Playing The Search Engines or Hacking The Box

Frédéric Philipp Thiele
contact@fp-thiele.de

Prof. Hendrik Speck
contact@hendrikspeck.com
Google Hacking
Google Hacking

"Search engines make it easier for everyone to gain information, hackers included."

Danny Sullivan
editor of searchenginewatch.com
Google Hacking
Concept

Database with more than 9 Billion documents

- Personal data
- Passwords
- User accounts
- Webcams
- Databases
- Password Secured Areas
- Customer Data
- Security Bugs
- …
Google Hacking

Database with over 9 Billion Documents + powerful Search Functionality = ?
Google Hacking

Definition

Google Hacking is a special search technique used to find hidden information, break into sites, and access supposedly secure information. Using search engines such as Google, "search engine hackers" can easily find exploitable targets and sensitive data.
Search Functionality
Google Advanced Search
Overview

You can do a lot more with Google search than just typing in search terms. With Advanced Search, you can search for pages:

- that contain ALL the search terms you type in
- that contain the exact phrase you type in
- that contain at least one of the words you type in
- that do NOT contain any of the words you type in
- written in a certain language
- created in a certain file format
- that have been updated within a certain period of time
- that contain numbers within a certain range
- within a certain domain, or website
- that don't contain "adult" material

Source: http://www.google.com/help/refinesearch.html
Advanced Search Functionality

link:wikipedia.org
Advanced Search Functionality

The Beatles - Octopus's Garden
"I'd ask my friends to come and see"

---

**Relevant Search Result:**
- "I'd ask my friends to come and see" - I've been trying to visit you for a long time but I don't have the strength to...
- Abbey Road Lyrics
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Abbey Road - Lyrics - Octopus's Garden (tracks 2 and 8)
- Octopus
- Abbey Road - Encyclopedia
- Fabrizio Costa
- Genevieve's Beatles
- Steve's Beatles Page - CDs - Abbey Road
- Abbey Road - Lyrics - Octopus's Garden (tracks 2 and 8)

---

**Irrelevant Search Result:**
- I don't think I listened to anything but rock and roll.
- Tech Support: How can I see my friend's LinkedIn Connections?... but even without that analogy, the key question you need to ask is what you're...
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road
- Steve's Beatles Page - CDs - Abbey Road

---

*The Beatles - Octopus's Garden*
Advanced Search Functionality
lecture search engines filetype:ppt
User Account Hacking
Getting your Virtual Identity
Many administrators fail to protect the admin panel (/admin/admin.php) with the .htaccess logon procedure. In such cases, customer information like email and passwords are in clear view of the attacker.
### Password Security

Don't use the same passwords for different services.

<table>
<thead>
<tr>
<th>Records 101 to 107 of 107</th>
</tr>
</thead>
<tbody>
<tr>
<td>first name</td>
</tr>
<tr>
<td>RON</td>
</tr>
<tr>
<td>Antoine MOUZING</td>
</tr>
<tr>
<td>Maryam Borzilli</td>
</tr>
<tr>
<td>percent</td>
</tr>
<tr>
<td>logan</td>
</tr>
<tr>
<td>ramesh rasal</td>
</tr>
<tr>
<td>Vinicius Filhaha</td>
</tr>
<tr>
<td>Vinicius Filhaha</td>
</tr>
</tbody>
</table>
## Password Security

### Search String: ronboy51

<table>
<thead>
<tr>
<th>Records 101 to 107 of 107</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>first name</td>
<td>last name</td>
<td>username</td>
<td>password</td>
<td>email</td>
</tr>
<tr>
<td>RON</td>
<td>STORY</td>
<td>ronboy51</td>
<td>timothy</td>
<td><a href="mailto:ron.story@officeemail.net">ron.story@officeemail.net</a></td>
</tr>
<tr>
<td>Antoine</td>
<td>MOUZINGAZoumra</td>
<td>mouz1953zoumra@yahoofr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MaryannBorzillary</td>
<td>genmb1953maryjo</td>
<td><a href="mailto:blessedwithcash@yahoo.com">blessedwithcash@yahoo.com</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>percent</td>
<td>aweek</td>
<td>percent</td>
<td>theone</td>
<td><a href="mailto:percentaweek@yahoo.com">percentaweek@yahoo.com</a></td>
</tr>
<tr>
<td>ltk</td>
<td></td>
<td></td>
<td></td>
<td><a href="mailto:03191991bigbadness365@yahoo.com">03191991bigbadness365@yahoo.com</a></td>
</tr>
<tr>
<td>ramesh</td>
<td>rasal</td>
<td><a href="mailto:rameshrasa.mh127896rameshrasal@yahoo.com">rameshrasa.mh127896rameshrasal@yahoo.com</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vinicius Filhaha</td>
<td>vinicius</td>
<td>vinicius</td>
<td><a href="mailto:vinicius@vinicius777.com">vinicius@vinicius777.com</a></td>
<td></td>
</tr>
<tr>
<td>Vinicius Filhaha</td>
<td>impacta</td>
<td>impacta</td>
<td><a href="mailto:vinicius777@gmail.com">vinicius777@gmail.com</a></td>
<td></td>
</tr>
</tbody>
</table>
### Password Security

