

Lord of the Bing!

Taking back search engine hacking from Google and Microsoft

03 MAY 2010



Presented by:
Vinnie Liu, Managing Partner
Stach & Liu, LLC
www.stachliu.com

Agenda

OVERVIEW

- Introduction
 - Quick History
- Advanced Attacks
 - Google/Bing Hacking
 - Other OSINT Attack Techniques
- Advanced Defenses
- Future Directions



Goals

DESIRED OUTCOME

- *To understand* Google Hacking
 - Attacks and defenses
 - Advanced tools and techniques
- *To survey* the best publicly available sources
- To blow your mind!



Introduction/ Background

GETTING UP TO SPEED



Open Source Intelligence

SEARCHING PUBLIC SOURCES

OSINT – is a form of intelligence collection management that involves finding, selecting, and acquiring information from *publicly available* sources and analyzing it to *produce actionable intelligence*.



Quick History

GOOGLE HACKING RECAP

2004 GHDB
Established

MSNPawn
v1 Released

May 2004
SiteDigger
v1 Released

Feb 2005
Google
Hacking
Book
Published

Quick History

GOOGLE HACKING RECAP

Dec 2006
Google stops
issuing SOAP
API keys

Nov 2007
Google
Hacking v2
Published

MSN
disables
inurl:, link:,
and
linkdomain:

Mar 2008
cDc Goolag
Released

Quick History

GOOGLE HACKING RECAP

Jun 2009
Bing goes
live

Nov 2009
Binging
Released

Sep 2009
Google
shuts down
SOAP Search
API

Dec 2009
SiteDigger
v3 Released

Quick History

GOOGLE HACKING RECAP

2010 Goolag.org
Disappears



April 2010
GoogleDiggity and
BingDiggity Released

Threat Areas

WHAT YOU SHOULD KNOW



Google/Bing Hacking



SEARCH ENGINE ATTACKS

- Our favorites are **Google** and **Bing**
- **Crawl** and **Index**
- **Cache** and **RSS** are forever
- **Query** modifiers
 - site:target.com
 - related:target.com
 - filetype:xls
 - ip:69.63.184.142

Attack Targets

GOOGLE HACKING DATABASE

- Advisories and Vulnerabilities (215)
- Error Messages (58)
- Files containing juicy info (230)
- Files containing passwords (135)
- Files containing usernames (15)
- Footholds (21)
- Pages containing login portals (232)
- Pages containing network or vulnerability data (59)
- Sensitive Directories (61)
- Sensitive Online Shopping Info (9)
- Various Online Devices (201)
- Vulnerable Files (57)
- Vulnerable Servers (48)
- Web Server Detection (72)

Attack Targets

GOOGLE HACKING DATABASE

Examples

Error Messages

- filetype:asp + "[ODBC SQL"
- "Warning: mysql_query()" "invalid query"

Files containing passwords

- inurl:passlist.txt

Google Hacking Toolkit

STATE OF THE ART

- SiteDigger v3.0
 - Uses Google AJAX API
 - Not blocked by Google
 - But restricted to 64 results/query
- Binging
 - Uses Microsoft Bing search engine
 - Limited domain/ip profiling utils
- Gooscan, Goolag
 - Work still, but get blocked by Google bot detection
 - Download sites not longer around



Google Hacking Toolkit

FOUNDSTONE SITEDIGGER



The screenshot shows the SiteDigger application window. The title bar reads "SiteDigger". The menu bar includes "File", "Edit", "Tools", and "Help". On the left, a tree view shows various search categories, with "FSDB(175)" and "GHDB(1467)" expanded. The main area has a "Site/Domain:" field containing "sample.com" and an "[Optional]" label. Below this are "Scan" and "Clear" buttons. The "Queries Scanned:" section lists 12 queries, with the first one selected. The "Selected Entry Info:" section displays the content of the selected query. At the bottom, a table shows the results of the scan.

Site/Domain: [Optional]

Queries Scanned:

- 1 "Index of /backup" F142
- 2 intitle:"Index of" ".htpasswd" htpasswd.bak F31
- 3 intitle:"Index of" index.html.bak F1
- 4 intitle:"Index of" index.html.bak F176
- 5 intitle:"Index of" index.html~ F178
- 6 intitle:"Index of" index.jsp.bak F3
- 7 intitle:"Index of" index.php.bak F2
- 8 intitle:"Index of" index.php.bak F177
- 9 intitle:"Index of" index.php~ F179
- 10 intitle:index.of .bash_history F19
- 11 intitle:index.of .sh_history F20
- 12 inurl:backup intitle:index.of inurl:admin F141

Selected Entry Info:

The robots.txt file serves as a set of instructions for web crawlers. The "disallow" tag tells a web crawler where not to look.

