# *Google Hacking Guide*


***For educational purposes only.***
***Not for use in matters or on sites you are not implicitly and legally approved to research***

*By Ken Foster – KMBL Security*

## What is Google hacking?

The purpose of Google Hacking is to leverage the vast amounts of data that are stored and indexed in search engines to produce unique results that quickly identify sensitive information, vulnerable systems, and network tactics and methods used by their hosts. These methods are beneficial to security professionals, ethical hackers, and unfortunately black hats.

## Who is Google Hacking?

Google hacking has become very popular among security testers, ethical hackers, and unfortunately script kiddies and hackers who all too well understand its value.

## Is it Illegal?

This is debatable and dependent upon the country or state you live in and the location of the target. As a standard practice you should never use these techniques against a target that you do not have written permission from the owner.

## What do will I Need?

A web browser, an internet connection, and a word processing program to save useful results for later review and analysis.

## Will the Target Know What I am Up to?

Just as a Hacker can perform these queries against the target, so can Google capture the query results and report them through mechanisms such as Google Alerts and Google Analytics. Additionally access to resources that are derived from these searches may be logged by the hosts' security systems, which may indicate your reconnaissance.

## Let's try something Simple…

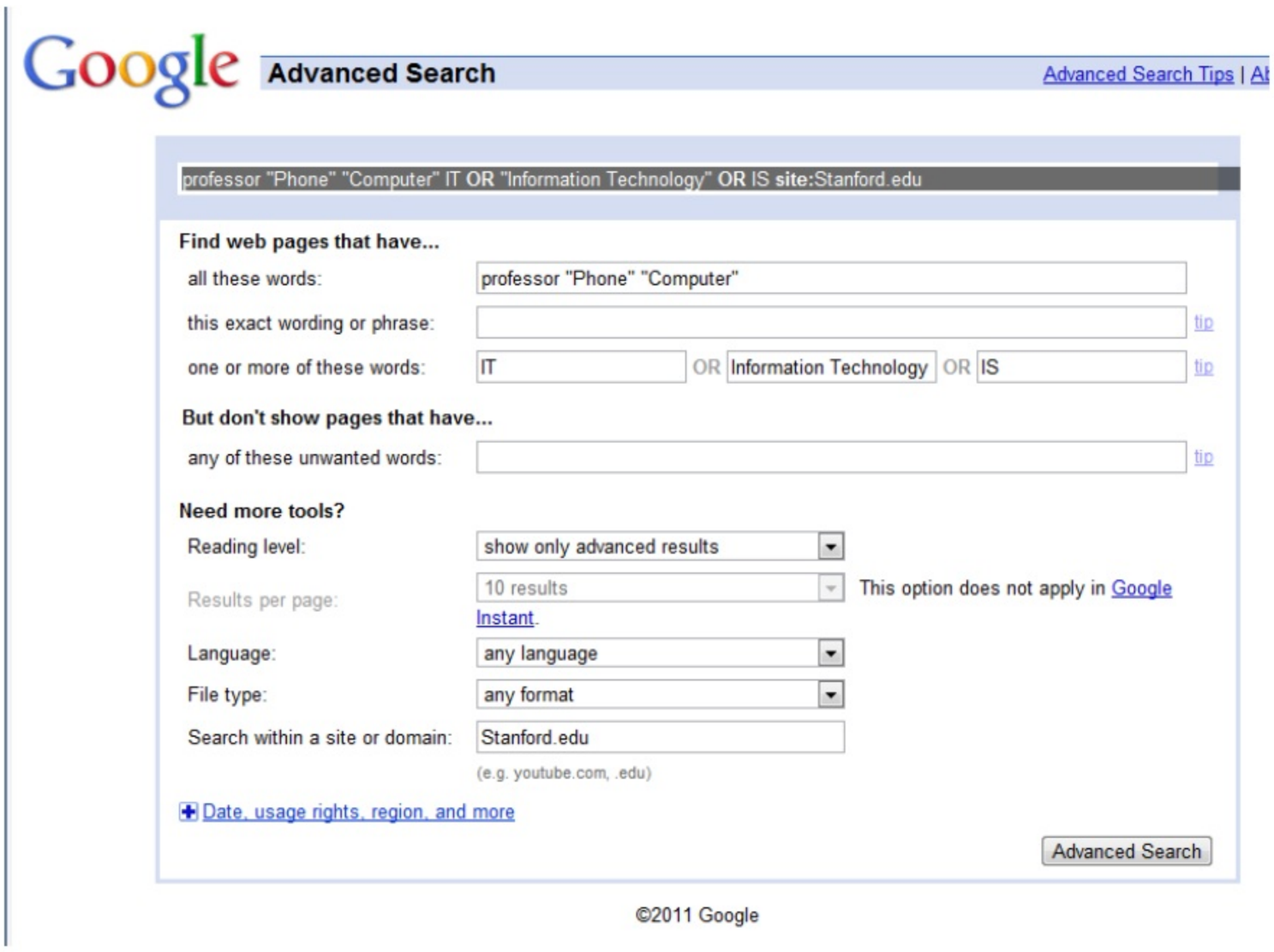
Let run a search for all the phone number listed on the Stanford University Public Web Site. We would do this using a specially tailored search phrase. In google query languard, just like SQL database language, there are key words. One such key word is the term *intext:,* which tells the search engine to review the collective text of the indexed page and search for the search term that follows. For example the results of **intext: 650723** provides us access to a broad listing of Stanford staff and instructor phone numbers, most with emails as well. How did we decide to use the search term 650723. By going to the home page of Stanford University, we were able to locate the Area code (650) and Phone exchange prefix (723) used by the University. Since most large organization utilize the entire prefix (the three numbers following the area code), this simplifies our efforts.

