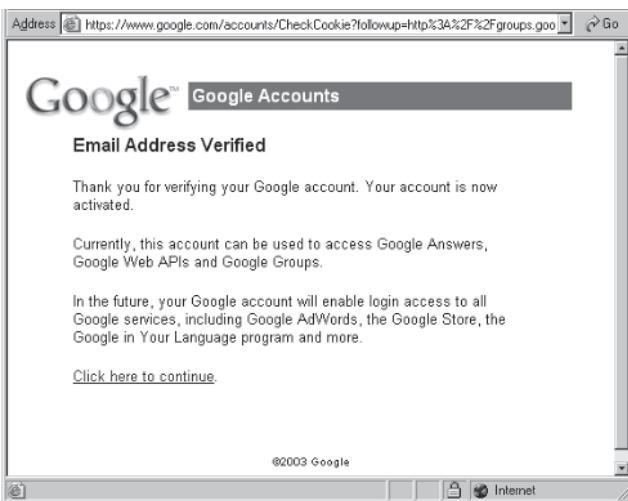


Figure 10.12: Signing up for a Google account

After left-clicking on *Click here to continue*, you are returned to the Google Groups home page. Now we start from the beginning again.

While we are discussing posts, an absolute inviolate “*do not do*” is never, ever post your e-mail address or any other personal information. Spammers have software that is constantly scouring the newsgroups for new e-mail addresses. (No, it does no good to add “remove” to your address, as the spammers’ software removes such trite attempts to avoid spam.) Additionally, if someone has your e-mail address, he or she can find you!

Here is some advice from a pro: Do not use Google Groups. Notice in Figure 10.13 that my e-mail address is displayed on the page that is used to post messages. See the top-left side of the figure. While my e-mail address will not show up in the body of the post (unless I deliberately place it there), the address will show up in the unseen (by typical newsreaders anyway) header information that I, or anyone else who knows how, can easily access with something as simple as WordPad. The underlying issue is Google Groups verifies the e-mail address with which you register. Google will send an e-mail to that address to confirm that it is a valid address so you



cannot opt out with a fake address. But other newsreaders, such as Free Agent, will let you enter any e-mail address you want. Free Agent does not verify your e-mail address. So, the choice is either use Google Groups with your e-mail address and get more spam or use Free Agent (or another newsreader) and get less spam. That is a very easy choice for me.

The screenshot shows a web browser window with the URL <http://posting.google.com/post?gs=%2Fgroups%3FH%3Den%26l%3D%26ie%3DUTF-8%26safe%3Doff%26group%3Delt.actr>. The page title is "Post a Message". The main content area is titled "Compose your message". It includes fields for "Your email address" (m\_busby@sbcglobal.net), "Send to Group(s)" (alt.activism), an unchecked checkbox for "email a copy to me", and a "Subject" field. Below these is a large text area for "Your message". At the bottom are buttons for "Preview message", "Post message - No preview", and "Cancel message". The status bar at the bottom right shows "Done" and "Internet".

Figure 10.13: Posting a Google Groups message

As a recent test of a person's ability to find information given only an e-mail address, a friend gave me the e-mail address of a woman I did not know. It took me only 20 minutes to acquire her telephone number, her street address, the name of her boyfriend who lived with her, her profession, and where she worked. Scary, is it not? We



tend to associate weird and predatory behavior with males, however females can be predatory also. I strongly recommend that no one put their e-mail address in a post to a newsgroup.

Don't believe how easy it is? I just went to a newsgroup that I have never been to before, alt.cooking-chat, and had the e-mail addresses of about 2,000 women in less than a minute without a single one of them needing to post their address! If Google wants to maximize its investment in Google Groups, Google must allow users to use an alias in the post header information. Until then, use Free Agent and use an alias. Alternately, get a "throwaway" e-mail address from Hotmail or Yahoo.

The archive is updated twice per day. All posts, except those containing binary content, are archived. Perhaps you do not want some researcher many years from now reading your comments. To prevent your post from being archived, include the "X-No-archive: yes" (without the quotes) text string in either the header or the first line of the message body. To be sure that the post is not archived, include the no-archive string in both places and do not include any additional text in the first line of the message body.

