

SYNGRESS

1100110011001100011101
010101010100110011100
1100110011001100011101

GOOGLE HACKING FOR PENETRATION TESTERS

THIRD EDITION

Johnny Long | Bill Gardner | Justin Brown

Google Hacking for Penetration Testers

Third Edition

Page left intentionally blank

Google Hacking for Penetration Testers

Third Edition

Johnny Long
Bill Gardner
Justin Brown



AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Syngress is an imprint of Elsevier

SYNGRESS

Acquiring Editor: Chris Katsaropoulos
Editorial Project Manager: Anna Valutkevich
Project Manager: Punithavathy Govindaradjane
Designer: Matthew Limbert

Syngress is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA

Copyright © 2016, 2008, 2005 Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-12-802964-0

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

For information on all Syngress publications
visit our website at <http://store.elsevier.com/Syngress>



Contents

CHAPTER 1	Google Search Basics.....	1
	Introduction	1
	Exploring Google’s web-based interface	1
	Summary	17
	Fast track solutions.....	18
CHAPTER 2	Advanced Operators	21
	Introduction	21
	Operator syntax	22
	Troubleshooting your syntax	23
	Introducing Google’s advanced operators	24
	<i>Intitle</i> and <i>allintitle</i> : search within the title of a page	24
	<i>Allintext</i> : locate a string within the text of a page	27
	<i>Inurl</i> and <i>allinurl</i> : finding text in a URL.....	27
	<i>Site</i> : narrow search to specific sites	29
	<i>Filetype</i> : search for files of a specific type	30
	<i>Link</i> : search for links to a page	32
	<i>Inanchor</i> : locate text within link text.....	35
	<i>Cache</i> : show the cached version of a page.....	36
	<i>Numrange</i> : search for a number	36
	<i>Daterange</i> : search for pages published within a certain date range.....	36
	<i>Info</i> : show Google’s summary information	37
	<i>Related</i> : show related sites	38
	<i>Stocks</i> : search for stock information	38
	<i>Define</i> : show the definition of a term	39
	Colliding operators and bad search-fu	40
	Summary	42
	Fast track solutions.....	43
	Links to sites	45

CHAPTER 3	Google Hacking Basics.....	47
	Introduction	47
	Anonymity with caches.....	48
	Directory listings	51
	Locating directory listings.....	52
	Finding specific directories	52
	Finding specific files.....	53
	Server versioning	53
	Going out on a limb: traversal techniques.....	55
	Summary	58
	Fast track solutions.....	59
CHAPTER 4	Document Grinding and Database Digging	61
	Introduction	61
	Configuration files	61
	Locating files	65
	Log files	66
	Office documents.....	67
	Database digging.....	67
	Login portals	68
	Support files	68
	Error messages.....	69
	Database dumps.....	70
	Actual database files	71
	Automated grinding.....	71
	Summary	76
	Fast track solutions.....	76
CHAPTER 5	Google's Part in an Information Collection Framework	79
	Introduction	79
	The principles of automating searches.....	80
	The original search term.....	82
	Expanding search terms	82
	Using "special" operators	87
	Getting the data from the source	88
	Scraping it yourself: requesting and receiving responses	88
	Scraping it yourself: the butcher shop.....	94
	Using other search engines	102
	Parsing the data	102
	Domains and subdomains.....	107
	Telephone numbers.....	108
	Postprocessing	109

	Collecting search terms	113
	Summary	118
CHAPTER 6	Locating Exploits and Finding Targets	119
	Introduction	119
	Locating exploit code	119
	Locating exploits via common code strings	121
	Locating vulnerable targets	122
	Locating targets via source code	122
	Summary	122
CHAPTER 7	Ten Simple Security Searches That Work	125
	Introduction	125
	Site	125
	Intitle:index.of	126
	Error Warning	126
	Login Logon	128
	Username Userid Employee.ID \ "Your username is"	129
	Password Passcode "Your password is"	129
	Admin Administrator	130
	-Ext:html -ext:htm -ext:shtml -ext:asp -ext:php	132
	Inurl:temp inurl:tmp inurl:backup Inurl.bak	134
	Intranet Help.desk	134
	Summary	136
CHAPTER 8	Tracking Down Web Servers, Login Portals, and Network Hardware	137
	Introduction	137
	Locating and profiling web servers	138
	Locating login portals	149
	Using and locating various web utilities	151
	Targeting web-enabled network devices	156
	Locating network reports	156
	Locating network hardware	157
	Summary	158
CHAPTER 9	Usernames, Passwords, and Secret Stuff, Oh My!	161
	Introduction	161
	Searching for usernames	162
	Searching for passwords	163
	Searching for credit card numbers, social security numbers, and more	165
	Social security numbers	167