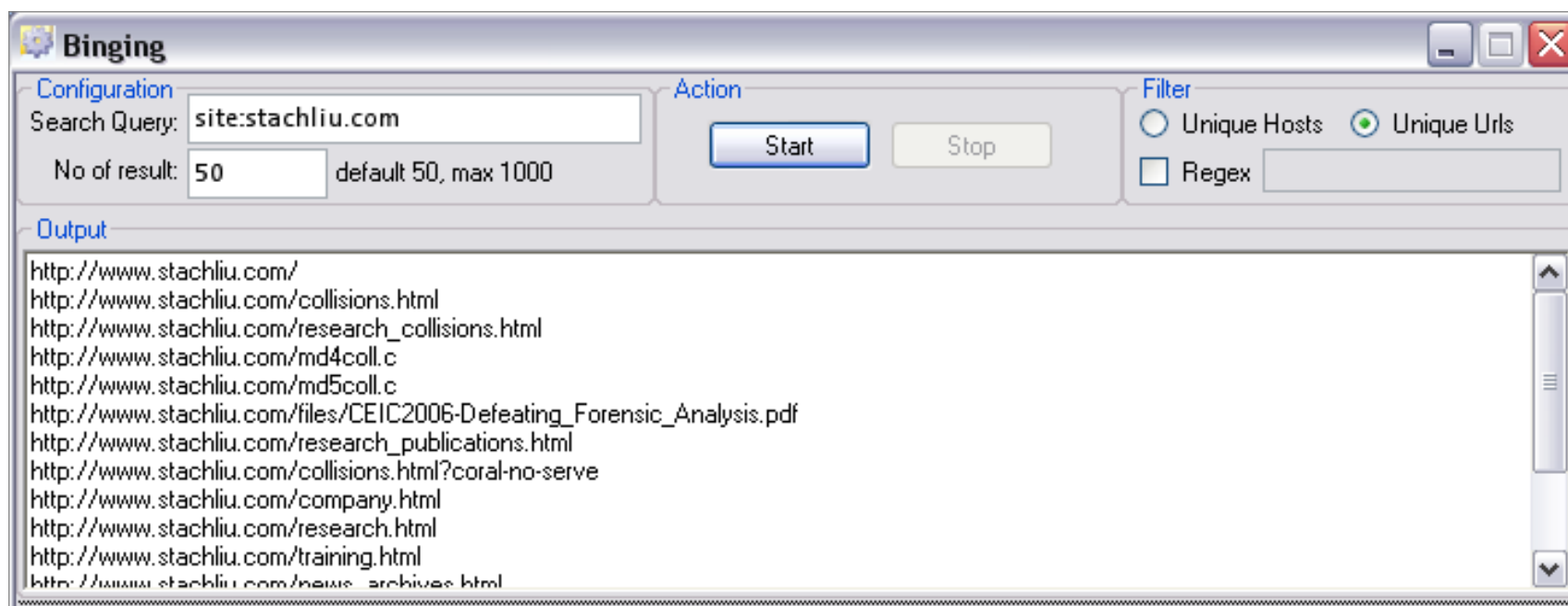
Cancelled

Results: [Double click a link to open in default browser]

URL	Query	Category
http://www.sample.com/robots.txt	"robots.txt" + "Disallow:" filetype:txt	Technology ...

