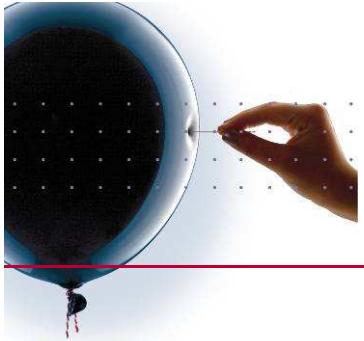


# **Modern information gathering**

Dave van Stein  
9 april 2009

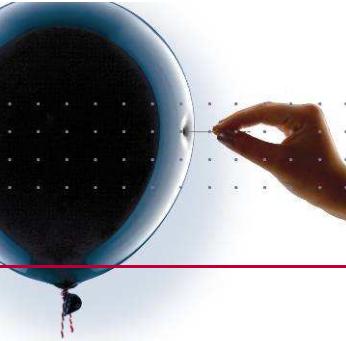


# Who Am I

---

- ✓ Dave van Stein
- ✓ 34 years
- ✓ Functional tester > 7 years
- ✓ Specializing in (Application) Security Testing
- ✓ "Certified Ethical Hacker"

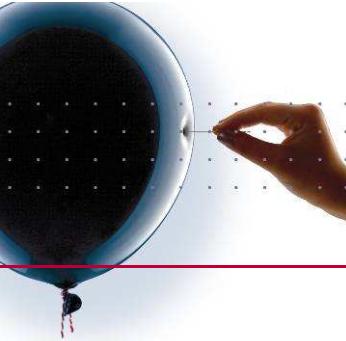




# Agenda

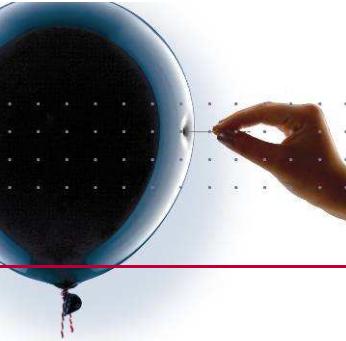
---

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



## Goal of this presentation

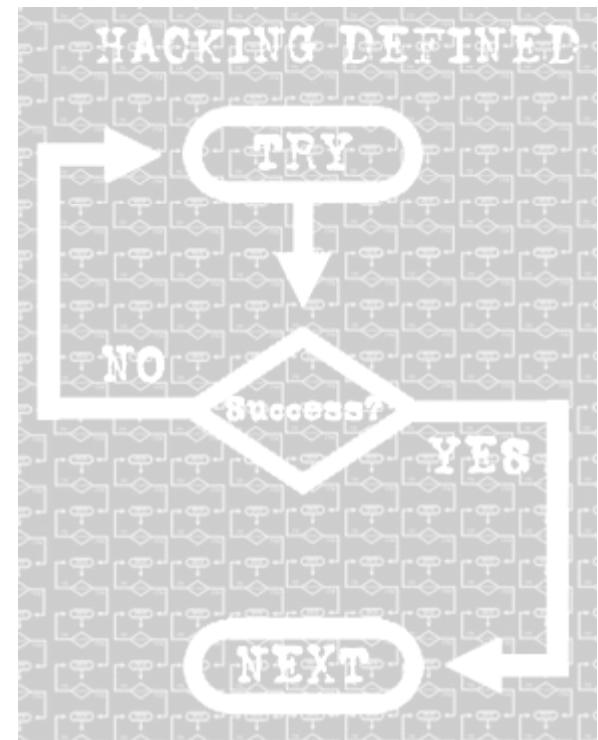
- ✓ Give insight in amount of information anonymously available on internet about your system (and users)
  
- ✓ Give insight in the amount and possibilities of tools freely available

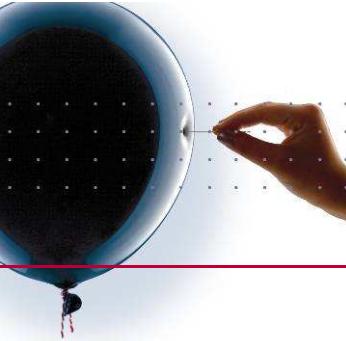


## Intermezzo: How to hack

- ✓ Identify entrypoint
- ✓ Gain access
- ✓ Secure access
- ✓ Do stuff
- ✓ Clear up the mess
- ✓ Come back another time

(simplified procedure)

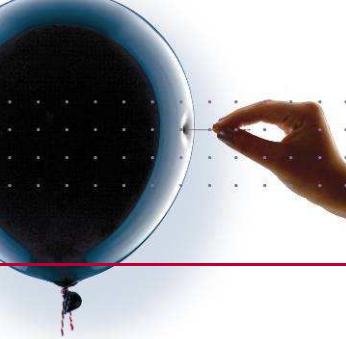




# Information Gathering

---

- ✓ Information gathering scans for:
  - Domains and subdomains
  - IP addresses
  - Applications and technologies
  - Hotspots (known vulnerabilities)
  - Usernames and passwords
  - Sensitive information
  
- ✓ Not only identifying risks, but also risk on exposure and exploiting



# Passive Reconnaissance

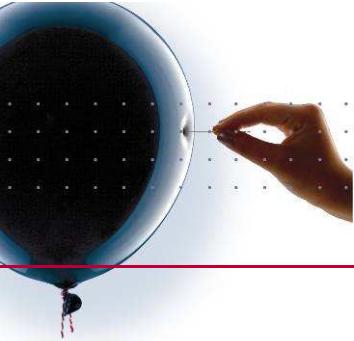
---

## ✓ Reconnaissance:

- Information gathering, fingerprinting
- Gaining information about a target

## ✓ Passive

- Without making contact with target
- No direct scanning, no intrusion
- No logging and no alarm triggering !



