

# F9 - Google and Beyond

Advanced Search Engine Hacking and Web-Based Intelligence Gathering  
21 April 2010



Presented by:  
Francis Brown and Rob Ragan  
Stach & Liu, LLC  
[www.stachliu.com](http://www.stachliu.com)

# Agenda

## OVERVIEW

- Introduction
- 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 think differently* about exposures caused by 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

Dates	Event
2004	Google Hacking Database (GHDB) begins
May 2004	Foundstone SiteDigger v1 released
2005	Google Hacking v1 released by Johnny Long
Jan. 2005	Foundstone SiteDigger v2 released
Feb. 13, 2005	Google Hack HoneyPot first release
Jan. 10, 2005	MSNPawn v1.0 released
Dec. 5, 2006	Google stops issuing Google SOAP API keys
...	...

# Quick History

## GOOGLE HACKING RECAP

Dates	Event
Mar. 2007	Bing disables inurl: link: and linkdomain:
Nov. 2, 2007	Google Hacking v2 released
Mar. 2008	cDc Goolag - gui tool released
June 3, 2009	Bing goes online
Sept. 7, 2009	Google shuts down SOAP Search API
Nov. 2009	Binging tool released
Dec. 1, 2009	FoundStone SiteDigger v 3.0 released
2010	Googlag.org disappears