## Removing Your Post

---

Google will honor remove requests if you have posted a message to a newsgroup and you want it removed from the archive. If you have the same e-mail address that you used when you posted your message, you can remove the post yourself in an automated process. Go to Google Groups' automatic removal tool at <http://services.google.com:8882/urlconsole/controller>.



The screenshot shows a web browser window with the address bar containing `http://services.google.com:8882/urlconsole/controller?cmd=reload&lastcmd=login`. The main content area features the Google logo and a "Remove your URL or Google Groups Post" button. A "First time here?" link is present. Below it, instructions explain how to remove a URL or article by verifying an email address with a password. A note specifies that users must register with the email address from which they posted the message. To the right, a login form is displayed with fields for Email and Password, a "Login" button, and links for forgot password and account creation. The status bar at the bottom indicates "Internet".

Figure 10.14: Removing a post

Fill in the appropriate fields. Google will then e-mail you a link. After you click on the e-mailed link, you will find yourself in the Remove your URL or Google Groups Post web page. See Figure 10.15.



The screenshot shows a web page with a "Logout" link at the top right. Below it is a "Remove your URL or Google Groups Post" button. On the left, there's a "Options" section containing text about removing URLs and Usenet posts. On the right, there's a "Status" section showing "No current requests." Below the options, there's a list titled "Select only one:" with four items: "Remove pages, subdirectories or images using a robots.txt file.", "Remove a single page using meta tags.", "Remove an outdated link.", and "Remove your usenet posts from Google Groups.". At the bottom, there's a note: "Please send email to [googlebot@google.com](mailto:googlebot@google.com) with further questions or problems regarding the removal of your URL, or to [groups-support@google.com](mailto:groups-support@google.com) if you encounter problems removing your Google Groups posts."

Logout

Remove your URL or Google Groups Post

Options

You may remove your URL from Google's results or your Usenet posts from Google Groups. URLs and posts will typically be removed within 24 hours of a successfully submitted request. You may also review the status of submitted requests in the column to the right.

Select only one:

- Remove pages, subdirectories or images using a [robots.txt file.](#)  
Your robots.txt file need not be in the root directory.
- Remove a single page using [meta tags.](#)
- Remove an outdated link.
- Remove your usenet posts from [Google Groups.](#)

Please send email to [googlebot@google.com](mailto:googlebot@google.com) with further questions or problems regarding the removal of your URL, or to [groups-support@google.com](mailto:groups-support@google.com) if you encounter problems removing your Google Groups posts.

Figure 10.15: Google Groups post removal tool

To remove your post, click on the bottom option, Remove your usenet posts from Google Groups. When I clicked on the link, I was taken back to the login page. I logged in (again) and was then taken to the Removal page (Figure 10.15) again. When I clicked on Remove your usenet posts from Google Groups (for the second time), I was finally taken to the actual page where I could enter the information necessary to identify my post and remove it. See Figure 10.16.