## Sources of information

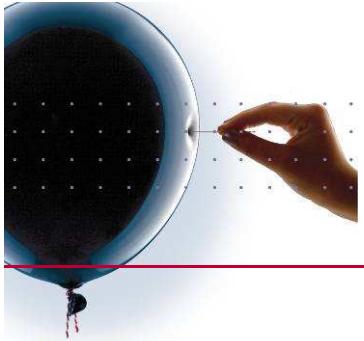
---

### ✓ Public records

- WHOIS: information about owner
- DNS : information about IP addresses
- Necessary for network functionality

### ✓ Search engines

- Often little restrictions on websites
- Cache all information gathered



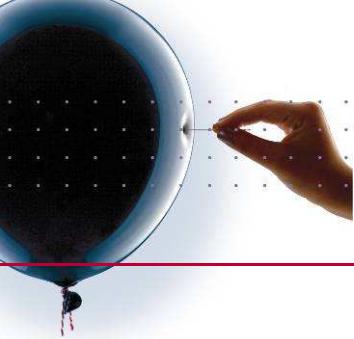
## Tools

### ✓ What do you need ?

- Webbrowser
- Internet access
- Creativity

### ✓ Advanced and Automated scanning:

- Specialized (offline) Tools



# '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:

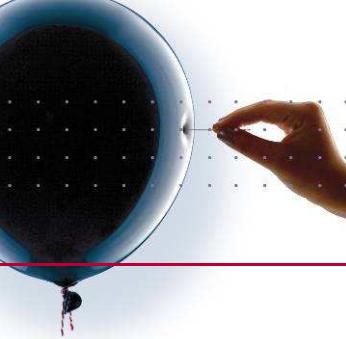
- DNS zone transfers often not authorized
- Active connection with target => detectable

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

C:\>tracert www.google.com

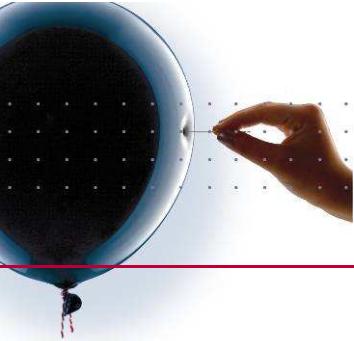
Tracing route to www.l.google.com [66.249.91.103]
over a maximum of 30 hops:
  1  2 ms   1 ms   1 ms  192.168.10.1
  2  58 ms   30 ms   31 ms  ip1-144-173-82.adsl2.static.versateli.nl [82.173.
144.1]
  3  73 ms   30 ms   31 ms  ge-0-1-0-1305.ncr0iasd2.versateli.net [217.16.44.
49]
  4  57 ms   32 ms   39 ms  ge-1-3-0-666.br0isara.versateli.net [212.53.18.18]
  5  69 ms   25 ms   54 ms  core1.ams.net.google.com [195.69.144.247]
  6  32 ms   22 ms   21 ms  69.85.248.93
  7  30 ms   45 ms   73 ms  4.233.175.246
  8  49 ms   135 ms   83 ms  66.249.94.146
  9  35 ms   35 ms   85 ms  ik-in-f103.google.com [66.249.91.103]

C:\>
```



## 'Modern' Domain Scanning

- ✓ Various websites
  - Anonymous
  - Combination of techniques
  - Sort results for nice presentation
  
- ✓ Search engine 'tweaking'
  - Additional information linked to domain
  
- Some examples



# Domain Scanning: ServerSniff

Reports

IP-Tools

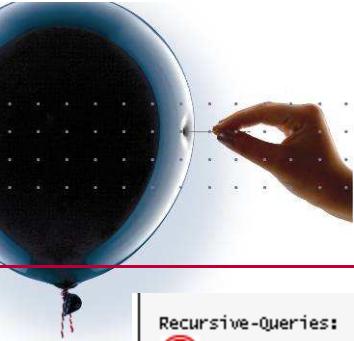
Nameserver

Webserver

## ✓ Server Sniff

- NS reports
- Domain reports
- Subdomains
- Various (trace)routes
- Various ping types
- Shows robots.txt
- Anonymous !





# Domain Scanning: Server Sniff

## Recursive-Queries:

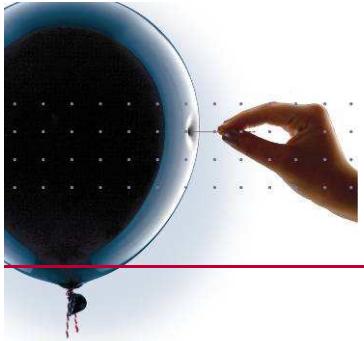
- ! ns1.secure.net. YES - recursive queries allowed!
- ! ns2.secure.net. YES - recursive queries allowed!

## 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. 2007060332 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
google6912a08c3a8cdff0b.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.rdrurkee.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. 2007060332 86400 7200 2592000 86400
```