# 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

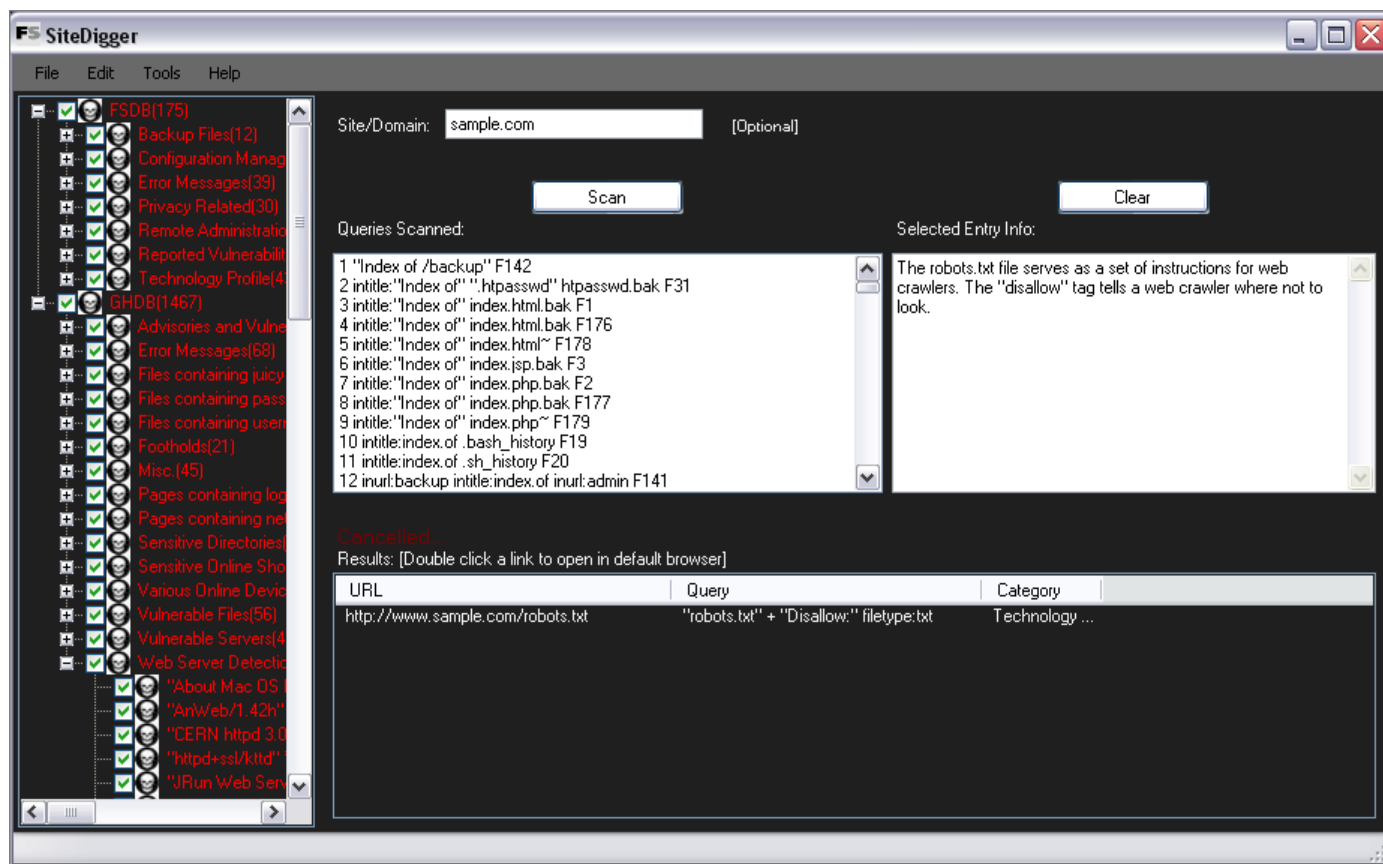
# Google Hacking Toolkit

## 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 no longer around

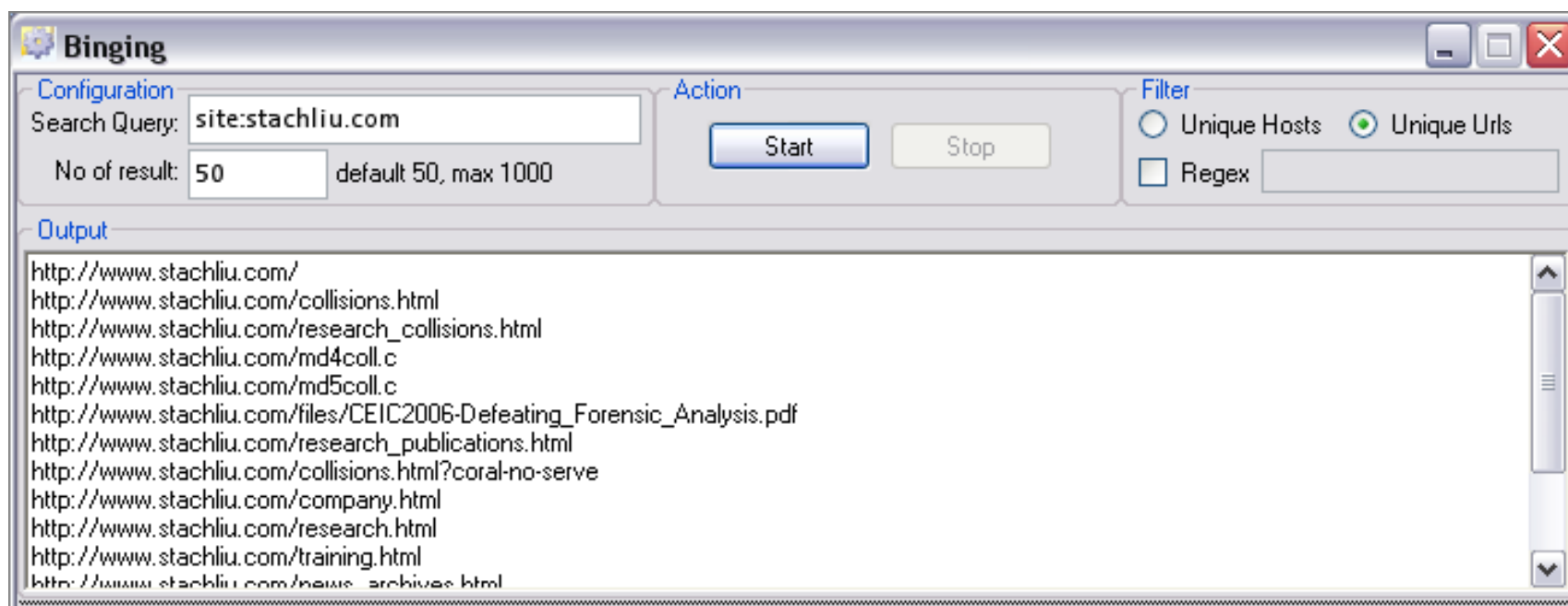
# Google Hacking Toolkit

FOUNDSTONE SITEDIGGER



# Google Hacking Toolkit

BINGING



NEW GOOGLE HACKING TOOLS

**DEMO**

# New Toolkit

STACH & LIU TOOLS

## GoogleDiggity

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

## BingDiggity

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



# New Toolkit

## GOOGLEDIGGITY

```
C:\GoogleDiggity\bin\Debug>GoogleDiggity.exe /?

+=====+
|           GoogleDiggity - Version 0.1           |
|   Google Hacking Utility - Using Google AJAX API   |
|           By: Fran Brown and Rob Ragan           |
|   Stach & Liu, LLC - http://www.stachliu.com     |
+=====+

** Note: Google AJAX API won't block you, but does limit you to 64 results per query.

Usages: GoogleDiggity.exe [-q Google_Search_Query] [-d Target_Domain] [-k Google_CustomSearchEngine_ID]
        GoogleDiggity.exe [-f Google_Search_Query_File] [-d TargetDomain] [-k GoogleCustomSearchEngineID]

-q [Google search query]
  Optional, can alternatively load queries from file using the -f option.
-f [Google Hack queries file]
  Optional, text file contain Google search queries, one per line.
  GHDB and FSDB txt files provided (e.g. "Error Messages.txt", "Files containing passwords.txt", etc.).
-d [Domain to search]
  Optional, example: whatever.com
-k [Google Custom Search Engine ID]
  Optional, this can also be set in the configuration file.
  Note, to use the Google CSE ID in the config file, specify the -k option without any params.
  For more info, go to http://www.google.com/cse/

Examples:
GoogleDiggity.exe -q test -d stachliu.com
GoogleDiggity.exe -f "Error Messages.txt" -k 006113852371342063645:ytae6rpigi4
GoogleDiggity.exe -f "Error Messages.txt" -k
```

# New Toolkit

## BINGDIGGITY

```
C:\BingDiggity\bin\Debug>BingDiggity.exe /?

+=====+
|           BingDiggity - Version 0.1           |
|       Bing Hacking Utility - Using Bing 2.0 API       |
|       By: Fran Brown and Rob Ragan                 |
|       Stach & Liu, LLC - http://www.stachliu.com      |