## What is you don't have the time or want to write Google Query Language Calls?

If you are in a hurry, no worries use the Google Advanced Search page to get your information. Up until recently this was easily accessible. Now you must know the pages address in order to access the page. Your best course of action is to bookmark the below address:
"http://www.google.com/advanced_search?hl=en". The Advanced Search page will allow you to mix a

combination of key words, phrases, and disqualifier words into a very complex query. In addition you can force the domain or site in which to narrow the results and even a specific type of file you may be seeking. Using our example from earlier, let's say we wanted to find the phone numbers of Stanford Professors that worked in IT related fields. Using the Advanced Search page we could build a query like:

**professor "Phone" "Computer" IT OR "Information Technology" OR IS site:Stanford.edu**




Which would return focused results based on the "intext" criteria we had specified above.

Mendel Rosenblum - Stanford School of Engineering - Personnel Profile
soe.stanford.edu/research/layout.php?sunetid=mendel - Cached
Affiliation(s): **Faculty** Director, Stanford **Computer** Forum Affiliates Program; **...**
**Phone**: 650.723.0474. Fax: 650.725.6949. E-mail: mendel@cs.stanford.edu **...**

Leading-Matters Bay Area Panels/Breakouts
leadingmatters.stanford.edu/bayarea/panels.htm - Cached
Sebastian Thrun, **professor** of **computer** science and electrical engineering Read More **....** There are now some 4 billion mobile **phone** subscriptions in the world. **...** on Human-**Computer** Interaction, **Information Technology** and People. Close **...**

Hector Garcia-Molina
infolab.stanford.edu/people/hector.html - Cached
**Professor**, Departments of **Computer** Science and Electrical Engineering **...**
Office: Gates Hall 4A, Room 434 **Phone**: (650) 723-0685 Fax: (650) 725-2588 **...**
From 1997 to 2001 he was a member the President's **Information Technology** Advisory **...**

## Google Search Operators

A search operator is used to create complex search criteria. For instance, you want to search for multiple items that are different in a single search request. You may want to specify certain items you do not want return if that keyword is located in the results. This is where search operators are used.

**And (+)**: Specified by using either the word *AND* or using the "+" symbol without quotes. Using this in the search term specified you want to find results that use multiple criteria. For example: dog + beagle + champion would return search results that include all three items in the results.

**Or (|):** Specified by using either the word *OR* or the "|", also known as the pipe symbol. You use this search term to specify conditional criteria. For example: 67 Nova | Malibu hotrods would return search results containing the result of 67 Nova or 67 Malibu hotrods.