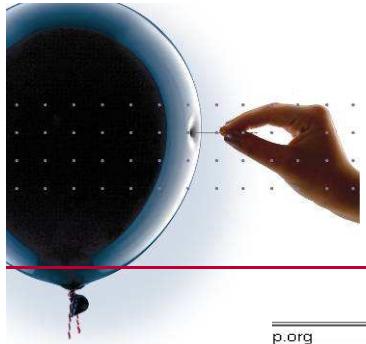
owasp.org. 86400 IN MX 2  
\*.owasp.org. 86400 IN CNAME  
austin.owasp.org. 86400 IN  
blogs.owasp.org. 86400 IN  
calendar.owasp.org. 86400 IN  
docs.owasp.org. 86400 IN  
es.owasp.org. 86400 IN A  
google6912a08c3a8cdff0b.ov  
jobs.owasp.org. 86400 IN  
lists.owasp.org. 86400 IN CNAME  
mail.owasp.org. 86400 IN CNAME  
old.owasp.org. 86400 IN A  
registration.owasp.org. 86400 IN CNAME  
stage.owasp.org. 86400 IN CNAME  
voip.owasp.org. 86400 IN A  
webmail.owasp.org. 86400 IN A  
www.owasp.org. 86400 IN CNAME  
owasp.org. 86400 IN SOA ns1.secure.net.  
hostmaster.secure.net. 2007060332 86400 7200 2592000 86400