Google Hacking Toolkit

BINGING



NEW GOOGLE HACKING TOOLS

DEMO

New Toolkit

STACH & LIU TOOLS



GoogleDignity

- Uses Google AJAX API
 - Not blocked by Google bot detection
- Can Leverage **Google** custom search

BingDignity

- Company/Webapp Profiling
 - Enumerate: URLs, IP-to-virtual hosts, etc.
- Bing Hacking Database (BHDB)
 - Regexs in Bing format

Defenses



GOOGLE/BING HACKING DEFENSES

- “Google Hack yourself” organization
 - Employ tools and techniques used by hackers
- Policy and Legal Restrictions
- Regularly update your robots.txt
- Data Loss Prevention/Extrusion Prevention Systems
- Social Sentry
 - Service to monitor employee Facebook and Twitter for \$2-\$8 per employee (MySpace, YouTube, and LinkedIn support by summer)

Google Apps Explosion

SO MANY APPLICATIONS TO ABUSE

Google alerts

Google reader

Google
PhoneBook

Google custom search

Google
trends

Google
code search labs

Google code

Google health

Google
public data explorer
labs

Google Insights for Search
beta

Google PhoneBook



SPEAR PHISHING

Google PhoneBook search interface showing results for "phonebook: schmidt, eric". The search bar contains "phonebook: schmidt, eric" and buttons for "Search PhoneBook" and "Search the Web". A "Preferences" link is also visible.

Residential Phonebook Results 121 - 150 of about 209 for phonebook: schmidt, e

Eric Schmidt	(952) 403-9689	1761 Countryside Dr, Shakopee, MN 55379-4451	Map
Eric Schmidt	(815) 522-6299	413 Prairie St, Kirkland, IL 60146-0000	Map
Eric Schmidt	(636) 942-3494	6404 Lipizzaner Dr, Imperial, MO 63052-4142	Map
Eric Schmidt	(618) 644-9277	2260 Steinkoenig School R, Highland, IL 62249-4006	Map
Eric Schmidt	(715) 324-5232	N17082 Roth Ln, Pembine, WI 54156-0000	Map
Eric Schmidt	(360) 854-0973	24400 Mckendree Ln, Sedro Woolley, WA 98284-7814	Map
Eric Schmidt	(650) 964-4017	Mountain View, CA 94043-0000	Map
Eric Schmidt	(207) 848-2221	41 Hardwood Dr, Hermon, ME U4401-U253	Map

Google CEO - Eric Schmidt



Google Code Search



VULNS IN OPEN SOURCE CODE

- Regex search for vulnerabilities in public code
- Example: SQL Injection in ASP querystring
 - `select.*from.*request\..QUERYSTRING`

The screenshot shows a Google Code Search interface. The search bar contains the query `select.*from.*request\..QUERYSTRING`. The search results show a file named `post.asp` with the following code snippet:

```
45: strSql = "SELECT * from reply where reply_id = " & Request.QueryString("reply_id")
46: msg = "<br><br>×çôâ±°ô»óð,ÂîÄöÂ×-ôß°í²ùÀíô±²ÂÄÜ±²ôâ,øìù×ó."
57: strSql = "SELECT T Message from Topics where Topic_id = " & Request.QueryString("reply_id")
58: msg = "<br><br>×çôâ±°ô»óð,ÂîÄöÂ×-ôß°í²ùÀíô±²ÂÄÜ±²ôâ,øìù×ó."
```

A red callout box points to the `reply_id` parameter in the SQL query, stating: `reply_id` is SQL injectable querystring parameter. The search results also indicate that there are about 2,000 results for this query.