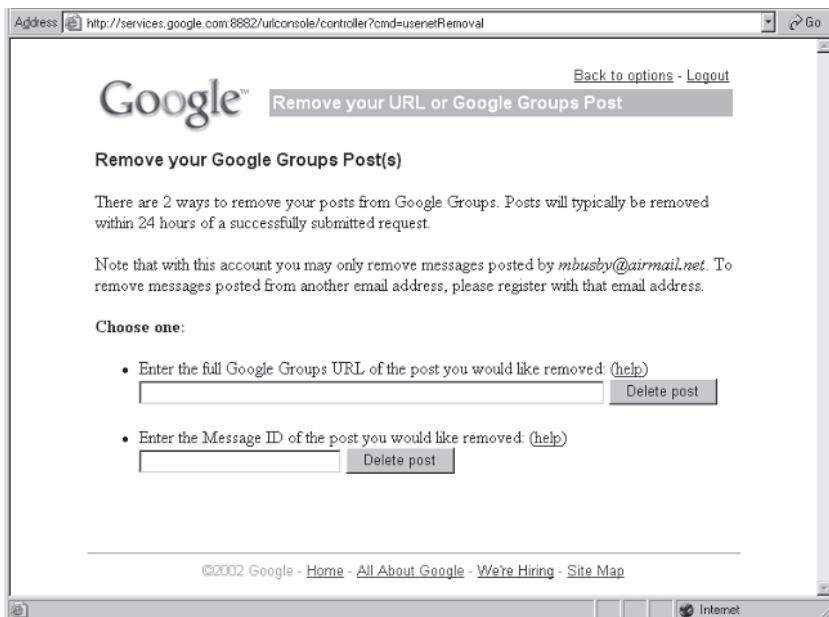


Figure 10.16: Removing a Google Groups post

At this point, enter the Message ID of the post you wish to remove from the archive. Then left-click on Delete post. If you entered the correct Message ID, your post will be deleted, and you will receive confirmation of the deletion.

If your current e-mail address is not the same one you used when you posted the message, you need to send an e-mail to groups-support@google.com requesting the post be removed. In your e-mail, include the following information:

- Your full name and appropriate contact information, including a verifiable e-mail address
- The Message ID for each individual message you want deleted
- The required statement “I swear under penalty of civil or criminal laws that I am the person who posted each of the foregoing messages or am authorized to request removal by the person who posted those messages”
- Your electronic signature



Google does not add new group topics to the Usenet. The Usenet community is responsible for adding new groups. Anyone may request that a new group be added to Usenet groups. Just practice what you are learning in this book by searching on “Starting a new newsgroup.” Include the text surrounded by quotes. You will find numerous sites offering advice and assistance.

Here is a tip for recruiters and salespeople who want to target their search to a specific company. You can do an @ symbol search. Determine the domain name of the company that you want to search. Then add the @ symbol as a prefix. Let’s say that we want to search for people at Lockheed. We would search for “@Lockheed.com” (without the quotation marks). See Figure 10.17. Our search results are shown in Figure 10.18. This is an excellent way to find new prospects at a targeted company. Impressive, eh? Try different company names and see what you get. Now, search on your e-mail address domain (the stuff on the right side of the @ character — be sure to include the @ character). Who says searching the web is frustrating? It’s becoming fun. But, remember,

The screenshot shows a Microsoft Internet Explorer window with the following details:

- Address Bar:** http://www.google.com/grphp?hl=en&tab=wg&ie=UTF-8&oe=UTF-8&q=@lockheed.com
- Google Groups Logo:** The main header features the "Google Groups" logo.
- Navigation Tabs:** Web, Images, Groups (highlighted), Directory, News.
- Search Bar:** Contains the query "@lockheed.com".
- Search Buttons:** Google Search, Advanced Groups Search, Preferences, Groups Help.
- Text Content:** "Post and read comments in Usenet discussion forums."  
List of newsgroups:
  - alt.** Any conceivable topic.
  - biz.** Business products, services, reviews...
  - comp.** Hardware, software, consumer info...
  - humanities.** Fine art, literature, philosophy...
  - misc.** Employment, health, and much more...
  - news.** Info about Usenet News...
  - rec.** Games, hobbies, sports...
  - sci.** Applied science, social science...
  - soc.** Social issues, culture...
  - talk.** Current issues and debates...[Browse complete list of groups...](#)
- Footer Links:** Advertise with Us - Business Solutions - Services & Tools - Jobs, Press, & Help
- Page Information:** ©2003 Google - Searching 800,000,000 messages
- Status Bar:** Internet

Figure 10.17: @lockheed.com search



if I can find you in Google Groups, so can your boss (or a friend, parent, relative, or significant other) if he or she bought this book. Maybe we need to rethink that post, eh?