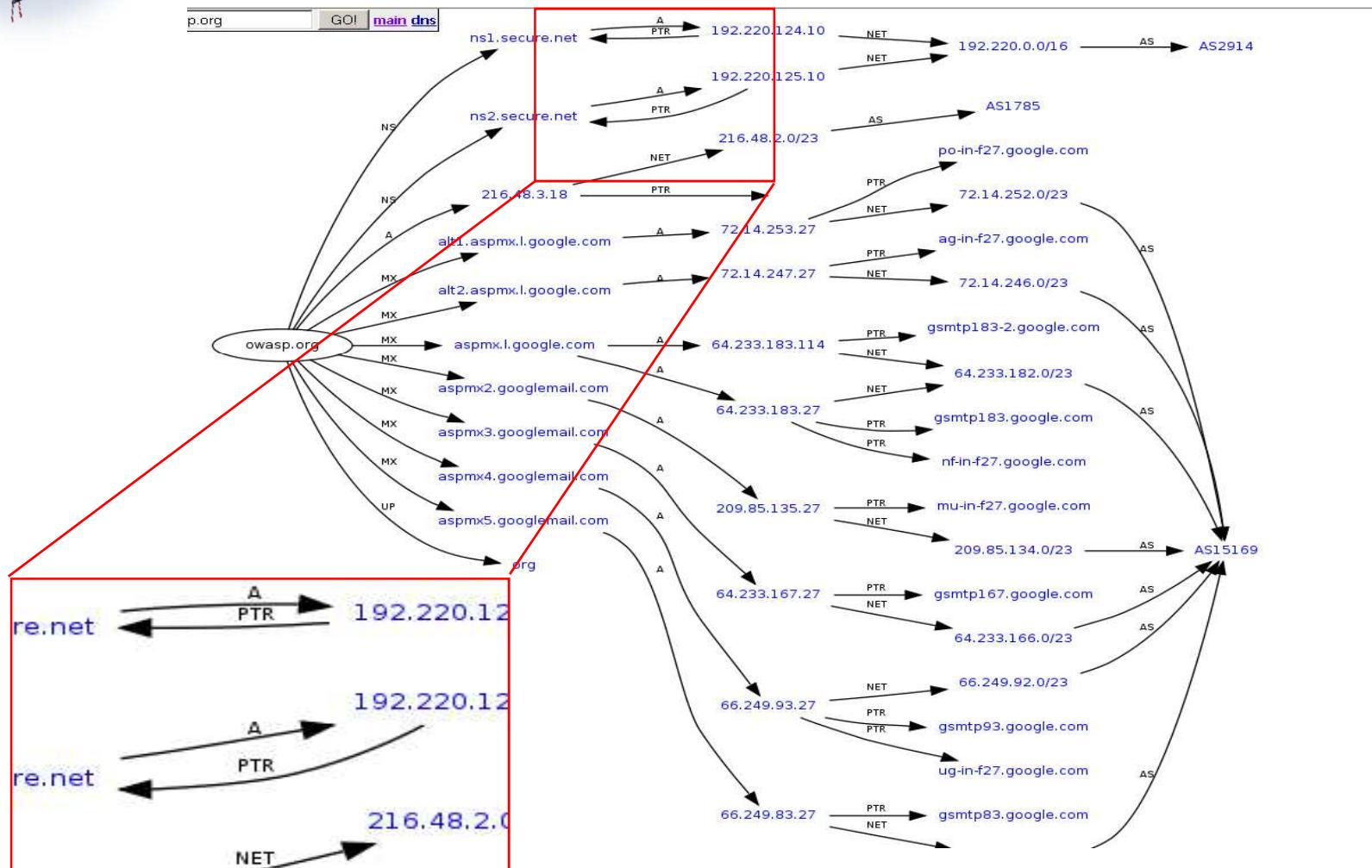
# Domain Scanning: Robtex

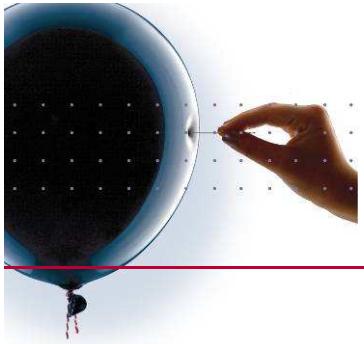
- ✓ Domain 'Swiss Army Knife'
  - Provides ALL information linked to a domain

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>
		<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 NTTC GIN AS N</a>
	mx	<a href="#">ns2.secure.net</a>	<a href="#">192.220.125.10</a>	-		
		<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="#">alt2.aspmx.l.google.com</a>	<a href="#">72.14.247.27</a>	<a href="#">aq-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="#">gsmt183.google.com</a>	<a href="#">64.233.182.0/23</a> Google	
			<a href="#">64.233.183.114</a>	<a href="#">nf-in-f27.google.com</a>		
		<a href="#">aspmx2.googlemail.com</a>	<a href="#">209.85.135.27</a>	<a href="#">gsmt183-2.google.com</a>	<a href="#">209.85.134.0/23</a> Google	<a href="#">AS15169 Google , Inc</a>
		<a href="#">aspmx3.googlemail.com</a>	<a href="#">64.233.167.27</a>	<a href="#">mu-in-f27.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="#">gsmt93.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="#">uq-in-f27.google.com</a>	<a href="#">66.249.82.0/23</a> Google	
org	ns	ptr	<a href="#">66.35.111.73</a>	-	<a href="#">66.35.111.0/24</a>	<a href="#">AS14955 N-V-C Northe</a>
		<a href="#">d0.org.afiliias-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 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="#">tld2.ultradns.net</a>	<a href="#">204.74.113.1</a>	-	<a href="#">204.74.113.0/24</a> UltraDNS	<a href="#">AS12008 UNSPECIFIED</a>
		<a href="#">a0.org.afiliias-nst.info</a>	<a href="#">199.19.56.1</a>	-	<a href="#">199.19.56.0/24</a> REACH (Customer Route)	
					<a href="#">199.19.54.0/24</a> REACH (Customer	<a href="#">AS12041 AFILIAS NST</a>



# Domain scanning: Robtex



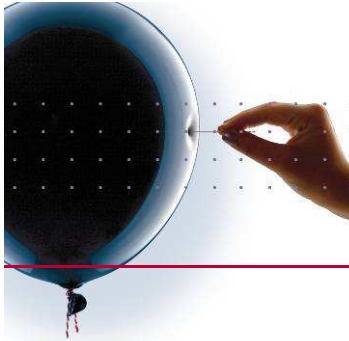


# Domain Scanning 'on-the-fly'

## ✓ Passive Recon (Firefox add-on)

The screenshot shows a portion of the Firefox context menu. On the left, under the 'More Tools' section, the 'PassiveRecon' option is highlighted with a dark blue background. To its right is a list of various passive reconnaissance tools and services:

- DNSReport.com - Domain Information
- Robtex.com (Root Domain) - DNS Information
- Robtex.com (WWW) - DNS Information
- DomainTools.com - Whois
- Sam Spade - Whois
- EmailStuff.org - MX Records
- Hashemian - Domain Mail Server/Exchanger (MX Records) Lookup
- Netcraft - Site Report
- Netcraft - What's this Site Running Report
- Network-Tools.com - Network Lookup
- Network-Tools.com - Traceroute



# Domain Scanning: Live search

✓ Finds subdomains with '*IP:x.x.x.x*'

Live Search | MSN | Windows Live

 **Live Search** ip:216.48.3.18 

Alleen [Nederlands](#)  Alleen websites in Nederland

---

**Web** 1-10 van 263.000 resultaten · Geavanceerd  
Zie ook: [Afbeeldingen](#), [Nieuws](#), [Alles weergeven...](#)

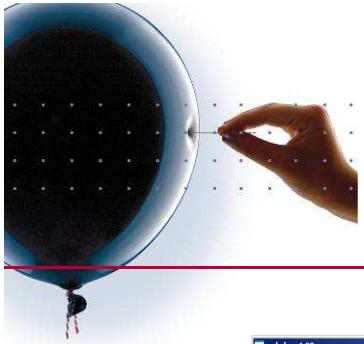
[ip](#) - www.Marktplaats.nl/telecommunicatie Ges  
Duizenden Mobiele Telefoons en Accessoires op Marktplaats .

