

# Modern information gathering

**Onderwerp: Modern Information Gathering**

**Datum: 26-JUN-2012**

**Aanwezigen: OWASP**

**Classificatie: Public**

# Who Am I

Dave van Stein

38 years

Tester > 11 years

(Application) Security Testing

"Certified Ethical Hacker"



# Agenda

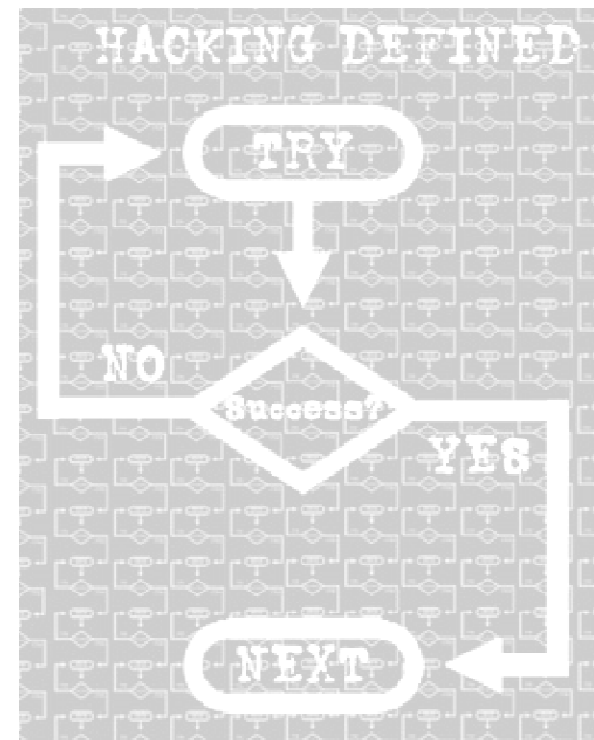
Goal of the presentation  
What is Information Gathering ?  
Domain scanning  
Search engine 'abuse'  
Other tools  
Some Social Engineering  
Remedies  
Conclusions

# Goal of this presentation

Give insight in amount of information anonymously available on internet about your system (and users)

Give insight in the amount and possibilities of tools freely available

Identify entrypoint  
Gain access  
Secure access  
Do stuff  
Clear up the mess  
Come back another time  
(simplified procedure)



# 'Classic' Domain Scanning

## Steps involved:

- Get network information with ping and traceroute
- Get DNS information with WHOIS and LOOKUP
- Do DNS zone transfer for subdomains
- Download website for extra info
- Scan servers

## Problems:

- DNS zone transfers often not authorized
- Active connection with target => detectable



```
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\WINDOWS>tracert www.google.com

Tracing route to www.l.google.com [66.249.91.103]
over a maximum of 30 hops:

  0  1 ms    1 ms    1 ms   192.168.10.1
  1  58 ms   30 ms   31 ms   ip1-144-173-82.ads12.static.versatel.nl [82.173.
144.11]
  2  73 ms   30 ms   31 ms   ge-0-1-0-1305.ncr01as2.versatel.net [217.16.44.
49]
  3  57 ms   32 ms   39 ms   ge-1-3-0-666.br01sara.versatel.net [212.53.18.18]
  4  32 ms   32 ms   31 ms   cnv1.ams.net.google.com [195.69.144.247]
  5  32 ms   32 ms   31 ms   209.85.248.93
  6  41 ms   35 ms   47 ms   64.233.175.246
  7  41 ms   35 ms   93 ms   209.85.255.23
  8  41 ms   35 ms   83 ms   66.249.94.146
  9  35 ms   35 ms   85 ms   ik-in-f103.google.com [66.249.91.103]

Trace complete.

C:\WINDOWS>
```

# Modern Information Gathering

## Interesting information:

- Domains and subdomains
- IP addresses
- Applications and technologies
- Hotspots (known vulnerabilities)
- Username and passwords
- Sensitive information

## Passive

- As little contact as possible with target
- No direct scanning, no intrusion
- No logging and no alarm triggering !