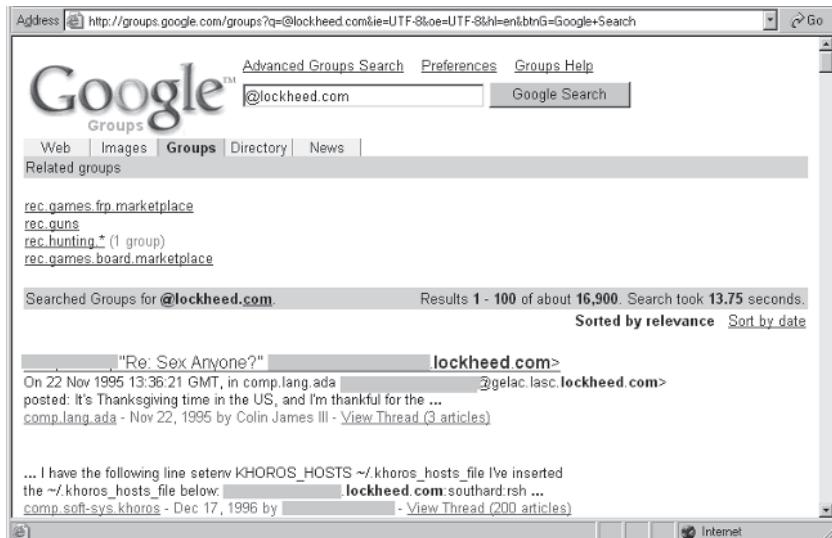


Figure 10.18: @lockheed.com search results sorted by relevance

The names of the individuals in Figure 10.18 have been grayed out in consideration of their privacy. However, Google Groups is a *public* forum. Some folks may wonder why I included this particular image of the @ search results. I included it so you can see for yourself that it is a public forum, and what you post will stay there forever (notice the date of the “Re: Sex Anyone?” post; it is November 22, 1995). The 16,900 search results are sorted by relevance. If we want more current e-mail addresses, then we left-click on the Sort by date link on the right side of the window just above the first search result. When we do that, we get the most current posts first. See Figure 10.19.



Figure 10.19: @lockheed.com search results sorted by date

You can see from Figure 10.19 that an @ search in Google Groups will yield a mine full of e-mail addresses. Unfortunately, Google does not support a character wildcard, or else a search on “@\*.com” would yield all e-mail addresses ever posted to Google Groups. But try that search from time to time and see if you get any useful results. It is reasonable and logical for Google to support a character wildcard some day.



## Summary

Google Groups is Google's gateway to the Usenet discussion groups, also called forums. Access Google Groups through Google's home page. These forums lie in the public domain, so whatever is posted there is available to the general public for reading and mining e-mail addresses or any other legal (and probably a few illegal) purposes. The newsgroup forums cover just about every conceivable topic imaginable, plus a few that are unimaginable. All of Google's search tools and operators are available for searching Google Groups.

This page intentionally left blank.

## **Afterword**

That is a wrap for *Learn Google*. I hope you have enjoyed your ramble through the world of Google and you return here often to discover new ways of improving your search skills. Be adventurous. Try new and odd combinations of Google’s tools and operators and discover new horizons. It is amazing what you can find with a little creativity and a small amount of time.

The purpose of this book was to help you zero in on your specific search goals, enabling you to find your “golden fleece,” and thereby reducing or eliminating your search frustration. I hope the purpose of this book was well served and you are now a happy searcher. If you have any unique or interesting search stories to share with me (for my next book), please contact me at [Michael\\_Busby@yahoo.com](mailto:Michael_Busby@yahoo.com).

I am a curious kind of individual. If I am standing at the beginning of something, I want to know what is at the end (or what is at the beginning, if I am standing at the end). Throughout the book, we have wondered what that last search result link was when our search resulted in millions of search results. What web page did that 3,749,561st link point to? Now, I can show you and satisfy your curiosity.



The End of the Internet (<http://www.shibumi.org/eoti.htm>)

Thanks to my lovely cousin, Karen Harwood, of Cottonwood, Arizona, for showing me where to find The End of the Internet. Both my quest and this book have now ended. Thank you for spending some time with me.