GOOGLE CODE SEARCH HACKING

DEMO

SHODAN

HACKER SEARCH ENGINE

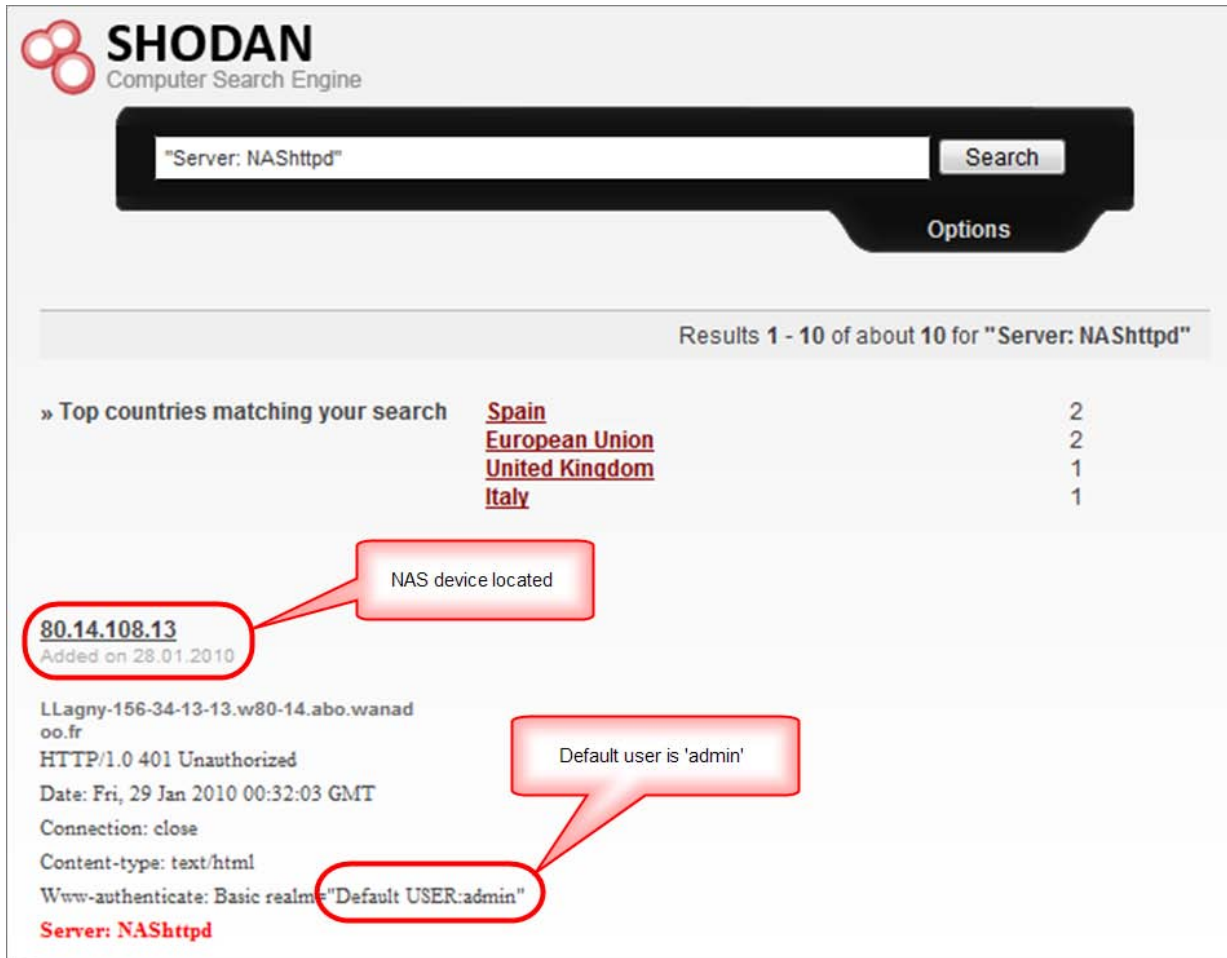


SHODAN Computer Search Engine

- Scans and probes the Internet for open HTTP ports and indexes the headers returned in the response
- Profile a target without directly probing their systems
- Discover specific network appliances
- **Easily find vulnerable systems!**



Target NAS Appliances



SHODAN
Computer Search Engine

Search: "Server: NAShttpd" [Search] [Options]

Results 1 - 10 of about 10 for "Server: NAShttpd"

» Top countries matching your search

Spain	2
European Union	2
United Kingdom	1
Italy	1

80.14.108.13
Added on 28.01.2010

NAS device located

LLagny-156-34-13-13.w80-14.abo.wanad
oo.fr
HTTP/1.0 401 Unauthorized
Date: Fri, 29 Jan 2010 00:32:03 GMT
Connection: close
Content-type: text/html
Www-authenticate: Basic realm="Default USER:admin"