**Search String: ronboy51**

<table>
<thead>
<tr>
<th>First Name</th>
<th>Last Name</th>
<th>Username</th>
<th>Password Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antoine</td>
<td>MOUZINGA</td>
<td>zoumra</td>
<td><a href="mailto:mouz1953zoumra@yahoo.fr">mouz1953zoumra@yahoo.fr</a></td>
</tr>
<tr>
<td>Maryann</td>
<td>Borzillar</td>
<td>maryjo</td>
<td><a href="mailto:blessedwithcash@yahoo.com">blessedwithcash@yahoo.com</a></td>
</tr>
<tr>
<td>percent</td>
<td>aweek</td>
<td>theone</td>
<td><a href="mailto:percentaweek@yahoo.com">percentaweek@yahoo.com</a></td>
</tr>
<tr>
<td>logan</td>
<td>knowles</td>
<td>ltc</td>
<td><a href="mailto:03191991bigbadness365@yahoo.com">03191991bigbadness365@yahoo.com</a></td>
</tr>
<tr>
<td>ramesh</td>
<td>rasal</td>
<td>rameshrasa <a href="mailto:mh127896rameshrasal@yahoo.com">mh127896rameshrasal@yahoo.com</a></td>
<td></td>
</tr>
<tr>
<td>Vinicius</td>
<td>Filhaha</td>
<td>vinicius</td>
<td><a href="mailto:vinicius@vinicius777.com">vinicius@vinicius777.com</a></td>
</tr>
<tr>
<td>Vinicius</td>
<td>Filhaha</td>
<td>impacta</td>
<td><a href="mailto:vinicius777@gmail.com">vinicius777@gmail.com</a></td>
</tr>
</tbody>
</table>

A search for ronboy51 offers about 100 potential targets.
Password Security

Don't use the same passwords for different services.
Password Security

Don't use the same passwords for different services.
WS FTP – FTP Passwords

Search String: intitle:index.of ws_ftp.ini

WS_FTP.ini is a configuration file for a popular win32 FTP client that stores usernames and weakly encoded passwords.
WS FTP – FTP Passwords
Search String: intitle:index.of ws_ftp.ini

WS_FTP.ini is a configuration file for a popular win32 FTP client that stores usernames and weakly encoded passwords.
WS FTP – FTP Passwords
Search String: intitle:index.of ws_ftp.ini
## Other Examples

<table>
<thead>
<tr>
<th>Query</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>profiles!Host=.*</td>
<td>VPN Profiles</td>
</tr>
<tr>
<td>intext:enc_UserPassword=* ext:pcf</td>
<td></td>
</tr>
<tr>
<td>ext:ini eudora.ini</td>
<td>Eudora Mailclient Configuration Files</td>
</tr>
<tr>
<td>&quot;parent directory&quot; +proftpdpasswd</td>
<td>ProFTPd. Password</td>
</tr>
<tr>
<td>ext:ini Version=4.0.0.4 password</td>
<td>servU FTP Server Configuration File</td>
</tr>
<tr>
<td>intitle:index.of trillian.ini</td>
<td>Trillian Messenger Password File</td>
</tr>
<tr>
<td>filetype:inf sysprep</td>
<td>MS Windows Configuration File Administrator's passwords, IP addresses and product IDs</td>
</tr>
</tbody>
</table>

Source: [http://johnny.ihackstuff.com](http://johnny.ihackstuff.com)
Database Hacking
phpMyAdmin

Search String: "phpMyAdmin" "running on" inurl:"main.php"

Onlineshop Database with customer details and credit card informations.

phpMyAdmin is a tool written in PHP intended to handle the administration of MySQL over the WWW.
phpMyAdmin

Search String: "phpMyAdmin" "running on" inurl:"main.php"
Digital Device Hacking and Default Passwords
Linksys Webcam
Search String: vr intitle:"Linksys wireless-G Internet Video Camera"

There are a lot of digital devices with HTTP configuration interface. Most users forget to change the default passwords of their digital devices.
Search Engine Market Analysis
Using Google as a Research Tool
Search Engine Market
comScore. March 2006 (USA)

Search Engine Market
WebHits. March 2006 (Germany)

Statistics - Search Engine Market

Introduction

Different technologies and methodologies have resulted in different measurement standards and indicators.