	Personal financial data	167
	Searching for other juicy info	167
	Summary	168
CHAPTER 10	Hacking Google Services.....	171
	Calendar	171
	Signaling alerts	172
	Google co-op.....	173
	Google's custom search engine	174
CHAPTER 11	Hacking Google Showcase	175
	Introduction	175
	Geek stuff	176
	Open network devices	179
	Open applications	186
	Cameras	191
	Telco gear	198
	Power.....	203
	Sensitive info	206
	Summary	207
CHAPTER 12	Protecting Yourself from Google Hackers	209
	Introduction	209
	A good solid security policy.....	209
	Web server safeguards.....	210
	Software default settings and programs	214
	Hacking your own site	214
	Wikto	215
	Advance dork	216
	Getting help from Google	216
	Summary	217
	Fast track solutions.....	217
	Links to sites	218
SUBJECT INDEX	219

Google Search Basics

INTRODUCTION

Google's Web interface is unmistakable. It is clean and simple. Its "look and feel" is copyright-protected for good reason. What most people fail to realize is that the interface is also extremely powerful. Throughout this book, we will see how you can use Google to uncover truly amazing things. However, as with most things in life, before you can run, you must learn to walk.

This chapter takes a look at the basics of Google searching. We begin by exploring the powerful Web-based interface that has made Google a household word. Even the most advanced Google users still rely on the Web-based interface for the majority of their day-to-day queries. Once we understand how to navigate and interpret the results from the various interfaces, we will explore basic search techniques.

Understanding basic search techniques will help us build a firm foundation on which to base more advanced queries. You will learn how to properly use the Boolean operators (AND, NOT, and OR), as well as explore the power and flexibility of grouping searches. You will also learn Google's unique implementation of several different wildcard characters. Finally, you will learn the syntax of Google's Uniform Resource Locator (URL) structure.

Learning the ins and outs of the Google URL structure will give you access to greater speed and flexibility when submitting a series of related Google searches. We will see that the Google URL structure provides excellent "shorthand" for exchanging interesting searches with friends and colleagues.

EXPLORING GOOGLE'S WEB-BASED INTERFACE

Google's Web Search Page

The main Google Web page, shown in [Figure 1.1](#), can be found at www.google.com. The interface is known for its clean lines, pleasingly uncluttered presentation and user-friendly layout.



FIGURE 1.1

Although the interface might seem relatively featureless at first glance, we will see that many different search functions can be performed right from the first page.

As shown in [Figure 1.1](#), there's only one place to type. This is the search field. In order to ask Google a question or query, you simply type what you're looking for, then either press Enter (if your browser supports it), or click the Google Search button to be taken to the results page for your query.

Google Web Results Page

After Google processes a search query, it displays a results page. This page lists the results of your search and provides links to the Web pages that contain your search text. The top part of the search result page mimics the main Web search page. Notice the Images, Video, News, Maps, and Gmail links at the top of the page. By clicking these links from a search page, you automatically resubmit your search as another type of search without having to retype your query.

The results line shows which results are displayed (1–10, in this case), the approximate total number of matches (here, over 8 million), the search query itself (including links to dictionary lookups of individual words), and the amount of time the query took to execute.

The speed of the query is often overlooked, but it is quite impressive. Even large queries resulting in millions of hits are returned within a fraction of a second. For each entry on the results page, Google lists the name of the site. This is followed by a summary of the site, usually with the first few lines of content, the URL of the page that matched, the size and date the page was last crawled, a

**FIGURE 1.2**

cached link that shows the page as it appeared when Google last crawled it, and a link to pages with similar content. If the result page is written in a language other than the default language, and Google supports the translation from that language to the default that is set in the preferences screen, a link titled “Translate this page” will appear, allowing you to read an approximation of that page in your own language (see [Figure 1.2](#)).