# Sources of information

Public records

WHOIS: information about owner

DNS : information about IP addresses

Search engines

Often little restrictions on websites

Cache all information gathered

Tweaking provides additional information

Various websites

Anonymous

Combine above techniques

Sort results for nice presentation

Advanced and Automated  
scanning

Specialized (offline) Tools

## Shodan

IP addresses

Server banner

X-Powered-by banner

Cookies

## Search filters

City, Country, Geo

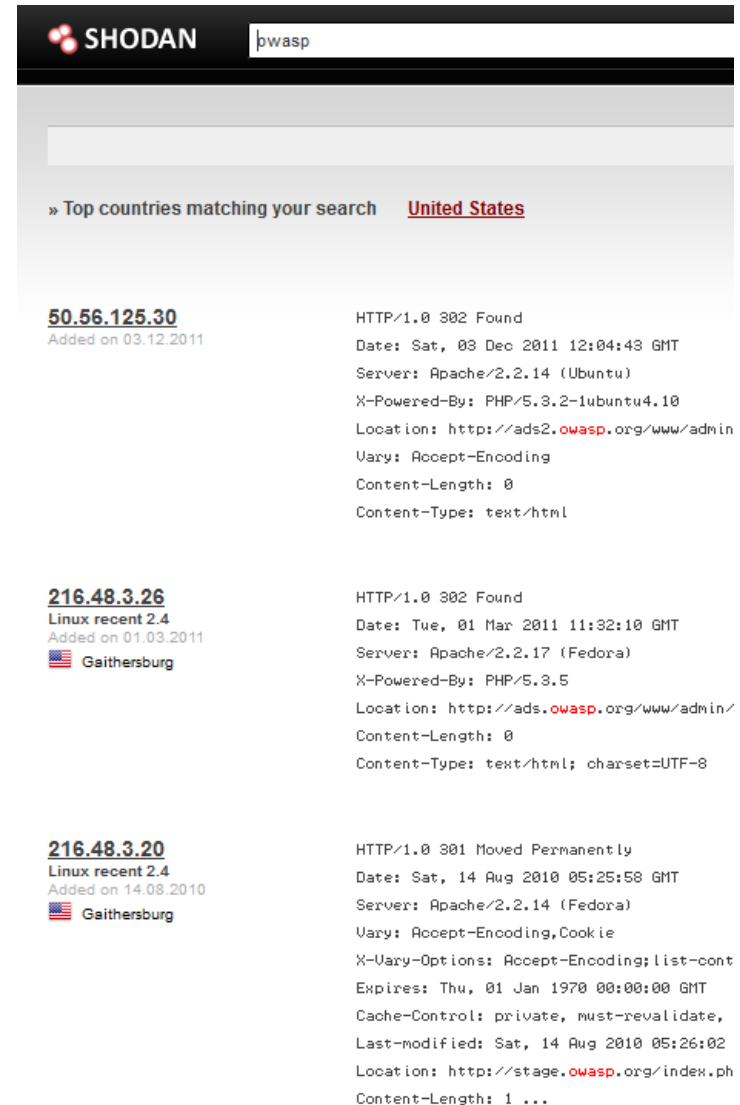
Hostname, ip address / net block

Os, port

date (before / after)



ssl cert version, bits, issuer

ssl cipher support, bit support , protocol



SHODAN owasp

» Top countries matching your search [United States](#)

<b><u>50.56.125.30</u></b> Added on 03.12.2011	HTTP/1.0 302 Found Date: Sat, 03 Dec 2011 12:04:43 GMT Server: Apache/2.2.14 (Ubuntu) X-Powered-By: PHP/5.3.2-1ubuntu4.10 Location: http://ads2.owasp.org/www/admin/ Vary: Accept-Encoding Content-Length: 0 Content-Type: text/html
<b><u>216.48.3.26</u></b> Linux recent 2.4 Added on 01.03.2011  Gaithersburg	HTTP/1.0 302 Found Date: Tue, 01 Mar 2011 11:32:10 GMT Server: Apache/2.2.17 (Fedora) X-Powered-By: PHP/5.3.5 Location: http://ads.owasp.org/www/admin/ Content-Length: 0 Content-Type: text/html; charset=UTF-8
<b><u>216.48.3.20</u></b> Linux recent 2.4 Added on 14.08.2010  Gaithersburg	HTTP/1.0 301 Moved Permanently Date: Sat, 14 Aug 2010 05:25:58 GMT Server: Apache/2.2.14 (Fedora) Vary: Accept-Encoding, Cookie X-Vary-Options: Accept-Encoding; list-cont Expires: Thu, 01 Jan 1970 00:00:00 GMT Cache-Control: private, must-revalidate, Last-Modified: Sat, 14 Aug 2010 05:26:02 Location: http://stage.owasp.org/index.ph Content-Length: 1 ...



Reports

IP-Tools

Nameserver

Webserver

## Server Sniff

NS reports

Domain reports

Subdomains

Various (trace)routes

Various ping types

Shows robots.txt

Anonymous !



# Domain Scanning: Server Sniff

## Recursive-Queries:

❗ ns1.secure.net. YES - recursive queries allowed!  
❗ ns2.secure.net. YES - recursive queries allowed!

## NS-AXFR:

❗ ns1.secure.net.: anonymous Zonetransfer (AXFR) allowed!!  
❗ ns2.secure.net.: anonymous Zonetransfer (AXFR) allowed!!  
owasp.org. 86400 IN SOA ns1.secure.net. hostmaster.secure.net. 2007080332 86400 7200 2592000 86400  
owasp.org. 86400 IN A 216.48.3.18  
owasp.org. 86400 IN NS ns2.secure.net.  
owasp.org. 86400 IN NS ns1.secure.net.  
owasp.org. 86400 IN MX 30 ASPMX2.GOOGLEMAIL.COM.  
owasp.org. 86400 IN MX 30 ASPMX3.GOOGLEMAIL.COM.  
owasp.org. 86400 IN MX 30 ASPMX4.GOOGLEMAIL.COM.  
owasp.org. 86400 IN MX 30 ASPMX5.GOOGLEMAIL.COM.  
owasp.org. 86400 IN MX 10 ASPMX.L.GOOGLE.COM.  
owasp.org. 86400 IN MX 20 ALT1.ASPMX.L.GOOGLE.COM.  
owasp.org. 86400 IN MX 20 ALT2.ASPMX.L.GOOGLE.COM.  
\*.owasp.org. 86400 IN CNAME owasp.org.  
austin.owasp.org. 86400 IN CNAME owasp.org.  
blogs.owasp.org. 86400 IN CNAME owasp.org.  
calendar.owasp.org. 86400 IN CNAME ghs.GOOGLE.COM.  
docs.owasp.org. 86400 IN CNAME ghs.GOOGLE.COM.  
es.owasp.org. 86400 IN A 216.48.3.18  
google6912a08c3a8cdf0b.owasp.org. 86400 IN CNAME GOOGLE.COM.  
jobs.owasp.org. 86400 IN CNAME owasp.org.  
lists.owasp.org. 86400 IN A 216.48.3.22  
lists.owasp.org. 86400 IN MX 10 lists.owasp.org.  
lists.owasp.org. 86400 IN MX 20 mailhost.rdurkee.COM.  
localhost.owasp.org. 86400 IN A 127.0.0.1  
mail.owasp.org. 86400 IN CNAME ghs.GOOGLE.COM.  
old.owasp.org. 86400 IN A 216.48.3.19  
registration.owasp.org. 86400 IN CNAME owasp.org.  
stage.owasp.org. 86400 IN CNAME owasp.org.  
voip.owasp.org. 86400 IN A 216.48.3.22  
webmail.owasp.org. 86400 IN A 216.48.3.24  
www.owasp.org. 86400 IN CNAME owasp.org.  
owasp.org. 86400 IN SOA ns1.secure.net. hostmaster.secure.net. 2007080332 86400 7200 2592000 86400

owasp.org. 86400 IN MX 20  
\*.owasp.org. 86400 IN CNAME  
austin.owasp.org. 86400 IN  
blogs.owasp.org. 86400 IN  
calendar.owasp.org. 86400 IN  
docs.owasp.org. 86400 IN  
es.owasp.org. 86400 IN A  
google6912a08c3a8cdf0b.owasp.org.  
jobs.owasp.org. 86400 IN  
lists.owasp.org. 86400 IN