+=====+

Usages: BingDiggity.exe [-q Bing_Search_Query] [-d Target_Domain] [-m Max_Search_Results] [-k Bing_2.0_App_ID]
        BingDiggity.exe [-f Bing_Search_Query_File] [-d TargetDomain] [-k Bing_2.0_App_ID]
        BingDiggity.exe [-i IP_Address | -r IP_Address_Range] [-m Max_Search_Results] [-k Bing_2.0_App_ID]

-q [Bing search query]
  Optional, can alternatively load queries from file using the -f option.
-f [Bing Hack queries file]
  Optional, text file contain Bing search queries, one per line.
-d [Domain to search]
  Optional, example: whatever.com
-i [IP address to search]
  Optional, example: 10.1.1.1
-r [IP address range to search]
  Optional, example: 10.1.1.1-10.1.1.15
-m [Max # of results to return per query]
  Optional, note Bing 2.0 doesnt limit your # of queries per day.
-k [Bing 2.0 App ID]
  Required, this can also alternatively be set in the configuration file.
  To obtain a Bing 2.0 App ID, go to http://www.bing.com/developer

Examples:
BingDiggity.exe -q "mySQL error with query"
BingDiggity.exe -d stachliu.com -k 2136A233ASDFGHC0CB87C73B9946PFJRYGBVFD32
BingDiggity.exe -q test -d stachliu.com
BingDiggity.exe -i 10.1.1.1
BingDiggity.exe -r 10.1.1.1-10.1.1.15
```

# 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
  - Free Tools: OpenDLP, Senf
- 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 buzz 

Google™  
code search labs

Google™ code

Google health

Google calendar

Google news

Google™ public data explorer  
labs

Google docs

Google™ Insights for Search  
beta

Google™ wave   
preview

Google blogs

Google maps

Google groups

# Google PhoneBook



## SPEAR PHISHING

Google PhoneBook

phonebook: schmidt, eric

Search PhoneBook Search the Web Preferences

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

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

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 result for the query `select.*from.*request\.QUERYSTRING`. The search results are displayed under the heading "Code". A red callout box points to the search query, stating "reply\_id is SQL injectable querystring parameter". 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>×çðâ±°ö»óð,ÃîÃðÃ×-ðß°í²ùÀíô±²ÅÄÜ±²Öá,øìú×ó."