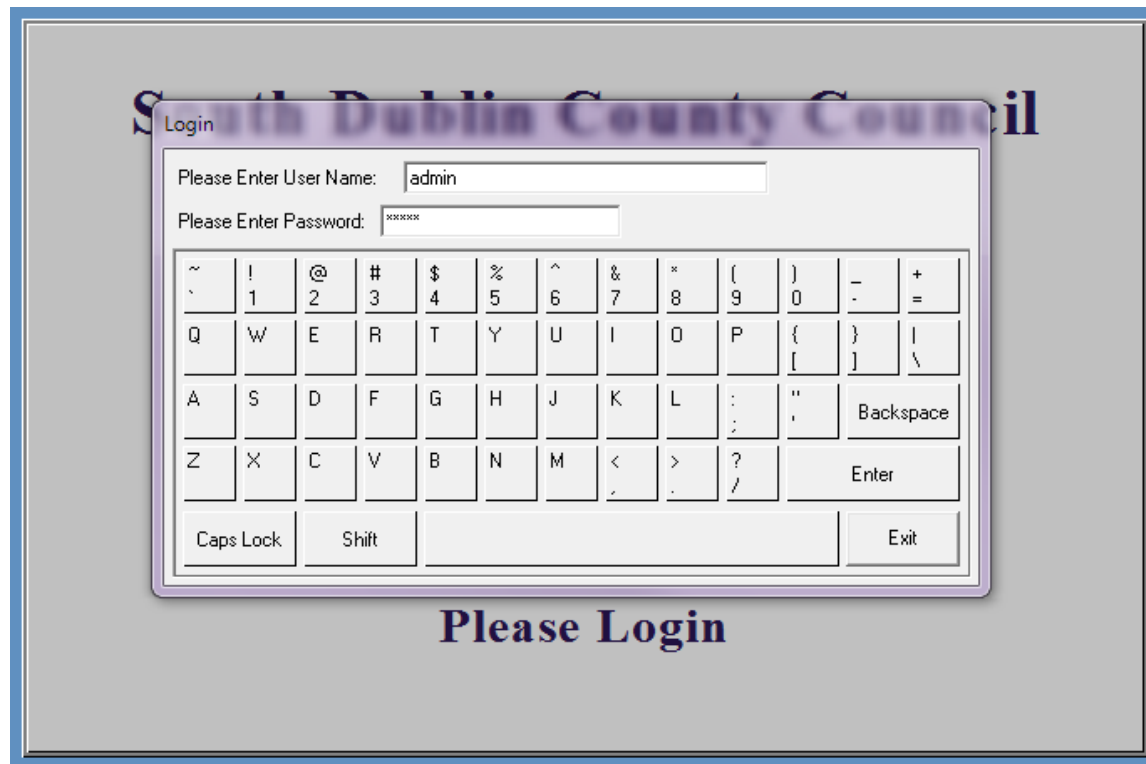
Default user is 'admin'

Server: NAShttpd

Target SCADA

CRITICAL INFRASTRUCTURE SECURITY

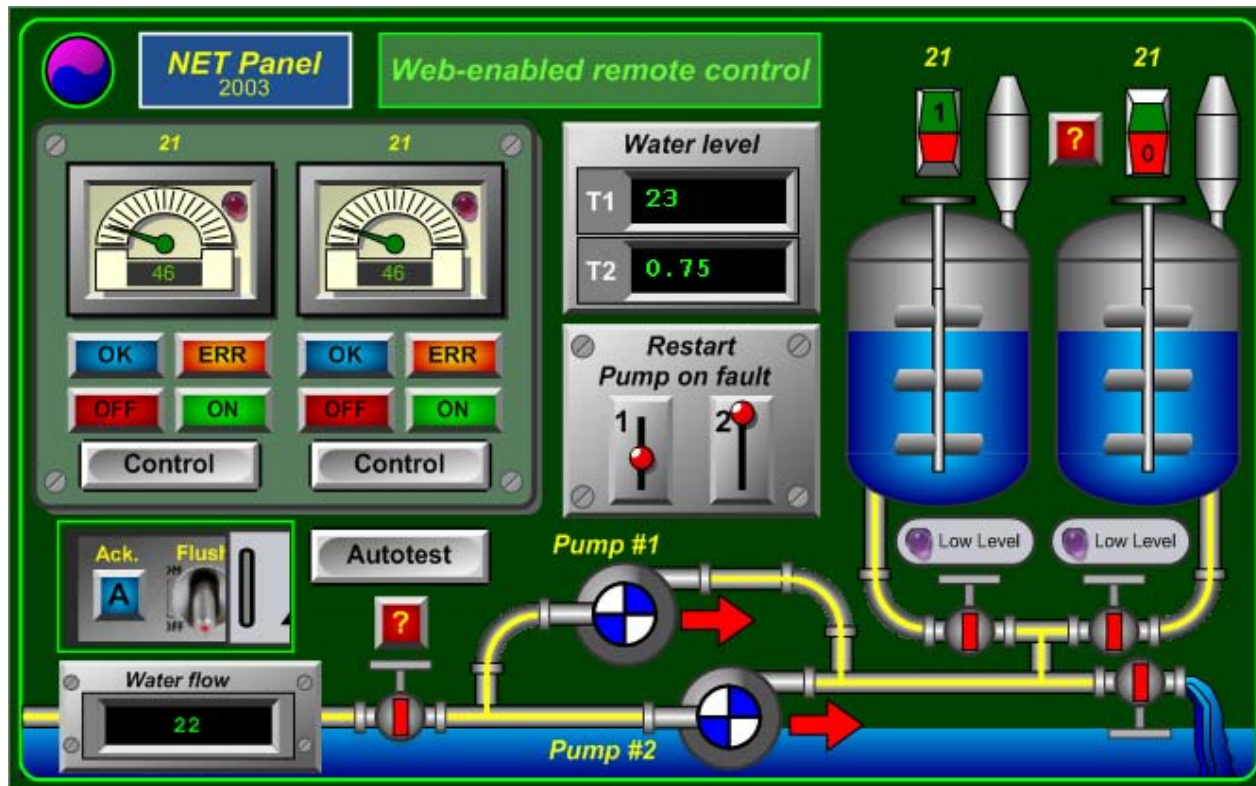
- Supervisory control and data acquisition