## Robtex Swiss Army Knife Internet Tool

In the searchbox above you can search for:

DNS checks detailed dns information for a hostname

( [www.facebook.com](http://www.facebook.com) , [www.yahoo.com](http://www.yahoo.com) , [www.youtube.com](http://www.youtube.com) )

IP-number checks ip number information such as dns reverse and forwards

A-net checks an entire a-network

B-net checks an entire b-network

C-net checks an entire c-network

whois lookup checks whois information for a domain

route checks a specific routed prefix

AS numbers checks information on an AS-number

BGP announcements checks prefixes originated from a specific AS-number

AS macros checks who belongs to an AS-macro

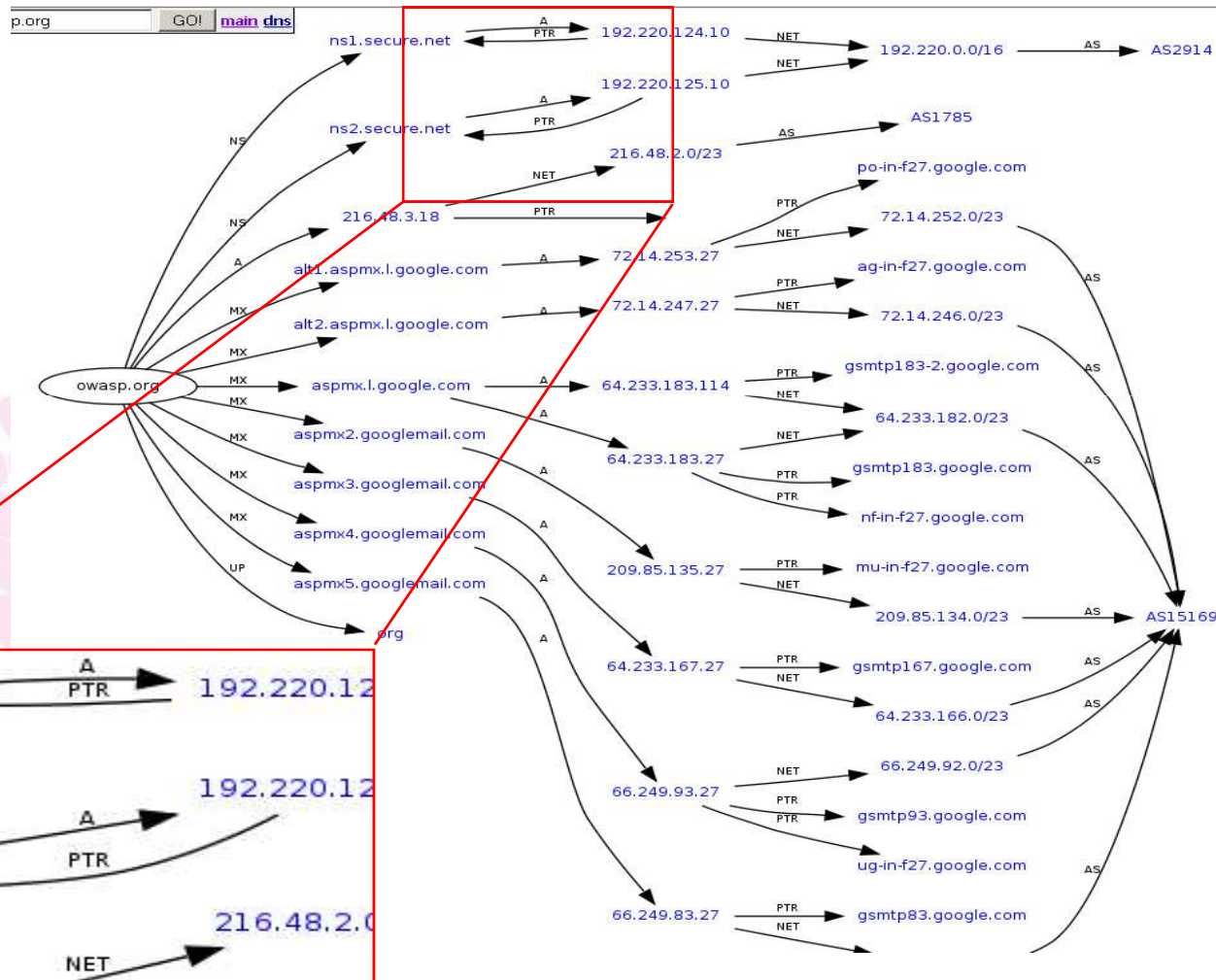
# Domain Scanning: Robtex

Domain 'Swiss Army Knife'