57: strSql = "SELECT T Message from Topics where Topic_id = " & Request.QueryString("reply_id")
58: msg = "<br><br>×çðâ±°ö»óð,ÃîÃðÃ×-ðß°í²ùÀíô±²ÅÄÜ±²Öá,øìú×ó."
```

The code snippet shows a SQL injection vulnerability in the `reply_id` parameter of the `Request.QueryString` function. The search results also show the source URL: [www.cnarts.net/eweb/download/software/bbs/tradeforum.zip](http://www.cnarts.net/eweb/download/software/bbs/tradeforum.zip) - Unknown - ASP - [More from tradeforum.zip](#) »

GOOGLE CODE SEARCH HACKING

**DEMO**

# SHODAN

HACKER SEARCH ENGINE

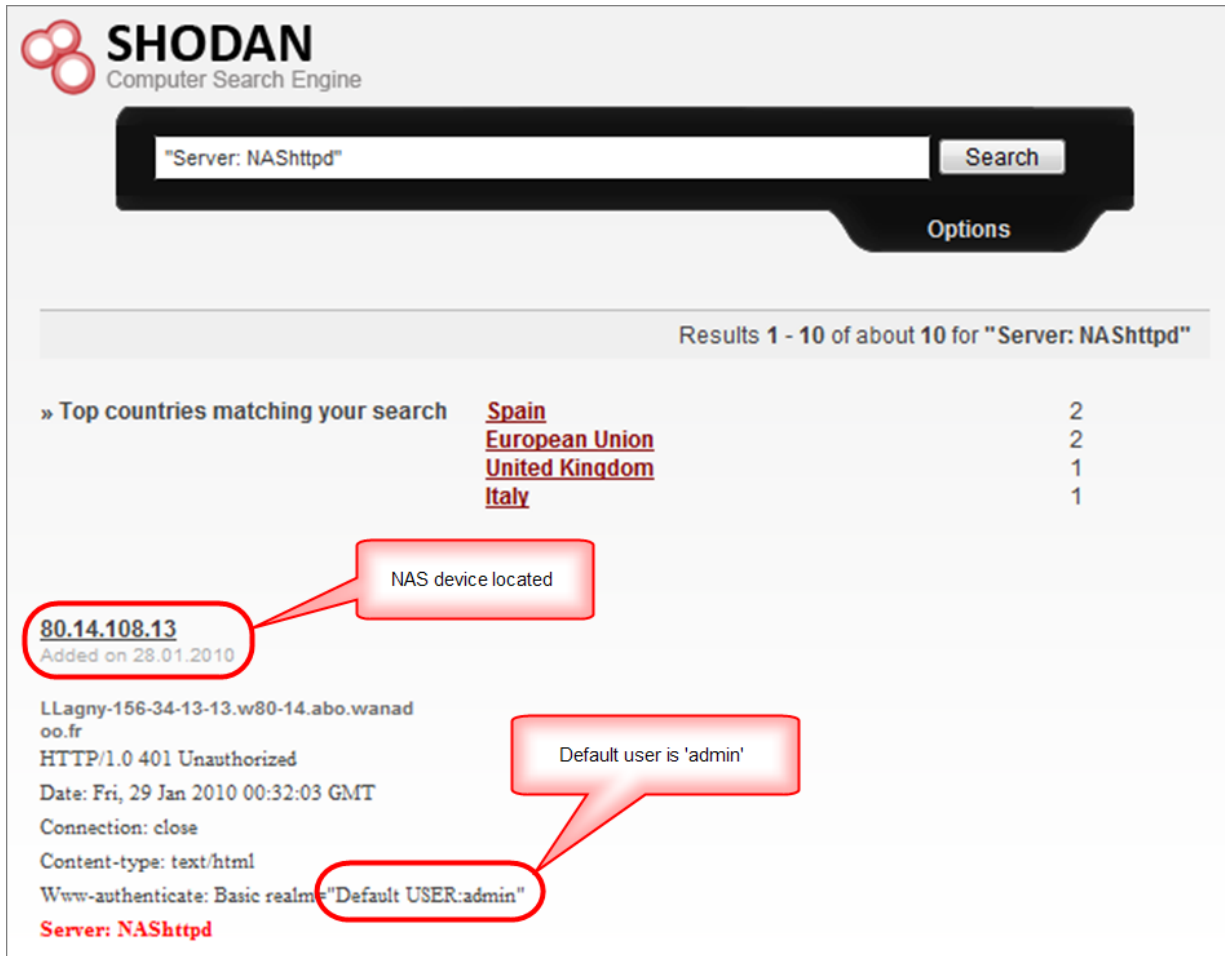
## SHODAN Computer Search Engine

- 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

<a href="#">Spain</a>	2
<a href="#">European Union</a>	2
<a href="#">United Kingdom</a>	1
<a href="#">Italy</a>	1

**80.14.108.13**  
Added on 28.01.2010

NAS device located

LLagny-156-34-13-13.v80-14.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

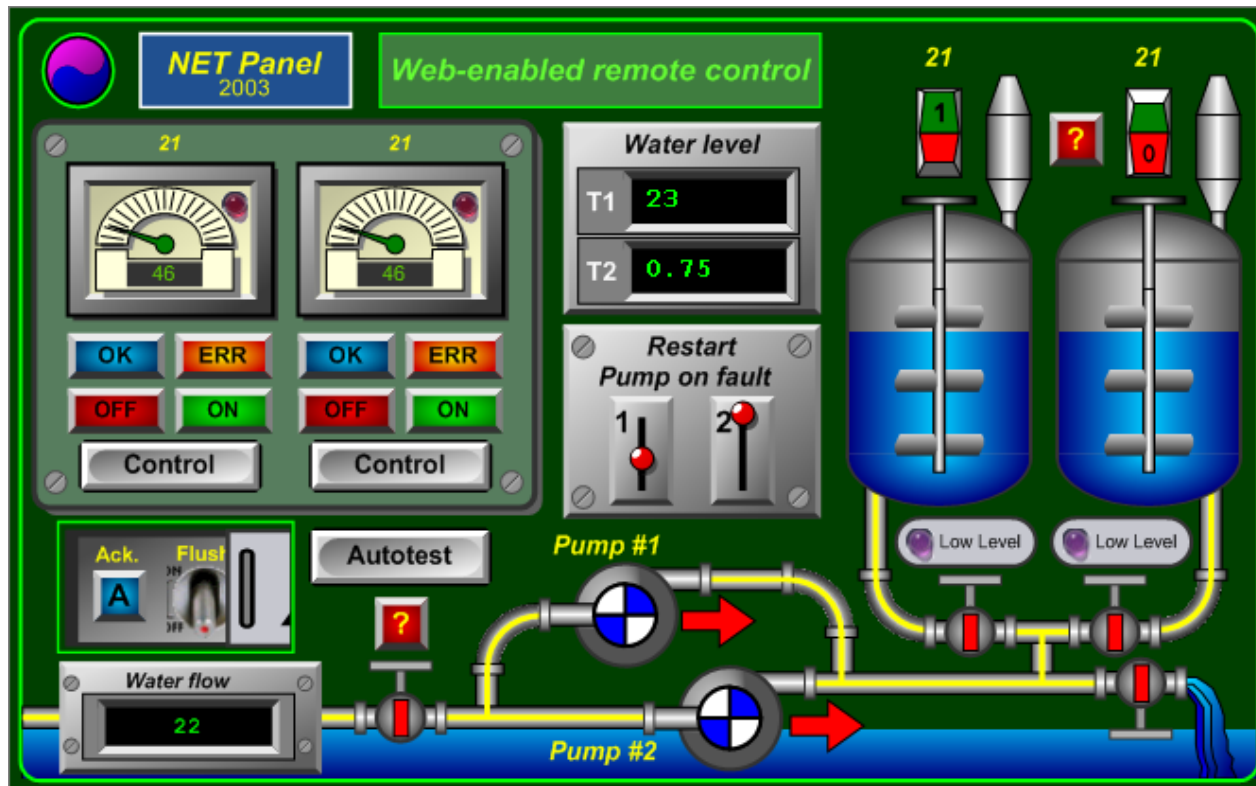
- Supervisory control and data acquisition



# Target SCADA

CRITICAL INFRASTRUCTURE SECURITY

- SHODAN: Target Acquired!



# 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

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)