# Index

## A

About Google Toolbar menu option, 235, 248  
advanced alternate query search operators, 181-182  
advanced features, enabling, 242  
Advanced Image Search page, 209-210  
advanced search features, 144-145  
Advanced Search page, 146  
Advanced Search Page menu option, 235, 237  
advertising with Google, 284  
All About Google, 258  
allintitle: operator, 182, 198-199  
allinurl: operator, 182, 201-202  
AND operator, 134-135  
Archie, 17  
arithmetic operators, 130-133  
    minus operator, 130, 132-133  
    plus operator, 130-132  
    vs. Boolean operators, 136  
author-controlled search engines, 21-22  
AutoFill feature, 249  
    using, 251-254

## B

basic search, 123  
basic search operators, 124, 127-133  
    minus operator, 130, 132-133  
    plus operator, 130-132  
    quote operator, 127-130  
BlogThis feature, 249  
bookmarks, organizing, 284-292  
Boolean operations, complex, 136-139  
Boolean operators, 134-136  
    AND, 134-135  
    NOT, 134-135  
    OR, 134  
    vs. arithmetic operators, 136  
Browser Control, 245

browser, configuring to accept cookies, 260-262

browser right-click menu, adding options to, 239

browser window, opening for each search, 110, 118, 238

business, searching for, 274

## C

cache: operator, 181, 182-190  
cached pages, 183, 185-186  
Calculator feature, 268, 269-271  
catalogs, searching with Google, 50-57  
Category feature, enabling, 242  
Clear Search History menu option, 235, 246  
Close Program dialog box, 67  
colleges, searching for, 97-98  
color format, specifying for images, 223  
Coloration option, 210, 223  
Combined Search feature, 244-245  
Contact Us menu option, 235, 248  
cookies,  
    accepting, 260-261  
    accepting from trusted sites, 262

country codes, 10-11

country domain, finding, 170-171

credit card information, storing in Google, 252-254

crossposts, 303-304

## D

Date option, 144, 150  
date, specifying for search results, 144, 150  
default home page, setting Google as, 263-264  
default search engine, setting Google as, 264-268  
definitions, searching for, 88-90

- Dictionary Definitions feature, 268, 271-272  
Direct Hit, 22  
domain, 153  
domain names, 11  
Domain option, 144, 153-158, 210, 223  
domain search, 153-158  
    example, 169-175  
double word search, 281  
drag-and-drop functionality in Google Toolbar, 229
- E**  
editor-controlled search engines, 22  
experimental features, 244-245
- F**  
File Format option, 144, 148-149  
file format, specifying, 148-149  
File Transfer Protocol, *see* FTP  
file type, selecting for image, 222  
Filetypes option, 209, 222  
filtering options, 109, 114-117  
Find Results text boxes,  
    Advanced Image Search, 209, 210-220  
    Advanced Search, 144, 147-148  
Froogle, 39, 41-47  
frustration with web searching, 2-3  
    reasons for, 23-24  
FTP, 17  
    sites, 16-17
- G**  
global unification, 7  
Google, 29-30  
    improvements to, 32-33  
    ranking of search results, 31-32  
    setting as default home page, 263-264  
    setting as default search engine, 264-268  
    using to search, 37  
Google Accounts, registering for, 305-308  
Google Answers, 39, 47-50  
    menu option, 235, 236
- Google Catalogs, 39, 50-57  
    Advanced Search page, 54-57  
Google Compute, 40, 79-84  
Google Compute button,  
    enabling, 244  
    installing on Google Toolbar, 80-83  
Google General FAQ, 259  
Google Glossary, 40, 88-90  
Google Groups, 39, 58, 297, 299  
    and personal security, 308-310  
    menu option, 235, 236  
    removing a post using, 310-314  
    targeting specific company in, 314-316  
    using to post to Usenet groups, 303-310  
    using to search newsgroups, 300-303  
Google Help Central, 255-256  
Google home page, 100, 122  
Google Home Page menu option, 234, 236  
Google Image Search, 39, 58-74  
    Advanced Image Search Page, 68-69  
Google Images menu option, 234, 236  
Google in Your Language, 165-169  
Google Labs, 39-41, 74-94  
    Google Compute, 40, 79-84  
    Google Glossary, 40, 88-90  
    Google News Alerts, 40, 74-79  
    Google Sets, 40, 90-91  
    Google Viewer, 40, 85-86  
    Google WebQuotes, 40, 86-88  
    Keyboard Shortcuts, 40, 92-94  
    Voice Search, 40, 92  
Google Links menu option, 249  
Google News, 39, 94-95  
    menu option, 235, 236  
Google News Alerts, 40, 74-79  
Google News button, enabling, 243  
Google Search feature, 229-231  
Google Services, 37-39, 41  
    Froogle, 39, 41-47  
    Google Answers, 39, 47-50  
    Google Catalogs, 39, 50-57  
    Google Groups, 39, 58, 297, 299  
    Google Image Search, 39, 58-74