Provides ALL information linked to a domain

base	record	name	ip	reverse	route	as
owasp.org	a		<a href="#">216.48.3.18</a>		<a href="#">216.48.2.0/23</a> Proxy-registered route object	<a href="#">AS1785 FASTNET-ASN-</a>
	ns	<a href="#">ns1.secure.net</a>	<a href="#">192.220.124.10</a>	-	<a href="#">192.220.0.0/16</a> VRIO-192-220	<a href="#">AS2914 NTT GIN AS N</a>
		<a href="#">ns2.secure.net</a>	<a href="#">192.220.125.10</a>	-		
	mx	<a href="#">alt1.aspmx.l.google.com</a>	<a href="#">72.14.253.27</a>	<a href="#">po-in-f27.google.com</a>	<a href="#">72.14.252.0/23</a> Google	<a href="#">AS15169 Google , Inc</a>
		<a href="#">alt2.aspmx.l.google.com</a>	<a href="#">72.14.247.27</a>	<a href="#">ag-in-f27.google.com</a>	<a href="#">72.14.246.0/23</a> Google	
		<a href="#">aspmx.l.google.com</a>	<a href="#">64.233.183.27</a>	<a href="#">gsmtpl83.google.com</a>	<a href="#">64.233.182.0/23</a> Google	
			<a href="#">64.233.183.114</a>	<a href="#">gsmtpl83-2.google.com</a>		
		<a href="#">aspmx2.googlemail.com</a>	<a href="#">209.85.135.27</a>	<a href="#">mu-in-f27.google.com</a>	<a href="#">209.85.134.0/23</a> Google	
		<a href="#">aspmx3.googlemail.com</a>	<a href="#">64.233.167.27</a>	<a href="#">gsmtpl67.google.com</a>	<a href="#">64.233.166.0/23</a> Google	
		<a href="#">aspmx4.googlemail.com</a>	<a href="#">66.249.93.27</a>	<a href="#">gsmtpl93.google.com</a>	<a href="#">66.249.92.0/23</a> Google	
		<a href="#">aspmx5.googlemail.com</a>	<a href="#">66.249.83.27</a>	<a href="#">gsmtpl83.google.com</a>		
ptr		<a href="#">66.35.111.73</a>	-	<a href="#">66.35.111.0/24</a>	<a href="#">AS14955 N-V-C North</a>	
	ns	<a href="#">d0.org.afilias-nst.org</a>	<a href="#">199.19.57.1</a>	-	<a href="#">199.19.57.0/24</a> REACH (Customer Route)	<a href="#">AS12041 AFILIAS NST</a> <a href="#">anycast from several p</a>
		<a href="#">tld1.ultradns.net</a>	<a href="#">204.74.112.1</a>	-	<a href="#">204.74.112.0/24</a> UltraDNS	<a href="#">AS12008 UNSPECIFIED</a>
		<a href="#">tld2.ultradns.net</a>	<a href="#">204.74.113.1</a>	-	<a href="#">204.74.113.0/24</a> UltraDNS	
		<a href="#">a0.org.afilias-nst.info</a>	<a href="#">199.19.56.1</a>	-	<a href="#">199.19.56.0/24</a> REACH (Customer Route)	<a href="#">AS12041 AFILIAS NST</a>
					<a href="#">199.19.54.0/24</a> REACH (Customer	

# Domain scanning: Robtex



# Google Advanced search

**filetype: (or ext:)**

Find documents of the specified type.

*E.g. PDF, XLS, DOC*

**intext:**

The terms must appear in the text of the page.

**intitle:**

The terms must appear in the title of the page.

**inurl:**

The terms must appear in the URL of the page.

## Alternative Query Types

Operators	Meaning
<a href="#">cache:</a>	Display Go
<a href="#">info:</a> (or <a href="#">id:</a> )	Find info at
<a href="#">related:</a>	List web pa

## Restrict Search to Sites where

Operators	Meaning
<a href="#">allinanchor:</a>	All query words
<a href="#">inanchor:</a>	Terms must app
<a href="#">allintext:</a>	All query words
<a href="#">intext:</a>	The terms must the text of the p
<a href="#">allintitle:</a>	All query words
<a href="#">intitle:</a>	The terms must
<a href="#">allinurl:</a>	All query words
<a href="#">inurl:</a>	The terms must

## Restrict Search to [Google Grc](#)

Operators	Meaning
<a href="#">author:</a>	Find Grou
<a href="#">group:</a>	Find Grou
<a href="#">insubject:</a>	Find Grou

## Restrict Search to [Google Nev](#)

Operators	Meaning
<a href="#">location:</a>	Find News
<a href="#">source:</a>	Find News

# Google Hacking Database

[www.johnny.ihackstuff.com](http://www.johnny.ihackstuff.com)

(edit: <http://johnny.ihackstuff.com/ghdb.php>)

Collection of queries for  
finding 'interesting' stuff

No longer updated

Possible results of GHD:

- Identify systems in use (including version)
- Identify known exploits
- Locations of sensitive information
- User-id's & passwords
- Logging files
- Many other things

[Advisories and Vulnerabilities](#) (215 entries)

These searches locate vulnerable servers. T

[Error Messages](#) (68 entries)

Really retarded error messages that say WA

[Files containing juicy info](#) (230 entries)

No usernames or passwords, but interesting

[Files containing passwords](#) (135 entries)

PASSWORDS, for the LOVE OF GOD!!! Googl

[Files containing usernames](#) (15 entries)

These files contain usernames, but no passw

[Footholds](#) (21 entries)

Examples of queries that can help a hacker

[Pages containing login portals](#) (232 entries)

These are login pages for various services. I

[Pages containing network or vulnerability da](#)

These pages contain such things as firewall

[Sensitive Directories](#) (61 entries)

Google's collection of web sites sharing sens

[Sensitive Online Shopping Info](#) (9 entries)

Examples of queries that can reveal online s



# The NEW and IMPROVED GHDB

## Welcome to the google hacking database

We call them 'googledorks': Inept or foolish people as revealed by Google. Whatever you call these fools, you've found the center of the Google Hacking Universe!

### Search Google Dorks

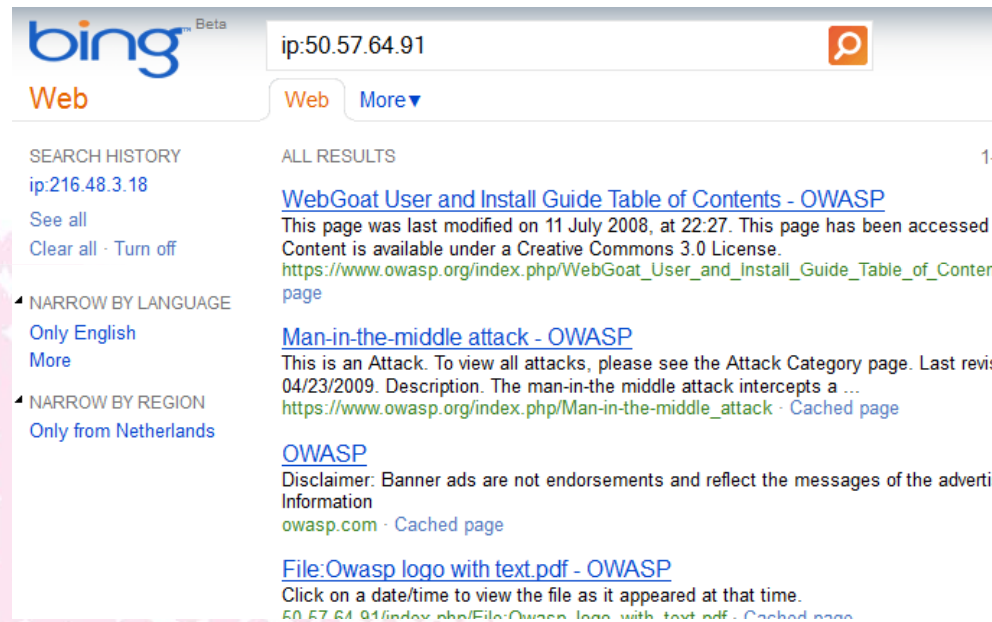
Category:  Free text search:

### Latest Google Hacking Entries

Date	Title	Category
2012-05-15	intitle:"HtmlAnvView:D7B039C1"	Various Online Devices
2012-05-15	intext:"~~Joomla1.txt" title:"Index..."	Files containing juicy info
2012-05-15	"Welcome to Sitecore" + "License Ho..."	Pages containing login portals
2012-05-15	intitle:"-N3t" filetype:php undetectable	Vulnerable Servers
2012-05-15	?intitle:index.of? ".mysql_history"	Files containing juicy info
2012-05-15	intitle:awen+intitle:asp.net	Vulnerable Servers
2012-05-15	"mailing list memberships reminder"	Pages containing login portals
2012-05-15	intext:"Thank you for your purchase/trial of ..."	Files containing juicy info
2012-05-15	inurl:"iki-index.php" filetype:php &quo...	Advisories and Vulnerabilities
2012-05-15	inurl:"*.php?*=*.php" intext:"Warni..."	Error Messages



## Finds subdomains with 'IP:x.x.x.x'



bing Beta

ip:50.57.64.91

Web More

SEARCH HISTORY

ip:216.48.3.18

See all

Clear all · Turn off

◀ NARROW BY LANGUAGE

Only English

More

◀ NARROW BY REGION

Only from Netherlands

ALL RESULTS 1-

[WebGoat User and Install Guide Table of Contents - OWASP](#)

This page was last modified on 11 July 2008, at 22:27. This page has been accessed : Content is available under a Creative Commons 3.0 License.  
[https://www.owasp.org/index.php/WebGoat\\_User\\_and\\_Install\\_Guide\\_Table\\_of\\_Content\\_page](https://www.owasp.org/index.php/WebGoat_User_and_Install_Guide_Table_of_Content_page)

[Man-in-the-middle attack - OWASP](#)

This is an Attack. To view all attacks, please see the Attack Category page. Last revis 04/23/2009. Description. The man-in-the middle attack intercepts a ...  
[https://www.owasp.org/index.php/Man-in-the-middle\\_attack](https://www.owasp.org/index.php/Man-in-the-middle_attack) · [Cached page](#)

[OWASP](#)

Disclaimer: Banner ads are not endorsements and reflect the messages of the advertis Information  
[owasp.com](#) · [Cached page](#)

[File:Owasp logo with text.pdf - OWASP](#)

Click on a date/time to view the file as it appeared at that time.  
[50.57.64.91/index.php/File:Owasp\\_logo\\_with\\_text.pdf](#) · [Cached page](#)

↑ Bing Query Language

Advanced Operator Reference

altloc:  
AND  
contains:  
define  
domain:  
ext:  
feed:  
filetype:  
hasfeed:  
imagesize:  
inanchor:  
inbody:  
instreamset:  
intitle:  
ip:  
keyword  
language:  
literalmeta:  
loc:  
location:  
meta:  
msite:  
near:  
NOT  
OR  
site:  
url:  
-  
&  
&&  
(  
)  
:  
[  
]  
|  
||  
"phrase"



[新闻](#) [网页](#) [贴吧](#) [知道](#) [MP3](#) [图片](#) [视频](#) [地图](#)

百度一下

[空间](#) [百科](#) [hao123](#) | [更多>>](#)

inurl:  
intitle:  
site:

**Google**

**Zoeken** Ongeveer 4.160 resultaten (0,23 seconden)

Alles  
Afbeeldingen  
Maps  
Video's  
Nieuws  
Shopping  
Meer

**Nijmegen**  
Locatie wijzigen

**Welcome To IIS 4.0!**  
[www.mckinleygrp.com/](http://www.mckinleygrp.com/) - Vertaal deze p  
29 Jun 2011 - Quality...Dependability...I  
commodities in today's market is over-t

**Welcome To IIS 4.0!**  
[www.mvidl.esu.k12.oh.us/](http://www.mvidl.esu.k12.oh.us/) - Vertaal dez  
Welcome to Microsoft® Windows NT®  
Option Pack provides enhanced Web, a

**welcome to iis 4.0**  
[www.keywordspy.com/.../keyw...](http://www.keywordspy.com/.../keyw...) - Vere  
Results 1 - 10 of 20 - 10. ctclix.com htt  
/business-and-economy-/1334-business

**bing** Beta

**Web**

**RELATED SEARCHES**  
IIS 4.0 Install  
IIS 4.0 Download XP  
Version 4.0 IIS  
Download IIS 4  
IIS 4.0 Free Download  
Windows XP IIS 4.0  
Welcome to IIS 4.0  
Internet Information  
Services IIS 4.0

**SEARCH HISTORY**  
Search more to see your  
history

**ALL RESULTS**

**Welcome To IIS 4.0!**  
Welcome to Microsoft® Windows NT® 4.0 Option Pack  
provides enhanced Web, application, and communicat  
[webdb.dmsc.moph.go.th](http://webdb.dmsc.moph.go.th) - Cached page

**Welcome To IIS 4.0!**  
Quality...Dependability...Reliability . The most efficient  
market is over-the-road motor carrier. McKinley Truckin  
[www.mckinleygrp.com](http://www.mckinleygrp.com) - Cached page

**Welcome To IIS 4.0!**  
**Internet Information Server 4.0**, the standards-base  
Server, brings unprecedented power to Web professor  
[206.222.19.242/iissamples/default/LEARN.asp](http://206.222.19.242/iissamples/default/LEARN.asp) - Cach

**Baidu** 百度 新闻 网页 贴吧 知道 MP3 图

去掉""获得更多 **intitle: Welcome to IIS 4.0** 的搜索!