[The Open Web Application Security Project](#)  
How to build, design and test the security of web applications and web services.  
[www.owasp.org](#) · [Pagina in cache](#) · [Vertaal deze pagina](#)

[www.owasp.org](#)  
[www.owasp.org/local/boston.html](#) · [Pagina in cache](#) · [Vertaal deze pagina](#)  
[Meer resultaten weergeven voor www.owasp.org](#)

[OWASP Blogs](#)  
5 Most Active Blogs  
[blogs.owasp.org](#) · [Pagina in cache](#) · [Vertaal deze pagina](#)

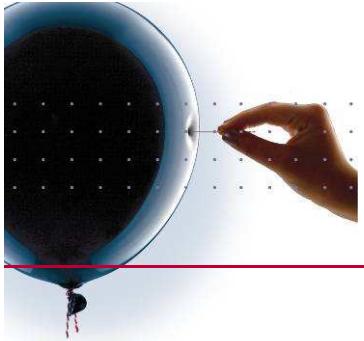
 [Dinic Cruz Blog](#)



# Live search automated: Webshag

The screenshot shows the Webshag 1.00 application window. In the top-left corner, there's a red circle highlighting the 'target [host | IPv4]' field which contains 'www.owasp.org'. A red arrow points from this field to a list of domains in the 'Results' section. Another red circle highlights the 'Console' output area at the bottom, which displays the message: 'INFO Domains of 216.48.3.18 retrieved' and 'INFO Found 13 domains.' Below the console is a 'Copy:' field with a large empty text area.

This screenshot shows a web-based interface for a live search. At the top, there is a 'Target [host | IPv4:]' input field containing 'www.owasp.org'. Below it, under the heading 'Results', is a list of 'Domains:' which matches the list shown in the Webshag interface. The domains listed are: blogs.owasp.org, www.owasp.org, austin.owasp.org, ww.owasp.org, forums.owasp.org, owasp.net, beta.owasp.org, forum.owasp.org, b.owasp.org, www.owasp.fr, WEBSCARAB.NET, WEBSCARAB.COM, and webscarab.org.