- Google Labs, 39-41, 74-94  
Google News, 39, 94-95  
Google Special Searches, 39, 95-97  
Google University Search, 39, 97-98  
Google Web Directory, 39, 99-100  
Google Web Search, 39, 100-103  
Google Wireless, 39, 104
- Google Sets, 40, 90-91  
Google Site Map, 256-257  
Google site, selecting, 238  
Google Special Searches, 39, 95-97  
Google Toolbar, 225  
    customizing, 237-246  
    displaying, 228  
    drag-and-drop functionality, 229  
    features, 229-230  
    help, 235, 246  
    installation, 80-83, 227-228  
    installation requirements, 226  
    menu, 229-230, 234-235  
    privacy policy, 226-227  
    returning to default condition, 245-246  
    right-click functionality, 229  
    updating, 248  
    version 2.0 features, 248-254
- Google University Search, 39, 97-98  
Google Viewer, 40, 85-86  
Google Web Directory, 39, 99-100  
    menu option, 235-236  
Google Web Search, 39, 100-103  
Google web search features, 268  
Google WebQuotes, 40, 86-88  
Google Wireless, 39, 104  
Gopher, 17  
Gopherspace, 17  
groups,  
    moderated, 303  
    unmoderated, 303
- H**  
help, *see* Google Help Central  
Help menu option, 235, 246  
Highlight feature, 230, 233-234  
    enabling, 243-244
- home page, 12  
    setting Google as default, 263-264
- host, 9
- HTML document, *see* web page
- hyperlink, 14
- I**  
I'm Feeling Lucky button, 101, 268, 272-273
- image searching, 58-74  
    advanced, 207-224
- image sizes, selecting, 220-222
- images, specifying file type, 67-74, 222
- info: operator, 181, 193
- Interface Language option, 108, 111
- Internet, 8, 16  
    and personal security, 292-294  
    compared to library, 7  
    end of, 319-320  
    popular activities on, 2
- Internet Explorer, configuring to accept cookies, 260-262
- Internet searching, *see* web searching
- Internet service provider, *see* ISP
- intitle: operator, 182, 199-201
- inurl: operator, 182, 202-206
- ISP, 8
- K**  
Keyboard Shortcuts, 40, 92-94
- keywords, 8, 21, 126  
    excluding, 130, 132-135, 148  
    importance of selecting, 125-127
- L**  
language,  
    setting for interface, 108, 111  
    setting for searches, 108-109, 112-114
- Language options, 144, 158-165
- Language Tools page, 159
- Language Tools Page menu option, 235, 237
- language translation, 159-165
- library, compared to Internet, 7
- link popularity, 20
- link: operator, 181, 190-191