Google Groups

Due to the surge in popularity of Web-based discussion forums, blogs, mailing lists, and instant messaging technologies, the oldest of public discussion forums, USENET newsgroups, has become an overlooked form of online public discussion. Thousands of users still post to USENET on a daily basis. (A thorough discussion about what USENET encompasses can be found at www.faqs.org/faqs/usenet/what-is/part1/.) DejaNews (www.deja.com) was once considered the authoritative collection point for all past and present newsgroup messages until Google acquired [deja.com](http://www.deja.com) in February 2001 (see www.google.com/press/pressrel/pressrelease48.html). This acquisition gave users the ability to search the entire archive of USENET messages posted since 1995 via the simple and straightforward Google search interface. Google now refers to USENET groups as Google Groups.

Today, Internet users around the globe turn to Google Groups for general discussion and problem solving. It is very common for Information Technology (IT) practitioners to turn to Google's Groups section for answers to all sorts of technology-related issues. The old USENET community still thrives and flourishes behind the sleek interface of the Google Groups search engine.

The Google Groups search can be accessed by clicking the Groups tab of the main Google Web page, or by surfing to <http://groups.google.com>. The search interface (shown in [Figure 1.3](#)) looks quite different from other Google search pages, yet the search capabilities operate in much the same way. The major difference between the Groups search page and the Web search page lies in the newsgroup browsing links.

Google Image Search

The Google Image search feature allows you to search (at the time of this writing) over a billion graphic files that match your search criteria. Google will attempt to locate your search terms in the image filename, the image caption, the



FIGURE 1.3

text surrounding the image, and/or in other undisclosed locations to return a somewhat “de-duplicated” list of images that match your search criteria. The Google Image search operates identically to the Web search with the exception of a few of the advanced search terms, which we will discuss in the next chapter.

The page header looks familiar but contains a few additions unique to the search results page. The Moderate SafeSearch link below the search field allows you to enable or disable images that may be sexually explicit. The Showing dropdown box (located in the Results line) allows you to narrow image results by size. Below the header, each matching image is shown in a thumbnail view with the original resolution and size, followed by the name of the site that hosts the image.

Google Preferences

You can access the Preferences page by clicking the Preferences link from any Google search page or by browsing to www.google.com/preferences. These options primarily pertain to language and locality settings. The Interface Language option describes the language that Google will use when printing tips and informational messages. In addition, this setting controls the language of text printed on Google’s navigation items, such as buttons and links. Google assumes that the language you select here is your native language and will “speak” to you in this language whenever possible. Setting this option is not the same as using the translation features of Google (discussed in the following

**FIGURE 1.4**

section). Web pages written in French will still appear in French, regardless of what you select here.

To get an idea of how Google's Web pages would be altered by a change in the interface language, take a look at [Figure 1.4](#) to see Google's main page rendered in "hacker speak." In addition to changing this setting on the preferences screen, you can access all language specific Google interfaces directly from the Language Tools screen at www.google.com/language_tools.

By default, Google will always try to locate Web pages written in any language. Even though the main Google Web page is now rendered in "hacker speak," Google is still searching for Web pages written in any language. If you are interested in locating Web pages that are written in a particular language, modify the Search Language setting on the Google preferences page.

SafeSearch Filtering blocks explicit sexual content from appearing in Web searches.

Although this is a welcome option from day-to-day Web searching, this option should be disabled when you're performing searches as part of a vulnerability assessment. If sexually explicit content exists on a Web site whose primary content is not sexual in nature, the existence of this material may be of interest to the site owner.

The Number of Results setting describes how many results are displayed on each search result page. This option is highly subjective, based on your tastes and Internet connection speed. However, you may quickly discover that the default setting of 10 hits per page is simply not enough. If you're on a relatively fast connection, you should consider setting this to 100, the maximum number of results per page as shown in [Figure 1.5](#).



FIGURE 1.5

When checked, the Results Window setting opens search results in a new browser window. This setting is subjective based on your personal tastes. Checking or unchecking this option should have no ill effects unless your browser (or other software) detects the new window as a pop-up advertisement and blocks it. If you notice that your Google results pages are not displaying after you click the Search button, you might want to uncheck this setting in your Google preferences. As noted at the bottom of this page, these changes won't stick unless you have enabled cookies in your browser.