## Google Hacking Search Terms

*Note: Do not place a space between the Keyword (e.g. cache) and the page url.*

**cache:**

If you include other words in the query, Google will highlight those words within the cached document. For instance, [cache:www.google.com web] will show the cached content with the word "web" highlighted. This functionality is also accessible by clicking on the "Cached" link on Google's main results page. The query [cache:] will show the version of the web page that Google has in its cache. For example the query *cache:www.google.com* will show Google's cache of the Google homepage.

**link:**

The query [link:] will list webpages that have links to the specified webpage. For example the query *link:www.google.com* will list webpages that have links pointing to the Google homepage.

**related:**

The query [related:] will list web pages that are "similar" to a specified web page. For example the query *related:www.google.com* will list web pages that are similar to the Google homepage.

**Info:**

The query [info:] will present some information that Google has about that web page. For example the query *info:www.google.com* will show information about the Google homepage.

**site:**

If you include [site:] in your query, Google will restrict the results to those websites in the given domain. For example in the query *help site:www.google.com* you will find pages about help within www.google.com.

**allintitle:**

If you start a query with [allintitle:], Google will restrict the results to those with all of the query words in the title. For example in the query *allintitle: google search* you will return only documents that have both “google” and “search” in the title.

**intitle:**

If you include [intitle:] in your query, Google will restrict the results to documents containing that word in the title. For example in the query *intitle:google search* you will return documents that mention the word “google” in their title, and mention the word “search” anywhere in the document (title or no).

*Notes:*

1. *there can be no space between the “intitle:” and the following word.*

2. *Note: Putting [intitle:] in front of every word in your query is equivalent to putting [allintitle:] at the front of your query: [intitle:google intitle:search] is the same as [allintitle: google search].*

**allinurl:**

If you start a query with [allinurl:], Google will restrict the results to those with all of the query words in the url. For example in the query *allinurl: google search* you will return only documents that have both “google” and “search” in the url.

*Note that [allinurl:] works on words, not url components. In particular, it ignores punctuation. Thus, [allinurl: foo/bar] will restrict the results to page with the words “foo” and “bar” in the url, but won’t require that they be separated by a slash within that url, that they be adjacent, or that they be in that particular word order. There is currently no way to enforce these constraints.*

**inurl:**

If you include [inurl:] in your query, Google will restrict the results to documents containing that word in the url. For example in the query inurl:google search you will return documents that mention the word “google” in their url, and mention the word “search” anywhere in the document (url or no). Note there can be no space between the “inurl:” and the following word.

Putting “inurl:” in front of every word in your query is equivalent to putting “allinurl:” at the front of your query: [inurl:google inurl:search] is the same as [allinurl: google search].

## Ready to try something more?

The following listing provides a collection of the more popular search terms that will aid in your security evaluation of your host targets.  Remember; always ensure you are performing these queries on your own hosts or with the written permission of the hosts' owner to avoid any potential legal issues.

## Common Google Hacking Terms and Expected Results

"#mysql dump" filetype:sql
"access denied for user" "using password"
"Certificate Practice Statement" filetype:PDF | DOC
"Generated by phpSystem"
"HTTP_FROM=googlebot" googlebot.com "Server_Software="
"Host Vulnerability Summary Report"
"Index of" / "chat/logs"
"Installed Objects Scanner" inurl:default.asp
"Mecury Version" "Infastructure Group"
"Microsoft (R) Windows * (TM) Version * DrWtsn32 C
"Microsoft (R) Windows * (TM) Version * DrWtsn32 Copyright (C)" ext:log
"Most Submitted Forms and Scripts" "this section"
“my webcamXP server!”
"Network Vulnerability Assessment Report"
"ORA-00921: unexpected end of SQL command"
"Request Details" "Control Tree" "Server Variables"
"Running in Child mode"
"Thank you for your order" +receipt
"These statistics were produced by getstats"
"This is a Shareaza Node"
"This report was generated by WebLog"
"This summary was generated by wwwstat"
"allow_call_time_pass_reference" "PATH_INFO"
"not for distribution" confidential
"phone * * *" "address *" "e-mail" intitle:"curriculum vitae"
"phpMyAdmin" "running on" inurl:"main.php"
"robots.txt" "Disallow:" filetype:txt
"sets mode: +p"
"sets mode: +s"
( filetype:mail | filetype:eml | filetype:mbox | filetype:mbx ) intext:password|subject
(inurl:"robot.txt" | inurl:"robots.txt" ) intext:disallow filetype:txt
+":8080" +":3128" +":80" filetype:txt
+"HSTSNR" -"netop.com"
-site:php.net -"The PHP Group" inurl:source inurl:url ext:pHp
94FBR "ADOBE PHOTOSHOP"
AIM buddy lists
Financial spreadsheets: finance.xls
Financial spreadsheets: finances.xls
Ganglia Cluster Reports