Target SCADA

CRITICAL INFRASTRUCTURE SECURITY

- SHODAN: Target Aquired!



Black Hat SEO

SEARCH ENGINE OPTIMIZATION



- Why use real news events?
- Black hats make their own fake news
- Faux celebrity sex tape anyone?
- Send to college students
- It works!
- Other scammers imitate what works

Google Trends



BLACK HAT SEO RECON

The screenshot shows the Google Insights for Search interface. The 'Compare by' section has 'Search terms' selected. The 'Search terms' input field contains 'All search terms'. The 'Filter' section is set to 'Web Search', 'United States', 'All subregions', 'All metros', '2004 - present', and 'All Categories'. A red callout box points to the '2004 - present' filter with the text 'Top Google searches over past 6 years'. Below the filters, the 'Web Search Interest' section shows 'United States, 2004 - present'. The 'Search terms' table lists 'lyrics' as the top search, circled in red. A red callout box points to this entry with the text 'Fake lyrics web sites setup to appear in Google search results, infect you with malware upon clicking'. To the right, a news snippet is visible with the headline 'Lada Gaga, Rihanna lyrics sites used to foist Java exploit' and the sub-headline 'As expected, virus writers now are actively exploiting a zero-day Sun...'. A red arrow points from the 'lyrics' search term to this news snippet.

Google Insights for Search beta

Help | Sign in | Download as CSV | English (US)

Compare by

- Search terms
- Locations
- Time Ranges

Search terms

Tip: Use a comma as shorthand to add comparison items. (tennis, squash)

All search terms

Filter

Web Search

United States | All subregions | All metros

2004 - present

All Categories

Search

Web Search Interest

United States, 2004 - present

Search terms

Top searches

- lyrics
- you
- yahoo

Lada Gaga, Rihanna lyrics sites used to foist Java exploit

Dan Kaplan April 14, 2010

PRINT | EMAIL | REPRINT | PERMISSIONS | FONT SIZE: A | A | A

As expected, virus writers now are actively exploiting a zero-day Sun...

RELATED ARTICLES

Defenses



BLACKHAT SEO DEFENSES

- Google SafeBrowsing plugin
- No-script and Ad-block browser plugins
- Install software security updates
- Stick to reputable sites!
 - Google results aren't safe.

Metadata Attacks



DATA ABOUT DATA

- It's everywhere!
 - In documents (doc, xls, pdf)
 - In images
- What can be data mined?
 - Usernames, emails
 - File paths
 - Operating systems, software versions
 - Printers
 - Network information
 - Device information

FOCA



AUTO METADATA MINING

- Automated doc search via Google/Bing
- Specify domains to target
- Automated download and analysis of docs

The screenshot shows the FOCA RC3.0.1 application window. The interface includes a menu bar with 'File', 'Search files', 'Analyze Metadata', 'Options', and 'About'. A left sidebar shows a tree view of the project structure, including 'CERN', 'Project', 'Configuration', 'Documents (1263)', 'Liste-tel-513-v2.xls', 'Users', 'Dates', 'Other Metadata', and 'MetaData' (with sub-items: Users (1), Folders (0), Printers (0), Software (0), Emails (0)). The main area features the FOCA logo, a 'METASHIELD Protector' banner with the text 'Protect your website against DLP through office documents', and a 'Custom search' link. On the right, there are sections for 'Search engines' (Google, Bing) and 'Extensions' (doc, ppt, pps, xls, docx, pptx, ppsx, xlsx, xds, docx, sxi, odt, ods, odg, odp, pdf, wpd). A table at the bottom displays search results with columns for Id, Type, URL, Download, Size, Analyzed, and M.

