

You found that on Google?

Gaining awareness about “Google Hackers”

Johnny Long

johnny@ihackstuff.com

HUGE DISCLAIMER!

- The print/CD version of this presentation is much smaller than the live version!
- The live version shows *many more* techniques and examples. After all, I can't leave a paper trail... =^P
- **DEFCON ATTENDEES:** This print version is the same as the Blackhat talk, but the live version is very different! I'm too lazy to make 2 print versions ;-)

What this is about

- We'll be talking about how hackers can use Google to locate vulnerable targets and sensitive information
- This process has been termed "Google hacking"
- We will be blowing through the basics
 - After all, this is DEFCON! =)

Advanced Operators

- Google advanced operators help refine searches
- Advanced operators use a syntax such as the following:
 - operator:search_term
- Notice that there's no space between the operator, the colon, and the search term

Advanced Operators

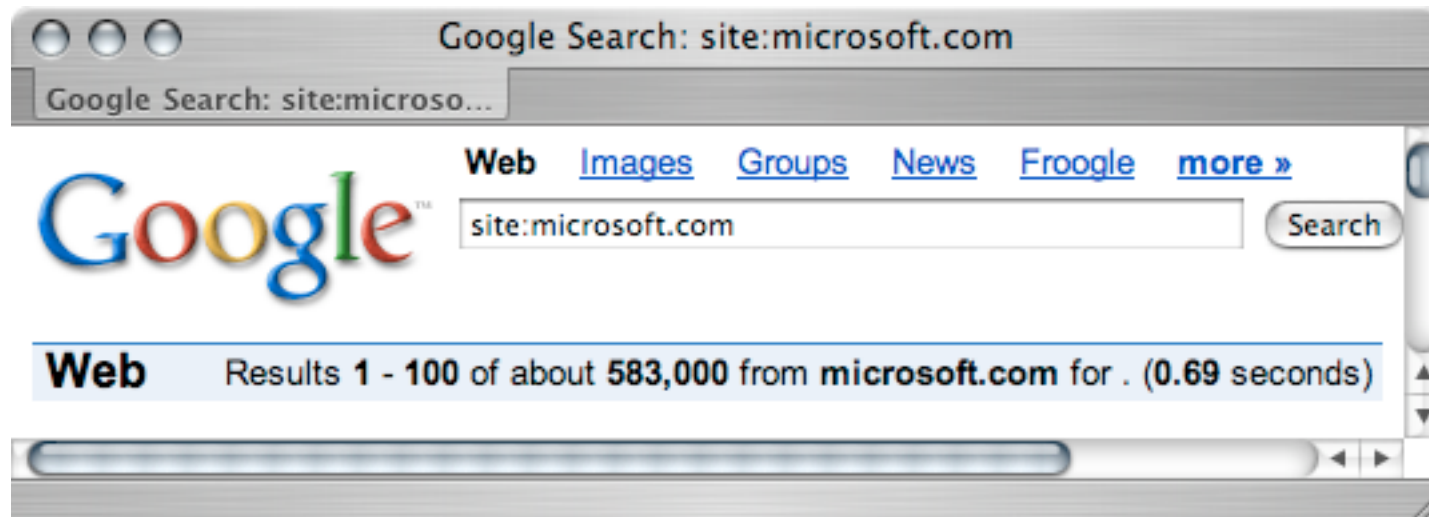
- site: restrict a search to a specific web site or domain
 - The web site to search must be supplied after the colon.
- filetype: search only within the text of a particular type of file
- link: search within hyperlinks
- cache: displays the version of a web page as it appeared when Google crawled the site
- intitle: search for a term in the title of a document
- inurl: search only within the URL (web address) of a document

Search Characters

- Some characters:
 - (+) force inclusion of a search term
 - (-) exclude a search term
 - (") use quotes around search phrases
 - (.) a single-character wildcard
 - (*) any word

Site Crawling

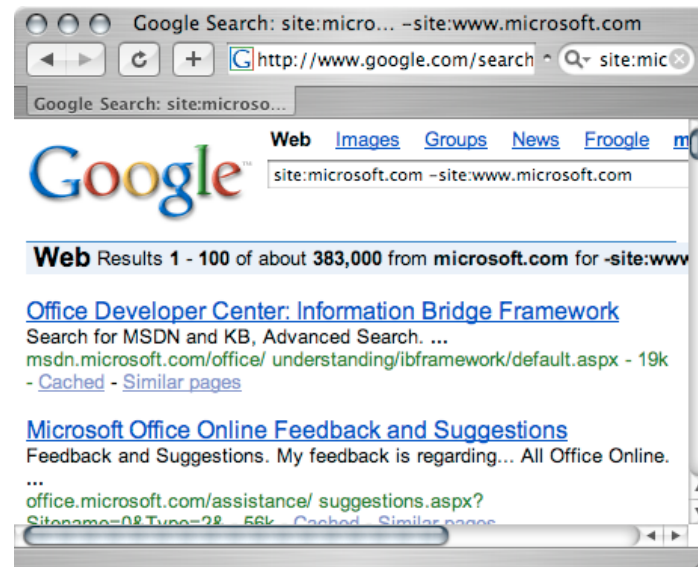
- To find every web page Google has crawled for a specific site, use the site: operator