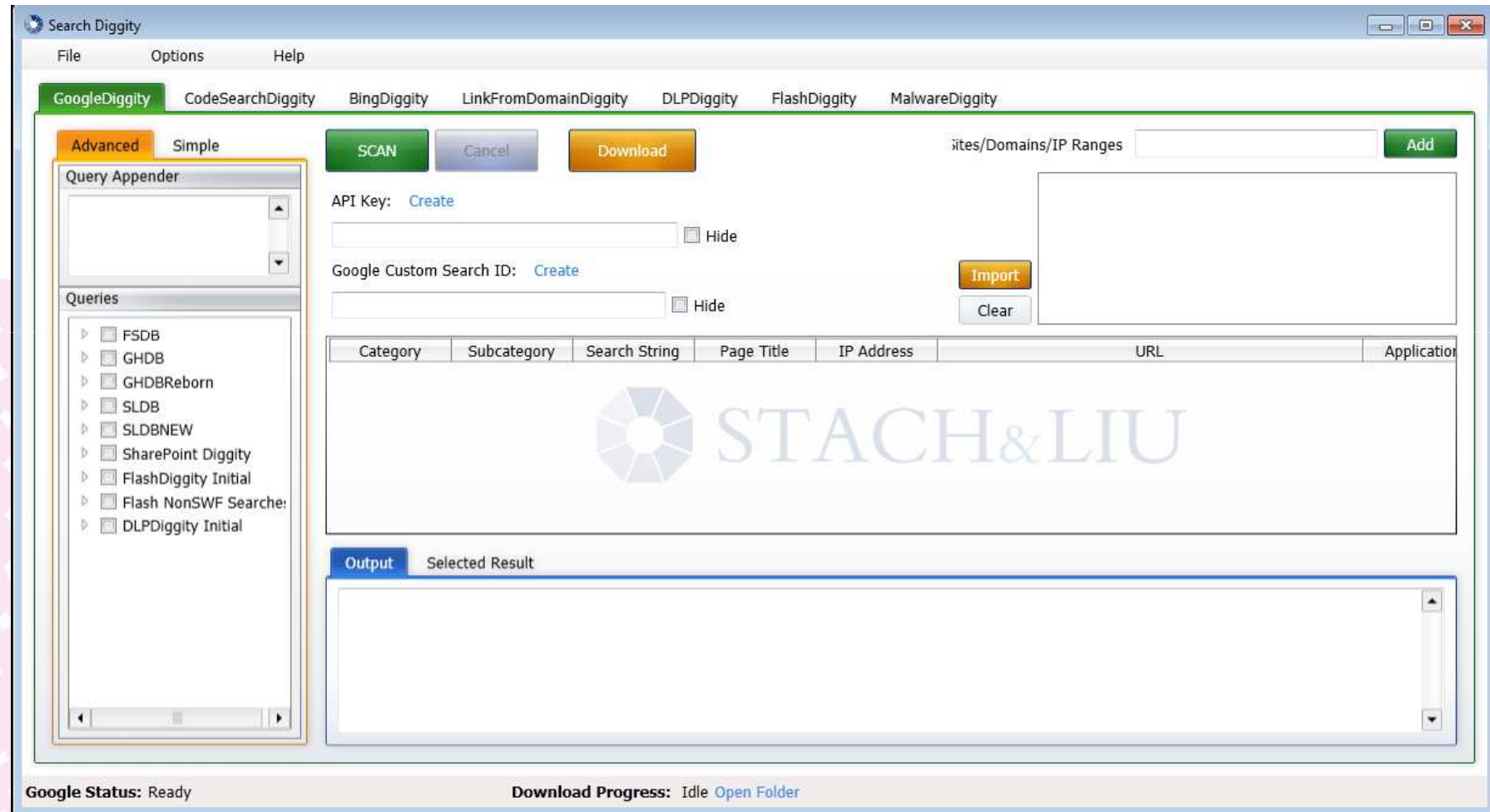
**Welcome To IIS 4.0!**

Welcome to Microsoft Windows NT 4.0 Option Pack Micro  
provides enhanced Web, application, and communication :  
[203.198.163.196/](http://203.198.163.196/) 2012-6-22 - 百度快照

**Welcome To IIS 4.0!**

Welcome to Microsoft Windows NT 4.0 Option Pack Micro  
provides enhanced Web, application, and communication :  
[worldcruising.net/](http://worldcruising.net/) 2012-6-15 - 百度快照

**Welcome To IIS 4.0!**



GoogleDiggity

BingDiggity

FlashDiggity

DLPDiggity

MalwareDiggity

Google CodeSearchDiggity

Bing LinkFromDomainDiggity

Bing Hacking Database (BHDB)

Sharepoint GoogleDiggity Dictionary File

Stach & Liu Database (SLDB)