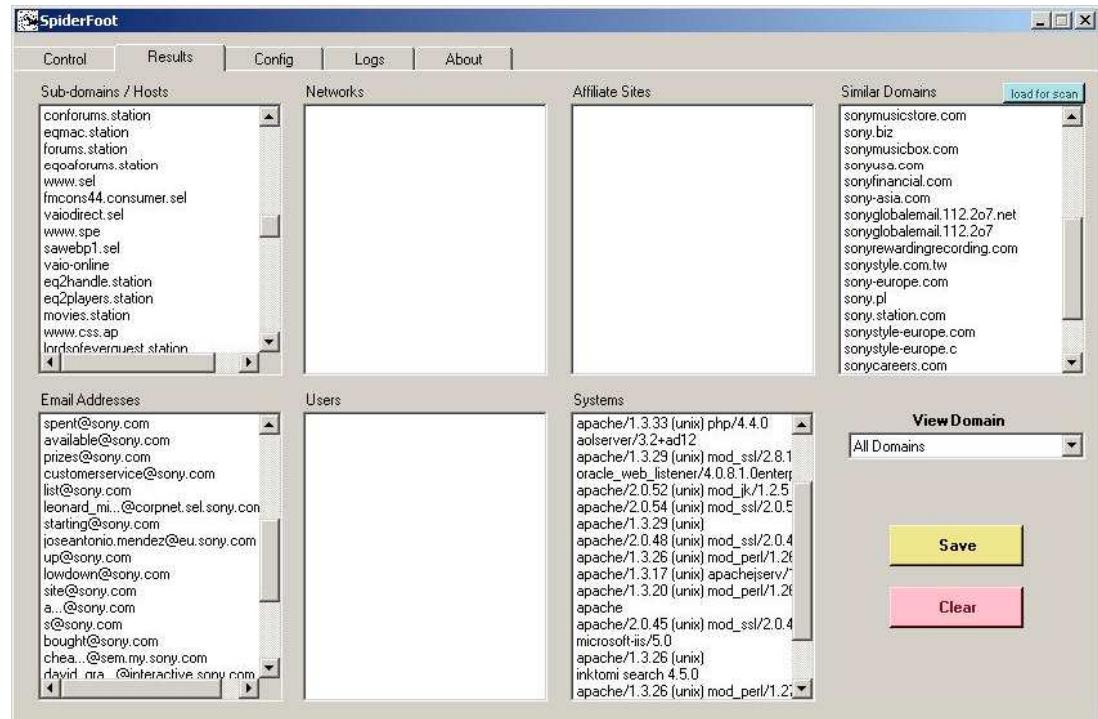


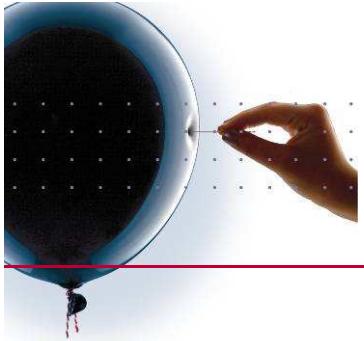
## Other tools

### ✓ Spiderfoot / Wikto

- Combine DNS / Google / Live Search / Yahoo

- Subdomains
- Directories
- IP's
- Email addresses
- Usernames
- Systems in use

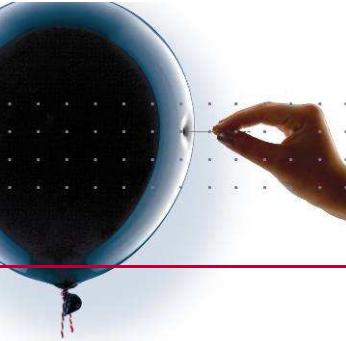




# Maltego

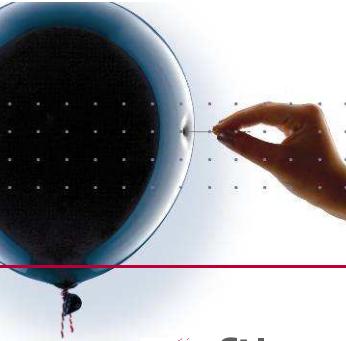
---

- ✓ Intelligence and forensics tool
- ✓ Connects many different sources of info
- ✓ Represents in graphical way
- ✓ Very extensive capabilities
  
- ✓ Too much to cover in this presentation
- ✓ <http://www.paterva.com/maltego>



## Modern Domain Scanning

- ✓ Anonymous
- ✓ Both online and offline
- ✓ Highly automated
- ✓ Graphical network mapping in less than 10 minutes !
- ✓ Lots of additional information



# Google Advanced search

## ✓ filetype: (or ext:)

- Find documents of the specified type.  
*E.g. PDF, XLS, DOC*

## ✓ intext:

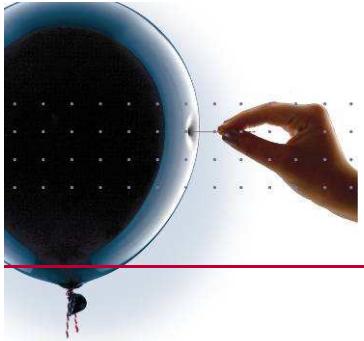
- The terms must appear in the text of the page.

## ✓ intitle:

- The terms must appear in the title of the page.

## ✓ inurl:

- The terms must appear in the URL of the page.



# 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

- ✓ Regular updates

[Advisories and Vulnerabilities](#) (215 entries)  
These searches locate vulnerable servers. 1

[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 pass

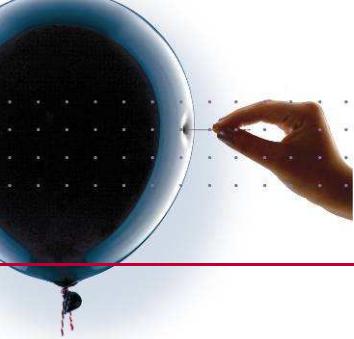
[Footholds](#) (21 entries)  
Examples of queries that can help a hacker

[Pages containing login portals](#) (232 entries)  
These are login pages for various services. !

[Pages containing network or vulnerability de](#)  
These pages contain such things as firewall

[Sensitive Directories](#) (61 entries)  
Google's collection of web sites sharing sensi

[Sensitive Online Shopping Info](#) (9 entries)  
Examples of queries that can reveal online s