site: microsoft.com

Server Crawling

- To locate additional servers, subtract common hostnames from the query



site: microsoft.com
-site:www.microsoft.com

Directory Listings

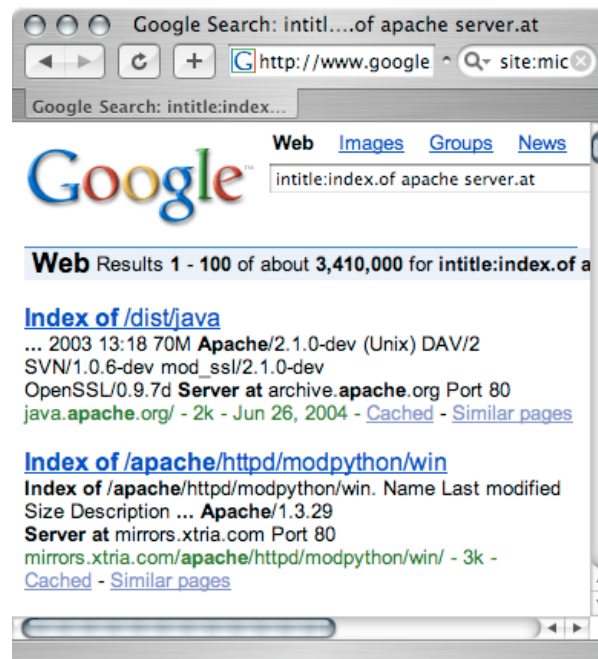
- Directory listings can be a source of great information



intitle:index.of/admin

Directory Listings

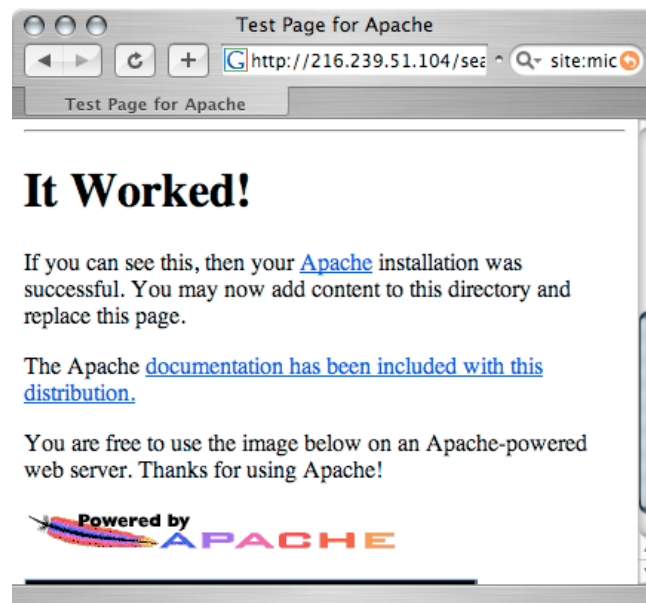
- Directory listings can provide server version information



intitle:index.of apache server.at

Default Server Pages

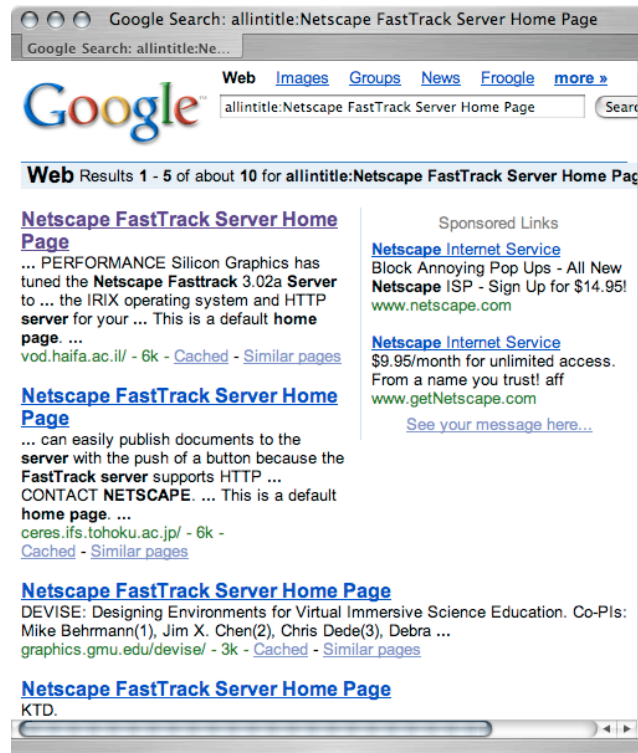
- Web servers with default pages can serve as juicy targets



intitle:test.page.for.apache "it worked"