Language Tools

The Language Tools screen, accessed from the main Google page, offers several different utilities for locating and translating Web pages written in different languages. If you rarely search for Web pages written in other languages, it can become cumbersome to modify your preferences before performing this type of search. The first portion of the Language Tools screen allows you to perform a quick search for documents written in other languages, as well as documents located in other countries. The Language Tools screen also includes a utility that performs basic translation services.

The translation form allows you to paste a block of text from the clipboard or supply a Web address to a page that Google will translate into a variety of languages.

In addition to the translation options available from this screen, Google integrates translation options into the search results page. The translation options available from the search results page are based on the language options that are set from the Preferences screen. In other words, if your interface language is set to English, and a Web page listed in a search result is French, Google will

**FIGURE 1.6**

give you the option to translate that page into language of your preference, English. The list of available language translations is shown in [Figure 1.6](#).

Building Google Queries

Google query building is a process. There's really no such thing as an incorrect search. It's entirely possible to create an ineffective search, but with the explosive growth of the Internet and increasing size of Google's cache, a query that's inefficient today may just provide good results tomorrow – or next month, or next year. The idea behind effective Google searching is to get a firm grasp on the basic syntax and then to get a good grasp of effective narrowing techniques.

Learning the Google query syntax is the easy part. Learning to effectively narrow searches can take some time and requires a bit of practice. Eventually, it will become second nature to find the required information from the plethora of available Web sites.

The Golden Rules of Google Searching

Before we discuss Google searching, we should understand some of the basic ground rules:

Google Queries are not Case Sensitive

Google doesn't care if you type your query in lowercase letters (hackers), uppercase (HACKERS), camel case (hAcKeR), or psycho-case (haCKeR). The word is always regarded the same way. This is especially important when you're searching things such as source code listings, when the case of the term carries a great deal of meaning for the programmer. The one notable exception is the word "or." When used as the Boolean operator, "or" must be written in uppercase as OR.

Google Wildcards

Google's concept of wildcards is not the same as a programmer's concept of wildcards. Most consider wildcards to be either a symbolic representation of any single letter (UNIX users may think of the question mark), or any series of letters represented by an asterisk. This type of technique is called stemming.

Google's wildcard, the asterisk (*), represents nothing more than a single word in a search phrase. Using an asterisk at the beginning or end of a word will not provide you any more hits than using the word by itself.

Google Reserves the Right to Ignore You

Google ignores certain common words, characters, and single digits in a search. These are sometimes called stop words. According to Google's basic search document (www.google.com/help/basics.html), these words include where and how. However, Google does seem to include those words in a search. For example, a search for WHERE 1 = 1 returns less results than a search for 1 = 1. This is an indication that the WHERE is being included in the search. A search for where pig returns significantly less results than a simple search for pig, again an indication that Google does in fact include words like how and where. Sometimes Google will silently ignore these stop words. For example, a search for HOW 1 = WHERE 4 returns the same number of results as a query for 1 = WHERE 4. This seems to indicate that the word HOW is irrelevant to the search results, and that Google silently ignored the word. There are no obvious rules for the word exclusion, but sometimes when Google ignores a search term, a notification will appear on the results page just below the query box.

32-Word Limit

Google limits searches up to 32 words, which is up from the previous limit of 10 words. This includes search terms as well as advanced operators, which we'll discuss in a moment. While this is sufficient for most users, there are ways to get beyond that limit. One way is to replace some terms with the wildcard character (*). Google does not count the wildcard character as a search term, allowing you to extend your searches quite a bit. Consider a query for the wording of the beginning of the US Constitution: "We the people of the United States in order to form a more perfect union establish justice."

This search term is seventeen words long. If we replace some of the words with the asterisk (the wildcard character) and submit it as "we * people * * united states * order * form * more perfect * establish *" including the quote, Google sees this as a nine-word query with eight uncounted wildcard characters. We could extend our search even further by two more real words and just about any number of wildcards.

Basic Searching

Google searching is a process, the goal of which is to find information about a topic. The process begins with a basic search, which is modified in a variety of ways until only the pages of relevant information are returned. Google's ranking technology helps this process along by placing the highest-ranking pages on the first results page. The details of this ranking system are complex and somewhat speculative, but it suffices to say that for our purposes. Google rarely gives us exactly what we need following a single search.