Id	Type	URL	Download	Size	Analyzed	M
117	xls	http://www.cern.ch/sm18-public/dipole/to%20do%20list/TO%20DO%20LIST%20FOR%2...	×	24.5 KB	×	-
118	xls	http://www.cern.ch/sm18-public/sss/template/Template_OHMS_CQR.xls	×	214.5 KB	×	-
119	xls	http://www.cern.ch/cms_tmb/PERSONAL/Vanhala/bigwheel/Photoreports/disk_def.xls	×	122 KB	×	-
120	xls	http://www.cern.ch/cms_tmb/PERSONAL/Vanhala/bigwheel/Photoreports/exterior_whe...	×	264.5 KB	×	-
121	xls	http://www.cern.ch/cms_tmb/PERSONAL/Vanhala/bigwheel/Photoreports/interior_whe...	×	245.5 KB	×	-
122	xls	http://www.cern.ch/cms_tmb/PERSONAL/Vanhala/bigwheel/Photoreports/tilted_wheel...	×	219.5 KB	×	-

Defenses



METADATA MINING DEFENSES

- Implement a policy to review files for sensitive metadata before they're released
- Run metadata extraction tools on your resources
- Utilize metadata cleaning tools
- Digital Rights Management (DRM) tools

Advanced Defenses

PROTECT YO NECK



Existing Defenses

"HACK YOURSELF"



- ✓ Tools exist
- ✗ Convenient
- ✗ Real-time updates
- ✗ Multi-engine results
- ✗ Historical archived data
- ✗ Multi-domain searching

Advanced Defenses

NEW HOT SIZZLE



Stach & Liu now proudly presents:

- Google Hacking Alerts
- Bing Hacking Alerts

ADVANCED DEFENSE TOOLS

DEMO

Advanced Defenses

GOOGLE HACKING ALERTS



Google Hacking Alerts

- All GHDB/FSDB regexs using **Google alerts**
- Real-time vuln updates to 1643 hack queries via RSS
- Organized and available via **Google reader** importable file

stachliu0@gmail.com | [Settings](#) | [FAQ](#) | [Sign out](#)

Google alerts Manage your Alerts

Your Google Alerts [Switch to text emails](#) | [Export alerts](#)

Search terms	Type	How often	Email length	Deliver to
<input type="checkbox"/> !Host=*.intext.enc_UserPassword=* ext:pcf	Web	as-it-happens	up to 50 results	Feed View in Google Reader edit
<input type="checkbox"/> "# Dumping data for table (username user users password)"	Web	as-it-happens	up to 50 results	Feed View in Google Reader edit
<input type="checkbox"/> "# Dumping data for table"	Web	as-it-happens	up to 50 results	Feed View in Google Reader edit
<input type="checkbox"/> "# phpMyAdmin MySQL-Dump" "INSERT INTO" -"the"	Web	as-it-happens	up to 50 results	Feed View in Google Reader edit

Advanced Defenses

GOOGLE HACKING ALERTS



Google reader

All items Search

+ Add a subscription

Home

All items (1000+)

Starred items ☆

Your stuff

Trends

Browse for stuff

People you follow

Explore

Subscriptions

- Advisories and Vulner... (1000+)
- Error Messages (1000+)
- Files containing juic... (1000+)
- Files containing pass... (668)
- Files containing user... (184)

Google Alerts - "mysql error with query"

Items - all items Mark all as read Refresh Feed settings...

James Bond needs help!
mysql error page snippet conveniently provided in RSS summary

Several thousand GHDB vuln alerts generated in a day

James Bond 007 :: MI6 - The Home Of James Bond

mysql error with query SELECT c.citem as itemid, c.cnumber as commentid, c.cbody as body, c.cuser as user, c.cmail as userid, c.cemail as email, ...
www.mi6.co.uk/mi6_php3/news/index.php?itemid...t...