- Links option, 144, 177  
links, checking, 190-191
- M**  
meta tags, 14-15  
metacrawlers, 21  
minus operator, 130, 132-133
- N**  
Navigation button, enabling, 245  
News Alerts, creating, 74-79  
newsgroups, 297-299 *see also Google Groups*  
    searching using Google Groups, 300-303  
NOT operator, 134-135  
Not related to the words text box, 219-220  
Number of Results option, 109, 118
- O**  
Occurrences option, 144, 150-153  
Open Directory Project, 99  
operators,  
    arithmetic, 130-133  
    Boolean, 134-136  
    comparing Boolean and arithmetic, 136  
    text, 127-130  
Options menu option, 249  
OR operator, 134  
Organize Favorites dialog, 285-291
- P**  
Page Info feature, 229-230, 233  
Page Info menu, enabling, 242-243  
PageRank feature, 31, 229, 232-233  
    enabling, 242  
paid placement, 22-23  
pecuniary-controlled search engines, 22-23  
person, searching for, 274-275  
PhoneBook, 268, 273-279  
    bookmarking, 277-278  
    removing information from, 278-279  
phrases, searching on, 127-130  
plus operator, 130-132
- Pop-up Blocker feature, 249  
portal, 16  
preferences,  
    saving, 119, 260  
    setting, 107-119  
Preferences page, 107-108  
Privacy Information menu option, 235, 247  
privacy policy for Google Toolbar, 226-227  
products, searching for with Google, 41-47
- Q**  
quality of results, 124-125  
query box, *see* search box  
query words, *see* keywords  
questions, getting answers to with Google, 47-50  
quote operator, 127-130
- R**  
rank-checking software, 19-20  
Related to all of the words text box, 212-215  
Related to any of the words text box, 218-219  
Related to the exact phrase text box, 216-217  
related: operator, 181, 191-192  
relevancy ranking, 23  
restricted domain search, 153-158  
    example, 169-175  
results, setting number of, 109, 118  
Results Window option, 110, 118  
right-click functionality in Google Toolbar, 229  
robots, 20-21
- S**  
SafeSearch filtering option, 109, 114-117, 144, 210, 224  
search box, 122  
    sizing, 239-240  
search buttons, displaying, 241-242  
search challenge, 3, 294

- search engine types,  
 author-controlled, 21-22  
 editor-controlled, 22  
 pecuniary-controlled, 22-23  
 user-controlled, 22
- search engines, 8, 16, 25  
 content not operated on, 26-27  
 content operated on, 25-26  
 development of, 18-19  
 effective, 29  
 future of, 33  
 how they work, 25-26  
 setting Google as default, 264-268
- search example, 169-175
- search history list, 231  
 clearing, 246  
 enabling, 240
- search integrity, 28
- Search Language option, 108-109,  
 112-114
- search modifiers, *see* search operators
- search operators,  
 advanced, 181-182  
 basic, 124
- Search Preferences Page menu option,  
 235, 237
- search results, setting number of, 109,  
 118
- Search Site feature, 229, 231
- search term button, enabling, 243-244
- search terms, specifying location of in  
 page, 144, 150-153
- search topics, most popular, 4-5
- search words, *see* keywords
- searches, types of, 123
- security  
 and Google Groups, 308-310  
 and the Internet, 292-294  
 concerns, 183-185
- shortcut keys, using in Google, 93
- Similar option, 144, 175-176
- simple search, 123-124
- site map, *see* Google Site Map
- site: operator, 182, 195-198
- site-specific search, 153-158
- Size option, 209, 220-222
- slide show, viewing search results as,  
 85-86
- Spell Checker feature, 268, 279
- spiders, 21
- sponsored links, 31
- stock: operator, 182, 194
- Street Maps feature, 268, 280
- T**
- Telenet, 17
- text operators, 127-130
- thumbnails, 212
- title, restricting search to, 198-201
- toolbar, *see* Google Toolbar
- toolbar button labels, displaying, 241
- Toolbar Help, 228
- Toolbar Options dialog, 250-252
- Toolbar Options menu option, 235, 237
- Topic-Specific Searches option, 144,  
 177-178
- Translate page, 160
- translation, 159-165
- U**
- Uniform Resource Locator, *see* URL
- Uninstall menu option, 235, 248
- universities, searching for, 97-98
- Up button, 230, 233  
 enabling, 243
- URL, 10, 15  
 restricting search to, 201-206  
 submitting to Google, 283-284
- Usenet, 298-299
- Usenet groups,  
 accessing using Google Groups,  
 300-301  
 posting to using Google Groups,  
 303-310
- user-controlled search engines, 22
- V**
- visitor popularity, 22
- Voice Search, 40, 92
- voting buttons, enabling, 243