The simplest Google query consists of a single word or a combination of individual words typed into the search interface. Some basic word searches could include:

- hacker
- FBI hacker Mitnick
- mad hacker dpak

Slightly more complex than a word search is a phrase search. A phrase is a group of words enclosed in double-quote marks. When Google encounters a phrase, it searches for all words in that phrase in the exact order you provide them. Google does not exclude common words found in a phrase. Phrase searches can include:

- "Google hacker"
- "adult humor"
- "Carolina gets pwnt"

Phrase and word searches can be combined and used with advanced operators, as we will see in the next chapter.

Using Boolean Operators and Special Characters

More advanced than basic word searches, phrase searches are still a basic form of a Google query. To perform advanced queries, it is necessary to understand the Boolean operators AND, OR, and NOT. To properly segment the various parts of an advanced Google query, we must also explore visual grouping techniques that use the parenthesis characters. Finally, we will combine these techniques with certain special characters that may serve as shorthand for certain operators, wildcard characters, or placeholders.

If you have used any other Web search engines, you have probably been exposed to Boolean operators. Boolean operators help specify the results that are returned from a query. If you are already familiar with Boolean operators, take a moment to skim this section to help you understand Google's particular implementation of these operators, since many search engines handle them

in different ways. Improper use of these operators could drastically alter the results that are returned.

The most commonly used Boolean operator is AND. This operator is used to include multiple terms in a query. For example, a simple query like hacker could be expanded with a Boolean operator by querying for hacker AND cracker. The latter query would include not only pages that talk about hackers, but also sites that talk about hackers and the snacks they might eat. Some search engines require the use of this operator, but Google does not. The term AND is redundant to Google. By default, Google automatically searches for all the terms you include in your query. In fact, Google will warn you when you have included terms that are obviously redundant.

The plus symbol (+) forces the inclusion of the word that follows it. There should be no space following the plus symbol. For example, if you were to search for "and," "justice," "for," and "all" as separate, distinct words, Google would warn that several of the words are too common and are excluded from the search. To force Google to search for those common words, preface them with the plus sign. It's okay to go overboard with the plus sign. It has no ill effects if it is used excessively. To perform this search with the inclusion of all words, consider a query such as +and justice for +all. In addition, the words could be enclosed in double quotes. This generally will force Google to include all the common words in the phrase. This query presented as a phrase would be: "and justice for all."

Another common Boolean operator is NOT. Functionally the opposite of the AND operator, the NOT operator excludes a word from a search. The best way to use this operator is to preface a search word with the minus sign (-). Be sure to leave no space between the minus sign and the search term. Consider a simple query, such as hacker. This query is very generic and will return hits for all sorts of occupations like golfers, woodchoppers, serial killers, and those with chronic bronchitis. With this type of query, you are most likely not interested in each and every form of the word hacker but rather a more specific rendition of the term. To narrow the search, you could include more terms, which Google would automatically AND together, or you could start narrowing the search by using NOT to remove certain terms from your search. To remove some of the more unsavory characters from your search, consider using queries such as hacker -golf or hacker -phlegm. This would allow you to get closer to the dastardly wood choppers you're looking for. Or, you could try a Google Video search for lumberjack song. Talk about twisted.

A less common and sometimes more confusing Boolean operator is OR. The OR operator, represented by the pipe symbol (|) or simply the word OR in uppercase letters, instructs Google to locate either one term or another in a query. Although this seems fairly straightforward when considering a simple

query, such as “evil cybercriminal” or hacker, things can get terribly confusing when you string together a bunch of ANDs, ORs and NOTs. To help alleviate this confusion, don't think of the query as anything more than a sentence read from left to right. Forget all that order of operations stuff you learned in high school algebra. For our purposes, an AND is weighed equally with an OR, which is weighed as equally as an advanced operator. These factors may affect the rank or order in which the search results appear on the page, but have no bearing on how Google handles the search query.