ICQ chat logs, please...
Lotus Domino address books
Microsoft Money Data Files
MySQL tabledata dumps
OWA Public Folders (direct view)
Peoples MSN contact lists
Quicken data files
SQL data dumps
Squid cache server reports
Unreal IRCd
Welcome to ntop!
allinurl:/examples/jsp/snp/snoop.jsp
allinurl:servlet/SnoopServlet
buddylist.blt
cgiirc.conf
e-mail address filetype:csv csv
exported email addresses
ext:asp inurl:pathto.asp
ext:cgi inurl:editcgi.cgi inurl:file=
ext:conf NoCatAuth -cvs
ext:conf inurl:rsyncd.conf -cvs -man
ext:dat bpk.dat
ext:gho gho
ext:ini intext:env.ini
ext:ldif ldif
ext:log "Software: Microsoft Internet Information
ext:log "Software: Microsoft Internet Information Services *.*"
ext:mdb inurl:*.mdb inurl:fpdb shop.mdb
ext:nsf nsf -gov
ext:pqi pqi -database
ext:reg "username=*" putty
ext:txt "Final encryption key"
ext:txt inurl:dxdiag
ext:vmdk vmdk
ext:vmx vmx
filetype:asp DBQ=" * Server.MapPath("*.mdb")
filetype:bkf bkf
filetype:blt "buddylist"
filetype:blt blt +intext:screenname
filetype:cfg auto_inst.cfg
filetype:cnf inurl:_vti_pvt access.cnf
filetype:conf inurl:firewall -intitle:cvs
filetype:config web.config -CVS
filetype:ctt Contact
filetype:ctt ctt messenger
filetype:eml eml +intext:"Subject" +intext:"From" +intext:"To"
filetype:fp3 fp3
filetype:fp5 fp5 -site:gov -"cvs log"

filetype:fp7 fp7
filetype:inf inurl:capolicy.inf
filetype:lic lic intext:key
filetype:log access.log -CVS
filetype:mbx mbx intext:Subject
filetype:myd myd -CVS
filetype:ns1 ns1
filetype:ora ora
filetype:pst inurl:"outlook.pst"
filetype:pst pst -from -to -date
filetype:rdp rdp
filetype:wab wab
filetype:xls -site:gov inurl:contact
filetype:xls inurl:"email.xls"
haccess.ctl
filetype:log cron.log
intext:"Session Start * * * *:*:* *" filetype:log
intext:"Tobias Oetiker" "traffic analysis"
intext:(password | passcode) intext:(username | userid | user) filetype:csv
intext:SQLiteManager inurl:main.php
intitle:"Apache::Status" (inurl:server-status | inurl:status.html | inurl:apache.html)
intitle:"AppServ Open Project" -site:www.appservnetwork.com
intitle:"Big Sister" +"OK Attention Trouble"
intitle:"FTP root at"
intitle:"Index Of" cookies.txt size
intitle:"index of" mysql.conf OR mysql_config
intitle:"index.of *" admin news.asp configview.asp
intitle:"wbem" compaq login
intitle:"web server status" SSH Telnet
intitle:admin intitle:login
intitle:index.of "Apache" "server at"
intitle:index.of dead.letter
intitle:index.of inbox dbx
intitle:intranet inurl:intranet +intext:"phone"
inurl:"newsletter/admin/"
inurl:"shopadmin.asp" "Shop Administrators only"
inurl:*db filetype:mdb
inurl:admin filetype:xls
inurl:backup filetype:mdb
inurl:changepassword.asp
inurl:forward filetype:forward -cvs
inurl:main.php Welcome to phpMyAdmin
inurl:main.php phpMyAdmin
inurl:netscape.hst
inurl:odbc.ini ext:ini -cvs
inurl:preferences.ini "[emule]"
inurl:ssl.conf filetype:conf
ipsec.conf

ipsec.secrets
private key files (.csr)
private key files (.key)