GHDB Reborn Dictionaries – Exploit-DB.com

## Alert RSS Feeds

Diggity Alerts FUNdle Bundle

Google Hacking Alerts

Bing Hacking Alerts

SharePoint Hacking Alerts

SHODAN Hacking Alerts

## Alert RSS Monitoring Tools

AlertDiggity

i-DiggityAlerts

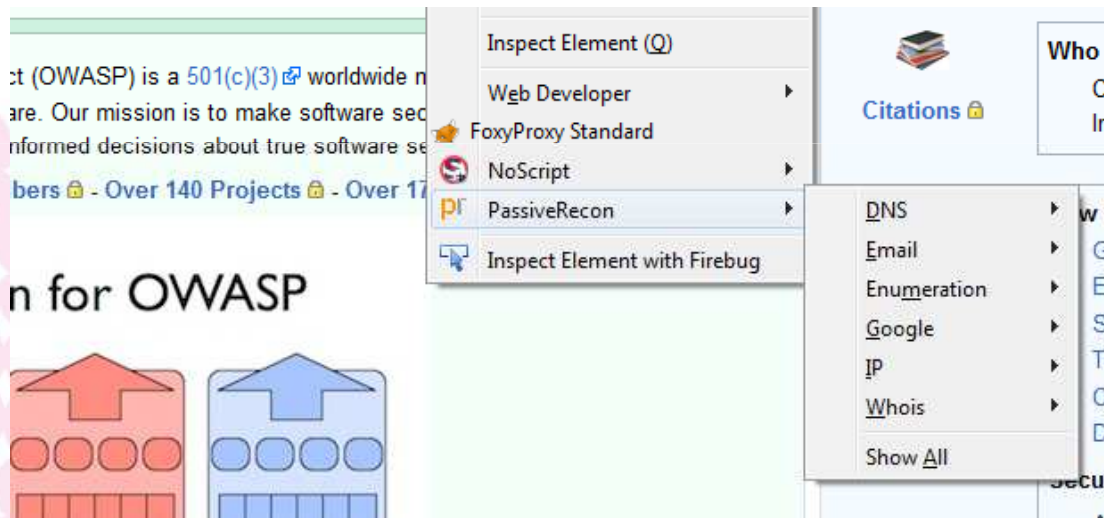
DroidDiggityAlerts

MalwareDiggityAlerts



# Domain Scanning 'on-the-fly'

Passive Recon (Firefox add-on)





OWASP - FOCA Free 3.0

Project Tools Options TaskList About Donate

OWASP

- Network
- Domains
- Roles
- Vulnerabilities**
- Metadata
  - Documents (123/1235)
    - .doc (123)
    - Metadata Summary
    - Users (42)
    - Folders (664)
    - Printers (0)
    - Software (6)
    - Emails (1)
    - Operating Systems (5)
    - Passwords (0)
    - Servers (0)

**Search engines**

☒ Google  
☒ Bing  
☒ Exalead

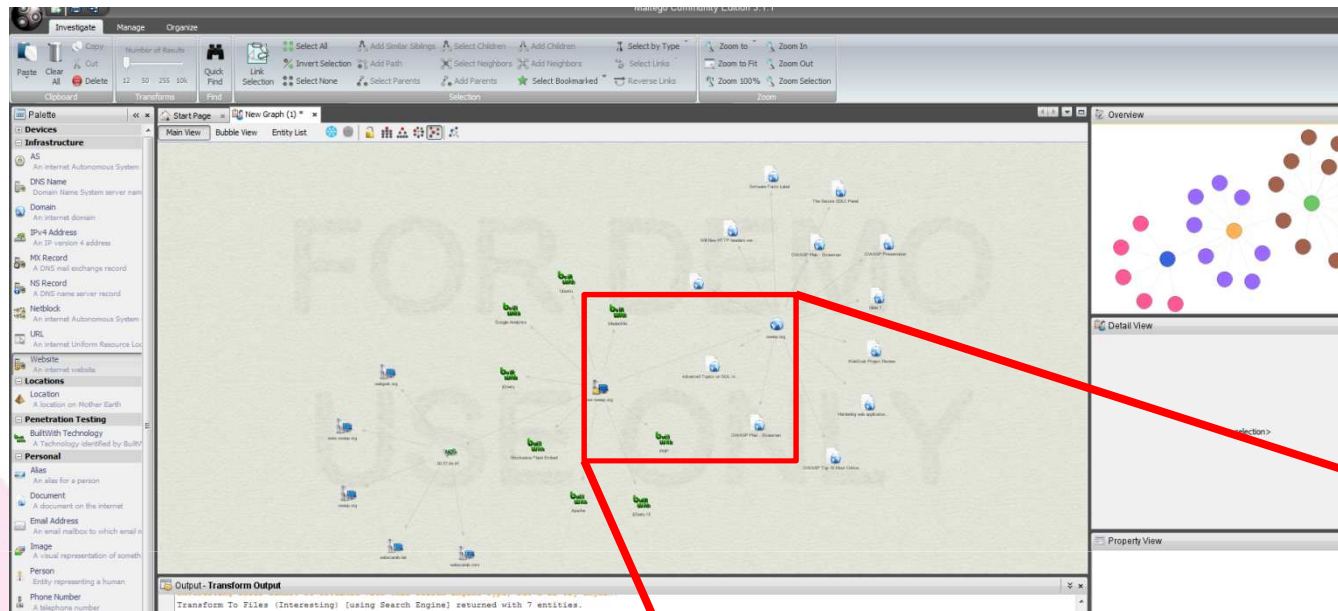
**Extensions**

☒ doc ☒ xls ☒ ppsx ☒ sxc  
☒ ppt ☒ docx ☒xlsx ☒ sxi  
☒ pps ☒ pptx ☒ sxw ☒ odt

[Custom search](#) Search All

Id	Type	URL	Download	Download Date	Size	Ans
100	doc	https://www.owasp.org/images/b/bd/%25E5%259F%25...	●	23-6-2012 11:39:52	0 bytes	✗
101	doc	https://www.owasp.org/images/a/aa/Legal_One_Page...	●	23-6-2012 11:39:53	139 KB	●
102	doc	https://www.owasp.org/images/archive/a/aa/2009012...	●	23-6-2012 11:39:54	123 KB	●
103	doc	https://www.owasp.org/images/8/81/OWASP_Code_R...	●	23-6-2012 11:39:57	1,57 MB	●
104	doc	https://www.owasp.org/images/6/60/ASVS_One_Page...	●	23-6-2012 11:39:56	347 KB	●
105	doc	https://www.owasp.org/images/archive/a/ac/2010111...	●	23-6-2012 11:39:58	272,5 KB	●
106	doc	https://www.owasp.org/images/4/4c/JavaEE-ESAPI_2...	●	23-6-2012 11:39:59	303,5 KB	●
107	doc	https://www.owasp.org/images/archive/a/a0/2008052...	●	23-6-2012 11:40:00	120,5 KB	●
108	doc	https://www.owasp.org/images/archive/a/a8/2008050...	●	23-6-2012 11:40:02	541 KB	●
109	doc	https://www.owasp.org/images/archive/a/aa/2009032...	●	23-6-2012 11:40:02	128 KB	●
110	doc	https://www.owasp.org/index.php/File:OotM-PaymentP...	●	23-6-2012 11:40:04	20,33 KB	●
111	doc	https://www.owasp.org/index.php/File:%25F5%259F%2...	●	23-6-2012 11:40:04	15,96 KB	●