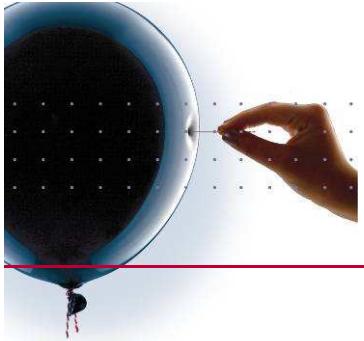


## GHD applications

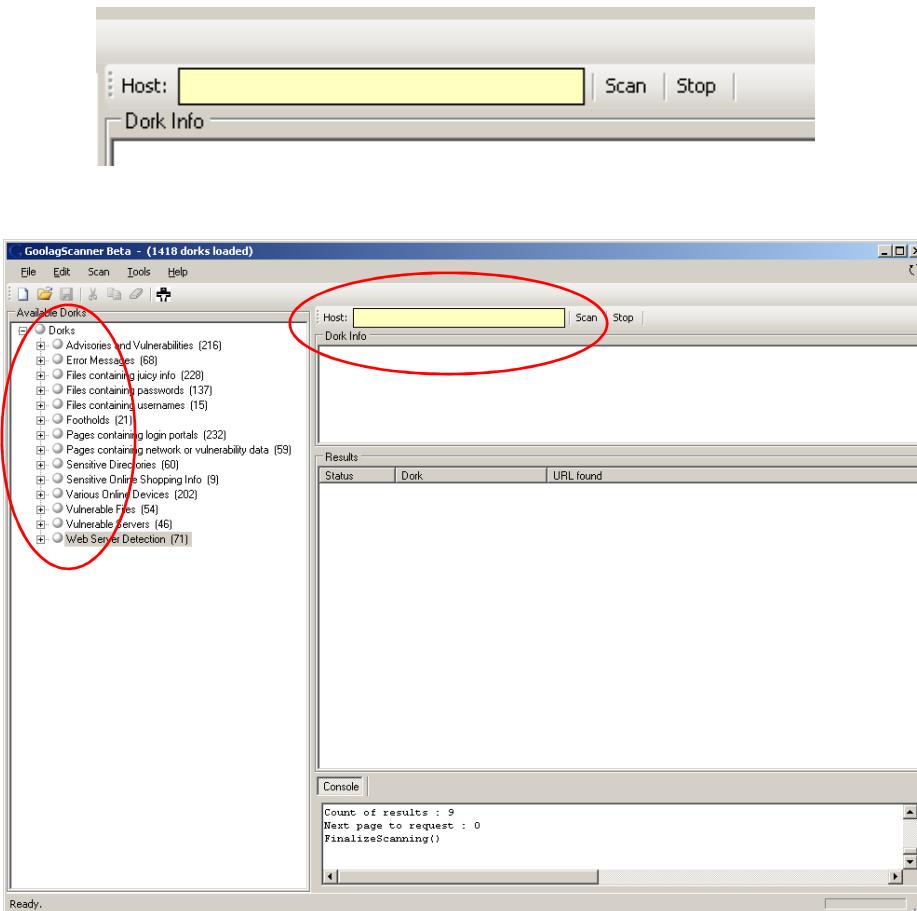
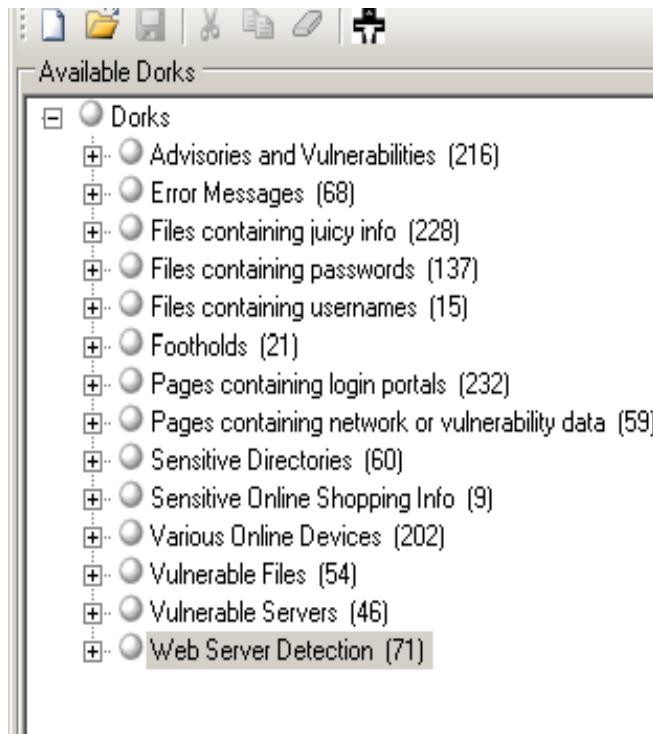
### ✓ Goolag scanner

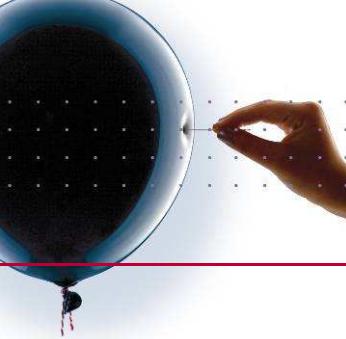
*Goolag Scanner is a Web auditing tool. It works by exploiting data- retention practices of popular search engines.*

- Contains Google Hacking Database
- Automated Google queries
- Automated result interpretation
- Single host or general scan



# Goolag scanner

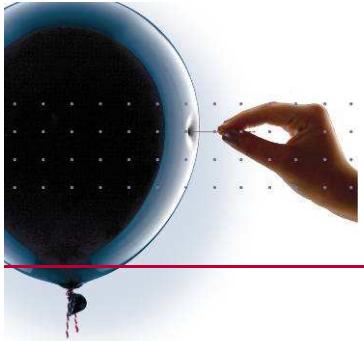




## More applications

---

- ✓ Modern vulnerability scanners use GHD:
  - IBM Rational Appscan
  - Acunetix Vulnerability Scanner
  - Others
  
- ✓ Several Firefox plug-ins for “on-the-fly” scanning

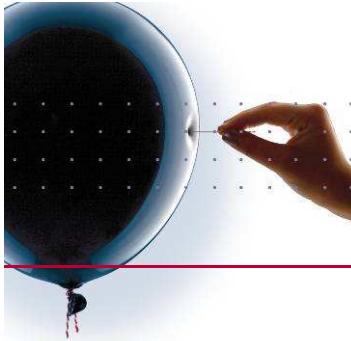


# Google Hacking Database

---

## ✓ 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



# Yahoo search: file explorer

## ✓ File explorer for the web

**WebLS: The file explorer for the Web. - Mozilla Firefox**

File Edit View History Bookmarks Tools Help  
http://marcoslot.net/apps/webls/ Google