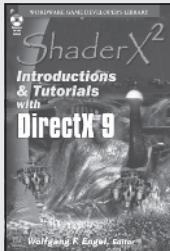
### W

- web, 8-9
  - address, 10
  - APIs, 282
  - page, 8, 12-14
- web search, using Google for, 100-103
- web searching, 2
  - frustration with, 2-3
  - reasons for, 6-8
  - time spent, 6
  - topics, 3-5
- web sites, 10
  - directory of, 99-100
  - finding comments about, 86-88
- WebCrawler, 17
- webmaster information, 283
- wildcard character, 280
- wireless services, using to connect to Google, 104
- With all of the words text box, 147
- With at least one of the words text box, 148
- With the exact phrase text box, 147
- Without the words text box, 148
- Word Find feature, 249
- World Wide Web, 8-10

# Looking for more?

Check out these and other titles from  
Wordware's complete list

## Game Developers Library



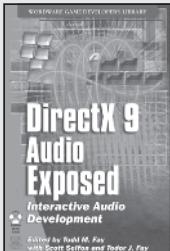
**ShaderX<sup>2</sup>: Introductions & Tutorials w/DirectX 9**

1-55622-902-X • \$44.95  
6 x 9 • 384 pp.



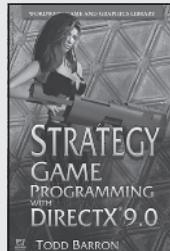
**ShaderX<sup>2</sup>: Shader Programming Tips & Tricks with DirectX 9**

1-55622-988-7 • \$59.95  
6 x 9 • 728 pp.



**DirectX 9 Audio Exposed:  
Interactive Audio Development**

1-55622-288-2 • \$59.95  
6 x 9 • 568 pp.



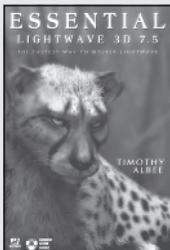
**Strategy Game Programming with DirectX 9.0**

1-55622-922-4 • \$59.95  
6 x 9 • 560 pp.

## Graphics Library

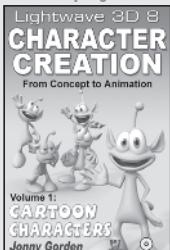


**LightWave 3D 7.5 Lighting**  
1-55622-354-4 • \$69.95  
6 x 9 • 496 pp.



**Essential LightWave 3D 7.5**  
1-55622-226-2 • \$44.95  
6 x 9 • 424 pp.

Available Spring 2004

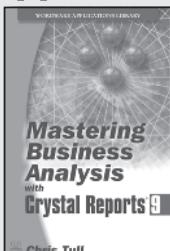


**LightWave 3D 8: Cartoon Character Creation**  
1-55622-083-9 • \$49.95  
6 x 9 • 550 pp.



**CGI Short Filmmaking:  
The Creation of Ghost Warrior**  
1-55622-227-0 • \$49.95  
9 x 7 • 300 pp.

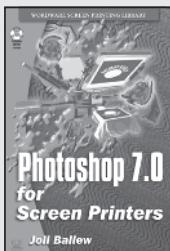
## Applications Library



**Mastering Business Analysis with Crystal Reports 9**  
1-55622-293-9 • \$44.95  
6 x 9 • 424 pp.



**Transac-SQL User-Defined Functions**  
1-55622-079-0 • \$49.95  
6 x 9 • 480 pp.

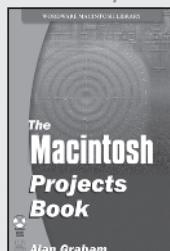


**Photoshop 7.0 for Screen Printers**  
1-55622-031-6 • \$59.95  
6 x 9 • 584 pp.



**Learn FileMaker Pro 6**  
1-55622-974-7 • \$39.95  
6 x 9 • 504 pp.

Available January 2004



**The Macintosh Projects Book**  
1-55622-228-9 • \$39.95  
6 x 9 • 400 pp.

Visit us online at [www.wordware.com](http://www.wordware.com) for more information. Use the following coupon code for online specials: **google-0383**