Intelligence and forensics tool

Connects many different sources of info

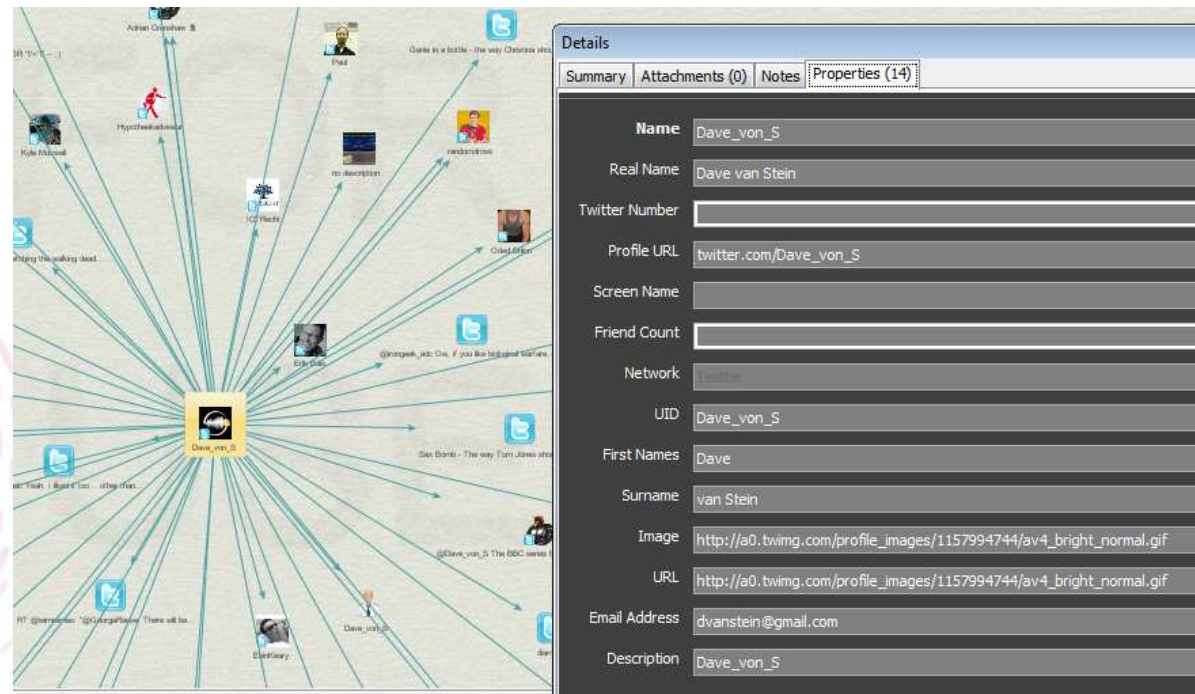
Represents in graphical way

Very extensive capabilities



Can also be used for social engineering

- Facebook & twitter
- Email addresses
- Phone numbers
- etc



# theHarvester

## The sources supported are:

Google - emails,subdomains/hostnames  
 Google profiles - Employee names  
 Bing search - emails, subdomains/hostnames,virtual hosts  
 Pgp servers - emails, subdomains/hostnames  
 Linkedin - Employee names  
 Exalead - emails,subdomain/hostnames

```
*****
*TheHarvester Ver. 2.2          *
*Coded by Christian Martorella  *
*Edge-Security Research        *
*cmartorella@edge-security.com  *
*****

Usage: theharvester options

-d: Domain to search or company name
-b: Data source (google,bing,bingapi,pgp,linkedin,google-p
23,jigsaw,all)
-s: Start in result number X (default 0)
-v: Verify host name via dns resolution and search for vir
-f: Save the results into an HTML and XML file
-n: Perform a DNS reverse query on all ranges discovered
-c: Perform a DNS brute force for the domain name
-t: Perform a DNS TLD expansion discovery
-e: Use this DNS server
-l: Limit the number of results to work with(bing goes fro
lts,
    -h: use SHODAN database to query discovered hosts
    google 100 to 100, and pgp doesn't use this option)

Examples:./theharvester.py -d microsoft.com -l 500 -b google
./theharvester.py -d microsoft.com -b pgp
./theharvester.py -d microsoft -l 200 -b linkedin
```

```
jim.manico@owasp.org
dirk.wetter@owasp.org
daniel.cuthbert@owasp.org
mark.roxberry@owasp.org
bradcausey@owasp.org
christian.edjenguele@owasp.org
dcampbell@owasp.org
edward@owasp.org
sebastien.gioria@owasp.org
delhi@lists.owasp.org
requ...@lists.owasp.org
houn...@lists.owasp.org
```

### [+] Hosts found in search engines:

```
50.57.64.91:www.owasp.org
50.56.58.227:lists.owasp.org
74.125.79.121:sl.owasp.org
50.56.58.227:Lists.owasp.org
67.215.65.132:40lists.owasp.org
67.215.65.132:www2.owasp.org
50.57.64.91:www.owasp.org
50.56.58.227:lists.owasp.org
[+] Virtual hosts:
=====
```

# Conclusions

What search engines see, hackers can abuse

Anonymous, online and offline, Highly automated

Many tools are freely available

Networks can be mapped with much detail in minutes

Much information about your company, systems and users  
available on internet

# Remedies (1/2)

## Limit access

- Allow search engines only to see what they need to see.
- Make sure unauthorized users are not able to look into or even see files they do not need to see.
- Force possible intruders to use methods that can be scanned and monitored.

## Use the tools of hackers

- Scan your systems with the tools hackers use and check the information that is found.
- Scan for error messages and other things that reveal information about the system and services and remove them.

## Check what spiders can see

- Use a spider simulator to check what spiders can see and if your application still functions correctly.



# Remedies (2/2)

## Awareness

- Be aware of all possible sources of information. Create awareness among employees. Assume all information will possibly abused

## Clean documents

- Remove al metadata from documents before publishing.

## Audit frequently

- Keep your knowledge up-to-date and scan regularly for information that can be found about your systems or hire professionals do to it for you.

# Interesting books on the subject