**Filter**  
Web Search  
United States | All subregions | All metros  
2004 - present  
All Categories

**Top Google searches over past 6 years**

**Web Search Interest**  
United States, 2004 - present

**Search terms**

**Top searches**

1. [lyrics](#)
2. [you](#)
3. [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  
vulnerable info routers through drive

RELATED ARTICLES

**Fake lyrics web sites setup to appear in Google search results, infect you with malware upon clicking**

**Top Google searches over past 6 years**

# Defenses

## BLACKHAT SEO DEFENSES

- Google SafeBrowsing plugin
- Microsoft SmartScreen Filter
- 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

**Search engines**

- Google
- Bing

**Extensions**

- doc
- ppt
- pps
- xls
- docx
- pptx
- ppsx
- xlsx
- sxx
- sxc
- sxi
- odt
- ods
- odg
- odp
- pdf
- wpd

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 1623 hack queries via RSS
- Organized and available via **Google reader** importable file

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

**Google alerts** Manage your Alerts

GHDB regexs made into Google Alerts

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

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

RSS Feeds generated that track new GHDB vulnerable pages in real-time

# 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...](http://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

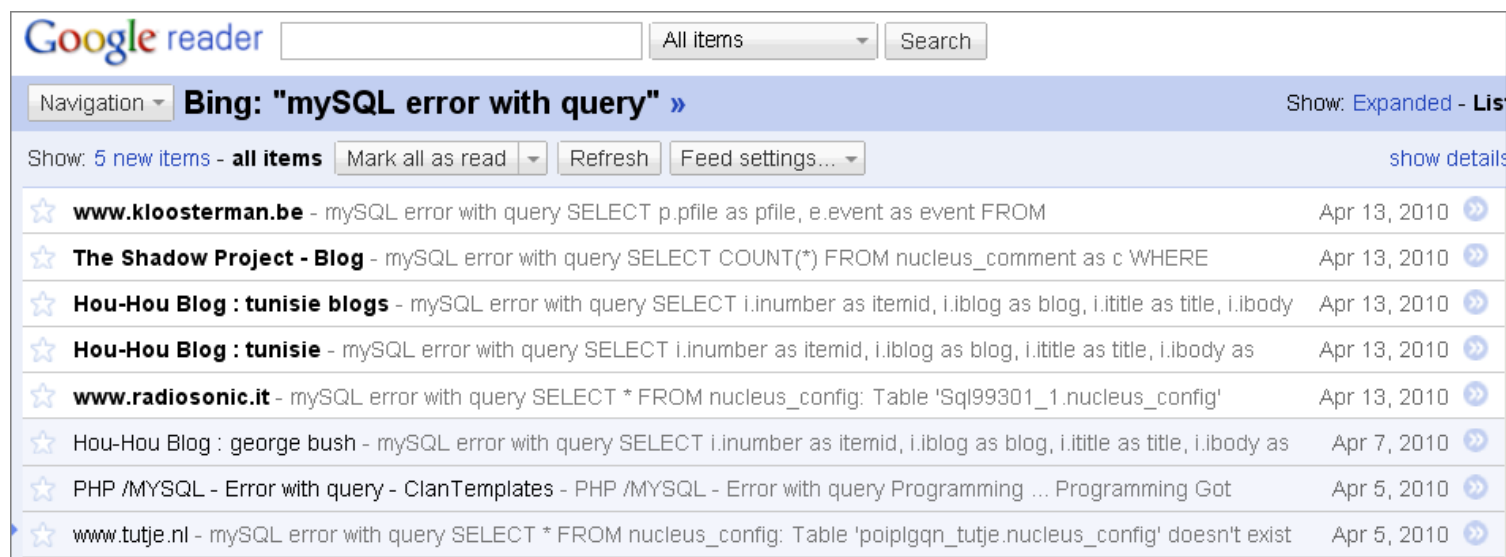
mysql error with query SELECT \* FROM nucleus\_blog WHERE bnumber=1 ... - mysql error with query SELECT \* FROM nucleus\_ca

# Advanced Defenses

## BING HACKING ALERTS

### Bing Hacking Alerts

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



The screenshot shows a Google Reader interface. At the top, there's a search bar and a dropdown menu set to 'All items'. Below that, the feed title is 'Bing: "mysql error with query"'. The feed contains several items, each with a star icon, a source name, a snippet of a MySQL error message, and a date. The items are:

Source	Snippet	Date
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

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  
[Obama falters on immigration reform promises](#) - 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-](#)

Real-time updates!

Questions?  
Ask us something  
We'll try to answer it.

For more info:  
Email: [contact@stachliu.com](mailto:contact@stachliu.com)  
Stach & Liu, LLC  
[www.stachliu.com](http://www.stachliu.com)

# Thank You