General News Blogs & Fora Testing Exploit lists IP tools Proxy & Mail Encrypt Tools Documentation Cheat Sheets

**WebLS: The file explorer for the Web.**

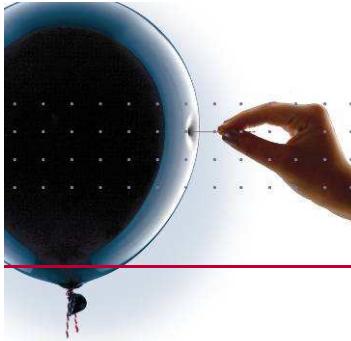
WebLS provides an alternative way of browsing the web. It uses the [Yahoo Search Engine API](#) to build a file tree for a specified URL and lets you browse through it. WebLS has many uses, but it was born from curiosity. You can now see what is behind a website.

**Note:** Some time after the release of WebLS Yahoo! recognized the merit of a tool like WebLS and released their [Site Explorer](#) service. For our purpose Site Explorer is much faster and more complete than regular search service we currently use. I will try to find the time to update WebLS to use this new great new service.

Go Refresh

Name	Size	Type	Last modified

Instructions About  
Cooked by [Marco Slot](#)  
Advertisement Ads by Google ▲▼  
▲

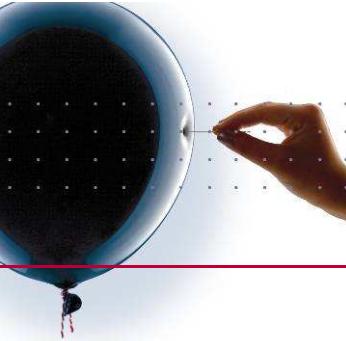


# Yahoo search: file explorer

## ✓ Examples

Name	Size	Type	Last modified
<a href="#">index.php</a>	0kB 17kB	text/html	Thu, 10 Jul 2008 07:00:00 GMT

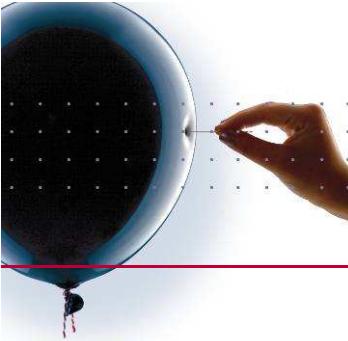
Name	Size	Type	Last modified
<a href="#">index.php</a>	0kB		
<a href="#">be</a>			
<a href="#">data</a>			
<a href="#">news</a>			
<a href="#">references</a>			
<a href="#">styles</a>			
<a href="#">function.mysql-list-tables</a>	0kB	text/html	Wed, 18 Jun 2008 07:00:00 GMT
<a href="#">index.php</a>	8kB	text/html	Wed, 16 Jul 2008 07:00:00 GMT



## Other tools

---

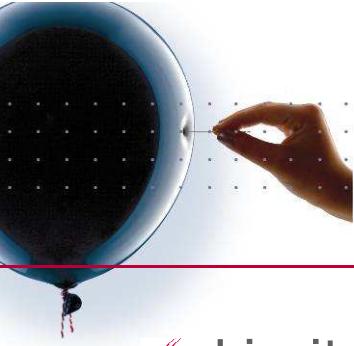
- ✓ Metagoofil : extract metadata from documents on website
  - User names, server names, path locations, software + versions, MAC addresses (!)
  
- ✓ Wikiscanner : check comments made on Wikipedia by company or domain
  - Company IP ranges
  
- ✓ Several “Social Site” extractors
  - Linkedin, twitter, hyves, etc, etc, etc



## Conclusions

---

- ✓ What search engines see, hackers can abuse
- ✓ 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

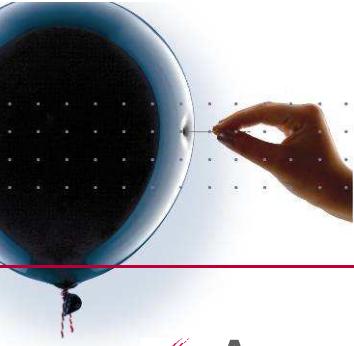
- Allow search engines only to see what they need to see. Make sure unauthorized users are not able to look into or even see files they do not need to see. Force possible intruders to use methods that can be scanned and monitored.

### ✓ 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

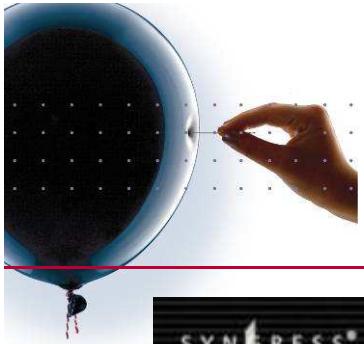
- Be aware of all possible sources of information. Create awareness among employees. Assume all information will possibly be abused.

### ✓ Clean documents

- Remove all 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 to do it for you.



## Interesting books on the subject