Remove star Like Share Share with note Email Keep unread Edit tags: Error Messages

わかの奇妙な日常 - mysql error with query SELECT COUNT(*) AS result FROM nucleus_actionlog: Can't open file: 'nucleus_actionlog.M

mysql error with query SELECT * FROM nucleus_blog WHERE bnumber=1 ... - mysql error with query SELECT * FROM nucleus_ca

Advanced Defenses

BING HACKING ALERTS



Bing Hacking Alerts

- Bing searches with regexs from BHDB
- Leverage '**&format=rss**' directive to turn into update feeds

Google reader All items

Navigation **Bing: "mysql error with query" »** Show: Expanded - Lis

Show: 5 new items - all items [show details](#)

☆ www.kloosterman.be - mysql error with query SELECT p.pfile as pfile, e.event as event FROM	Apr 13, 2010	»
☆ The Shadow Project - Blog - mysql error with query SELECT COUNT(*) FROM nucleus_comment as c WHERE	Apr 13, 2010	»
☆ Hou-Hou Blog : tunisie blogs - mysql error with query SELECT i.inumber as itemid, i.iblog as blog, i.ititle as title, i.ibody	Apr 13, 2010	»
☆ Hou-Hou Blog : tunisie - mysql error with query SELECT i.inumber as itemid, i.iblog as blog, i.ititle as title, i.ibody as	Apr 13, 2010	»
☆ www.radiosonic.it - mysql error with query SELECT * FROM nucleus_config: Table 'Sql99301_1.nucleus_config'	Apr 13, 2010	»
☆ Hou-Hou Blog : george bush - mysql error with query SELECT i.inumber as itemid, i.iblog as blog, i.ititle as title, i.ibody as	Apr 7, 2010	»
☆ PHP /MYSQL - Error with query - ClanTemplates - PHP /MYSQL - Error with query Programming ... Programming Got	Apr 5, 2010	»
☆ www.tutje.nl - mysql error with query SELECT * FROM nucleus_config: Table 'poiplgqn_tutje.nucleus_config' doesn't exist	Apr 5, 2010	»

Future Direction

PREDICTIONS



Future Directions

PREDICTIONS



Data Explosion

- More data indexed, searchable
- Real-time, streaming updates
- Faster, more robust search interfaces

Google Involvement

- Filtering of search results
- Better GH detection and tool blocking

Renewed Tool Dev

- Google Ajax API based
- Bing/Yahoo/other engines
 - Search engine aggregators
- Google Code and Other Open Source Repositories
 - MS CodePlex, SourceForge, ...
- More automation in tools
 - Real-time detection and exploitation
 - Google worms

Future Directions

REAL-TIME UPDATES



Google obama Search [Advanced Search](#)

Web > Updates [Hide options](#) Results 1 - 10 of about 4 for obama. (0.58 seconds)

[All results](#)
[Images](#)
[Videos](#)
[News](#)
[Blogs](#)
Updates
[Books](#)
[Discussions](#)

[Any time](#)
Latest
[Reset options](#)

2010 > April > 20 - 21

7am 1pm 7pm 1am

New results will appear below as they become available. [Pause](#)

Helen Thomas on her one question for **Obama**
[YouTube - Helen Thomas on her one question for Obama](#) - youtube.com

[Idanah](#) - **Twitter** - 1 minute ago

Obama falters on immigration reform promises
m #usnews #news
n reform, Obama's priorities shift -
latimes.com - latimes.com

[filterednews](#) - **Twitter** - 1 minute ago

Top links

[Obama to discuss Supreme Court pick with party leaders - CNN.com](#)
President **Obama** is expected to meet with key Republican and Democratic leaders Wednesday to discuss a replacement for retiring Supreme ...
<http://www.cnn.com/2010/.../jobama.../index.html>
[All mentions >](#)

[Obama Supreme Court Pick: President Talking With Possible High ...](#)
WASHINGTON — Pushing forward with one of his most consequential decisions, President Barack **Obama** has begun informal talks with ...
<http://www.huffingtonpost.com/.../jobama-supreme->

Questions?

I'll try to answer it...

For more info:

Email: contact@stachliu.com

Stach & Liu, LLC

www.stachliu.com

Thank You