Let's take a look at a very complex example, the exact mechanics of which we will discuss in [Chapter 2](#): `intext:password | passcode intext:username | userid | user filetype:csv`. This example uses advanced operators combined with the OR Boolean to create a query that reads like a sentence written as a polite request. The request reads, “Locate all pages that have either password or passcode in the text of the document. From those pages, show me only the pages that contain either the words username, userid, or user in the text of the document. From those pages, only show me documents that are CSV files.” Google doesn't get confused by the fact that technically those OR symbols break up the query into all sorts of possible interpretations. Google isn't bothered by the fact that from an algebraic standpoint, your query is syntactically wrong. For the purposes of learning how to create queries, all we need to remember is that Google reads our query from left to right.

Google's cut-and-dried approach to combining Boolean operators is still very confusing to the reader. Fortunately, Google is not offended (or affected by) parenthesis. The previous query can also be submitted as `intext:(password | passcode) intext:(username | userid | user) filetype:csv`. This query is infinitely more readable for us humans, and it produces exactly the same results as the more confusing query that lacked parentheses.

Search Reduction

To achieve the most relevant results, you'll often need to narrow your search by modifying the search query. Although Google tends to provide very relevant results for most basic searches, we will begin looking at fairly complex searches aimed at locating a very narrow subset of Web sites. The vast majority of this book focuses on search reduction techniques and suggestions, but it's important that you at least understand the basics of search reduction.

As a simple example, we'll take a look at GNU Zebra, free software that manages Transmission Control Protocol (TCP)/Internet Protocol (IP)-based routing protocols. GNU Zebra uses a file called `zebra.conf` to store configuration settings, including interface information and passwords. After downloading the latest version of Zebra from the Web, we learn that the included `zebra.conf` sample file looks like this:

```

! *-zebra.*
!
! zebra sample configuration file
!
!
! SId: zebra.conf.sample,v1.1# 1999/02/15 17:26:58 developer Exp 5
!
hostname Router
password zebra
enable password zebra
!
! Interface's description.
!
interface fa
! description test of desc.
!
interface s0/0
! multicast
!
! Static default route example.
!
ip route 0.0.0.0/0 205.181.89.241
!
!log file zebra.log

```

To attempt to locate these files with Google, we might try a simple search such as:

"! Interface's description." This is considered the base search. Base searches should be as unique as possible in order to get as close to our desired results as possible, remembering the old adage, "Garbage in, garbage out." Starting with a poor base search completely negates all the hard work you'll put into reduction. Our base search is unique not only because we have focused on



FIGURE 1.7

the words Interface's and description, but we have also included the exclamation mark, the spaces, and the period following the phrase as part of our search. This is the exact syntax that the configuration file itself uses, so this seems like a very good place to start. However, Google takes some liberties with this search query, making the results less than adequate, as shown in [Figure 1.7](#). Looking for zebra.conf files. So let's add this to our search to help narrow the results. This makes our next query: "! Interface's description." zebra.conf.

As [Figure 1.8](#) shows, the results are slightly different but not necessarily better.

For starters, the SeattleWireless hit we had in our first search is missing. This was a valid hit, but because the configuration file was not named zebra.conf, (it was named ZebraConfig) our "improved" search doesn't see it. This is a great lesson to learn about search reduction: don't reduce your way past valid results.

These sample files may clutter valid results, so we'll add to our existing query, reducing hits that contain this phrase. This makes our new query: "! Interface's description." - "zebra.conf.sample".

Now, it helps to step into the shoes of the software's users for just a moment. Software installations like this one often ship with a sample configuration file to help guide the process of setting up a custom configuration. Most users will simply edit this file, changing only the settings that need to be changed for their environments, saving the file not as a .sample file but as a .conf file.



FIGURE 1.8

In this situation, the user could have a live configuration file with the term `zebra.conf.sample` still in place. Reduction based on this term may remove valid configuration files created in this manner.

There's yet another reduction angle. Notice that our `zebra.conf.sample` file contained the term `hostname Router`. This is most likely one of the settings that a user will change; although we're making an assumption that his machine is not named `Router`. This is less a gamble than reducing based on `zebra.conf.sample`, however. Adding the reduction term "hostname Router" to our query brings our results number down and reduces our hits on potential sample files, all without sacrificing potential live hits.