Statistical problems, including active and latent errors, different target groups, samples, sample sizes and objectives lead to different results.

Different results support misuse, bias, and intentional errors.
Statistics - Search Engine Market

Problem Sources

1. Miuse, Bias, and Intentional Errors
2. Environment and Conditions
3. Methods, Procedures, and Processing
4. Indicators and Understanding Data
5. Data, Sample, and Sample Size
Statistics - Search Engine Market

Problem Sources

Search Engines
*(Bias and Business Secrets)*

Examples: Google, MSN, Yahoo, Ask Jeeves

User Surveys/Proxies
*(Procedure, Reliability, Market, Sample, Sample Size)*

Example. Nielsen.

Logfile Analyzer
*(Real Time, Source, Completeness, Focus, Sample, Sample Size)*

Examples. Analog, AWStats
Statistics - Search Engine Market
Problem Sources

Web Tracker - Tracking Software/Counter
installed on every Page
(Real Time, Source, Completeness, Market, Focus,
Sample, Sample Size)

Examples. Google Analytics, Webposition

Web Analytics Software - Logfile Analyzer
(Based on Logfiles, Source, Focus, Sample Size, Focus, Resources)

Examples. Sawmill
Searching for Logfiles
intitle: "Usage Statistics for" "Generated by Webalizer"
Searching for Logfiles

intitle:"statistics of" "advanced web statistics"
Searching for Logfiles

intitle:"statistics of" "advanced web statistics"

Links from an Internet Search Engine

<table>
<thead>
<tr>
<th>Search Engine</th>
<th>Hits</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>1315</td>
<td>49.2%</td>
</tr>
<tr>
<td>Virgilio</td>
<td>728</td>
<td>27.2%</td>
</tr>
<tr>
<td>Yahoo</td>
<td>506</td>
<td>18.9%</td>
</tr>
<tr>
<td>AltaVista</td>
<td>60</td>
<td>2.2%</td>
</tr>
<tr>
<td>Other search engines</td>
<td>25</td>
<td>0.9%</td>
</tr>
<tr>
<td>WebCrawler</td>
<td>6</td>
<td>0.2%</td>
</tr>
<tr>
<td>Ilse</td>
<td>6</td>
<td>0.2%</td>
</tr>
<tr>
<td>MSN</td>
<td>4</td>
<td>0.1%</td>
</tr>
<tr>
<td>Netscape</td>
<td>4</td>
<td>0.1%</td>
</tr>
<tr>
<td>AllTheWeb</td>
<td>3</td>
<td>0.1%</td>
</tr>
<tr>
<td>MetaCrawler (Metsmoteur)</td>
<td>3</td>
<td>0.1%</td>
</tr>
<tr>
<td>Lycos</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>AOL</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>Setnam.cz</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>Overture</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>Excite</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>Web.de</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>Search.com</td>
<td>1</td>
<td>0%</td>
</tr>
</tbody>
</table>
Regular Expression

Definition

A regular expression (abbreviated as regexp or regex, with plural forms regexps, regexes, or regexen) is a string that describes or matches a set of strings, according to certain syntax rules. Regular expressions are used by many text editors and utilities to search and manipulate bodies of text based on certain patterns.

Searching for Logfiles
Different Search Queries

intitle:"statistics of" "advanced web statistics" 80,000
allinurl:awstats.pl?month= 294,000
allinurl:?output=refererse 11,500
Searching for Logfiles
Different Search Queries

Portuguese Language

"Estatísticas de acesso ao servidor WEB"

Brazil

site:.br "Estatísticas de acesso ao servidor WEB"
Searching for Logfiles

Different Search Queries

Chinese

Taiwanese

British

site:.cn  ....

site:.tw  ....

site:.co.uk  ....
Searching for Logfiles

Functional Overview

Collecting
1. search engine query
2. parsing search engine result pages (regex)
3. identifying domain and cc
4. crawling search engine results

Analysing
5. parsing data sets (regex)
6. data computing, consolidation and analysis (ip2dns)

Selecting
7. database
8. filtering and dataset creation
Questions?

Frédéric Philipp Thiele
contact@fp-thiele.de

Prof. Hendrik Speck
contact@hendrikspeck.com