Although it's certainly possible to keep reducing, often it's enough to make just a few minor reductions that can be validated by eye than to spend too much time coming up with the perfect search reduction. Our final (that's four qualifiers for just one word!) query becomes: "Interface's description." – "hostname Router". This is not the best query for locating these files, but it's good enough to give you an idea about how search reduction works. As we'll see in [Chapter 2](#), advanced operators will get us even closer to that perfect query.

Working With Google URLs

Advanced Google users begin testing advanced queries right from the Web interface's search field, refining queries until they are just right. Every Google query can be represented with a URL that points to the results page. Google's results pages are not static pages. They are dynamic and are created on the fly

when you click the Search button or activate a URL that links to a results page. Submitting a search through the Web interface takes you to a results page that can be represented by a single URL. For example, consider the query `ihackstuff`. Once you enter this query, you are whisked away to a URL similar to the following: www.google.com/search?q=ihackstuff. If you bookmark this URL and return to it later, or simply enter the URL into your browser's address bar, Google will reprocess your search for `ihackstuff` and display the results.

This URL then becomes not only an active connection to a list of results, but it also serves as a nice, compact sort of shorthand for a Google query. Any experienced Google searcher can take a look at this URL and realize the search subject. This URL can also be modified fairly easily. By changing the word `ihackstuff` to `iwritestuff`, the Google query is changed to find the term `iwritestuff`. This simple example illustrates the usefulness of the Google URL for advanced searching. A quick modification of the URL can make changes happen fast!

URL Syntax

To fully understand the power of the URL, we need to understand the syntax. The first part of the URL, www.google.com/search, is the location of Google's search script. I refer to this URL, as well as the question mark that follows it, as the base or starting URL. Browsing to this URL presents you with a nice, blank search page. The question mark after the word `search` indicates that parameters are about to be passed into the search script. Parameters are options that instruct the search script to actually do something. Parameters are separated by the ampersand (&) and consist of a variable followed by the equal sign (=), followed by the value that the variable should be set to. The basic syntax will look something like this: www.google.com/search?variable1=value&variable2=value. This URL contains very simple characters. More complex URLs will contain special characters, which must be represented with hex code equivalents. Let's take a second to talk about hex encoding.

Special Characters

Hex encoding is definitely geek stuff, but sooner or later you may need to include a special character in your search URL. When that time comes, it's best to just let your browser help you out. Most modern browsers will adjust a typed URL, replacing special characters and spaces with hex-encoded equivalents. If your browser supports this behavior, your job of URL construction is that much easier. Try this simple test: Type the following URL in your browser's address bar, making sure to use spaces between `i`, `hack`, and `stuff`: [www.google.com/search?q="i hack stuff"](http://www.google.com/search?q=). If your browser supports this autocorrecting feature, after you press Enter in the address bar, the URL should be corrected to [www.google.com/search?q="i%20hack%20stuff"](http://www.google.com/search?q=), or something similar. Notice that the spaces were changed to `%20`. The percent sign indicates that the next two

- [read *Princes of the Yen: Japan's Central Bankers and the Transformation of the Economy* online](#)
- [download online *A Red Death \(Easy Rawlins Mystery, Book 2\)*](#)
- [download online *Vicious Circle \(Hector Cross, Book 2\)* pdf, azw \(kindle\)](#)
- [*Malinowski, Rivers, Benedict and Others: Essays on Culture and Personality \(History of Anthropology, Volume 4\)* book](#)
- [click *Gothic Sports, Volume 1*](#)

- <http://www.netc-bd.com/ebooks/Princes-of-the-Yen--Japan-s-Central-Bankers-and-the-Transformation-of-the-Economy.pdf>
- <http://flog.co.id/library/Ghost-Fleet--A-Novel-of-the-Next-World-War.pdf>
- <http://monkeybubblemedia.com/lib/Vicious-Circle--Hector-Cross--Book-2-.pdf>
- [http://thermco.pl/library/Malinowski--Rivers--Benedict-and-Others--Essays-on-Culture-and-Personality--History-of-Anthropology--Volume-4-](http://thermco.pl/library/Malinowski--Rivers--Benedict-and-Others--Essays-on-Culture-and-Personality--History-of-Anthropology--Volume-4-.)
- <http://flog.co.id/library/MongoDB-Cookbook--2nd-Edition-.pdf>