Default Server Pages

- Web servers with default pages can serve as juicy targets



**allintitle:Netscape FastTrack Server
Home Page**

Default Server Pages

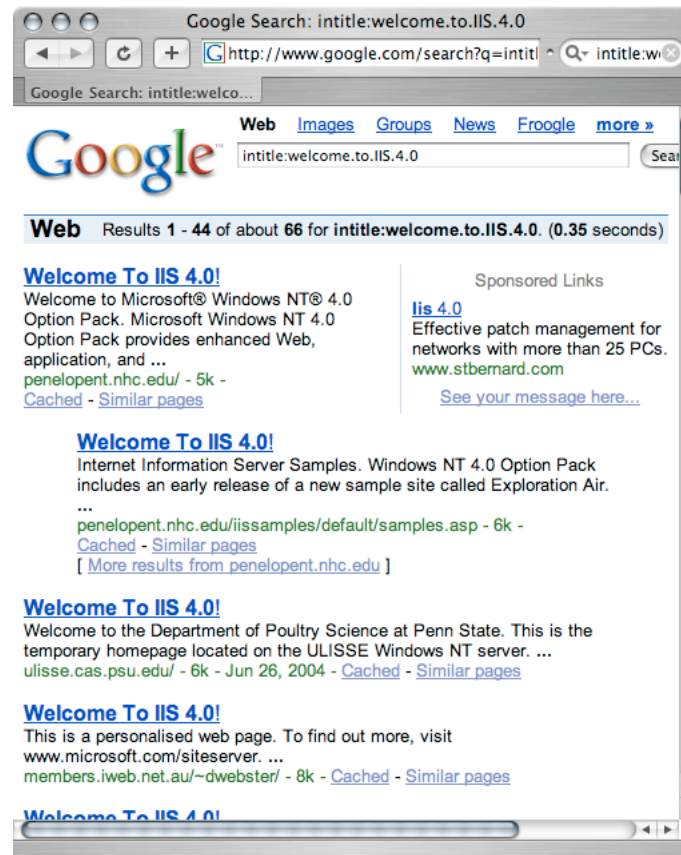
- Web servers with default pages can serve as juicy targets



intitle:"Welcome to Windows 2000 Internet Services"

Default Server Pages

- Web servers with default pages can serve as juicy targets



intitle:welcome.to.IIS.4.0

Default Server Pages

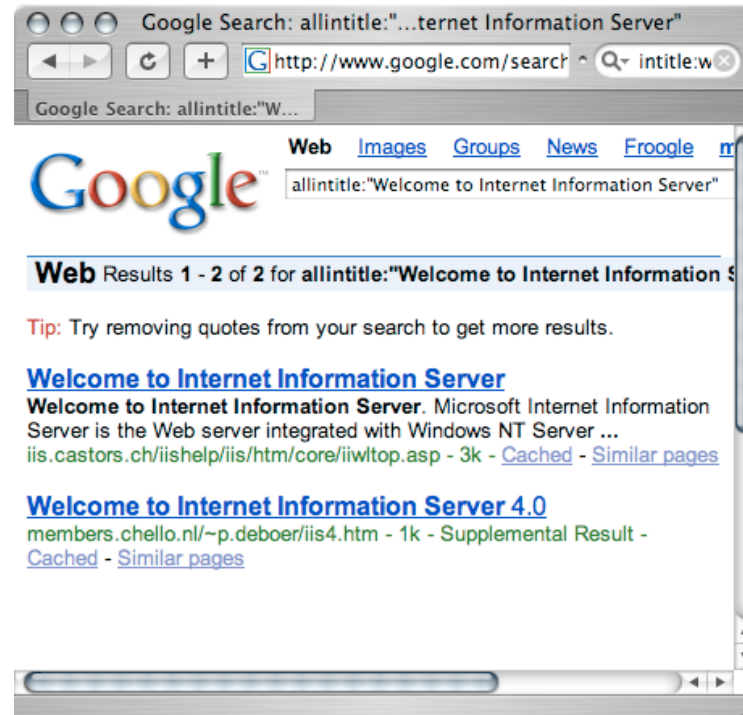
- Web servers with default pages can serve as juicy targets



**allintitle:Welcome to Windows XP
Server Internet Services**

Default Server Pages

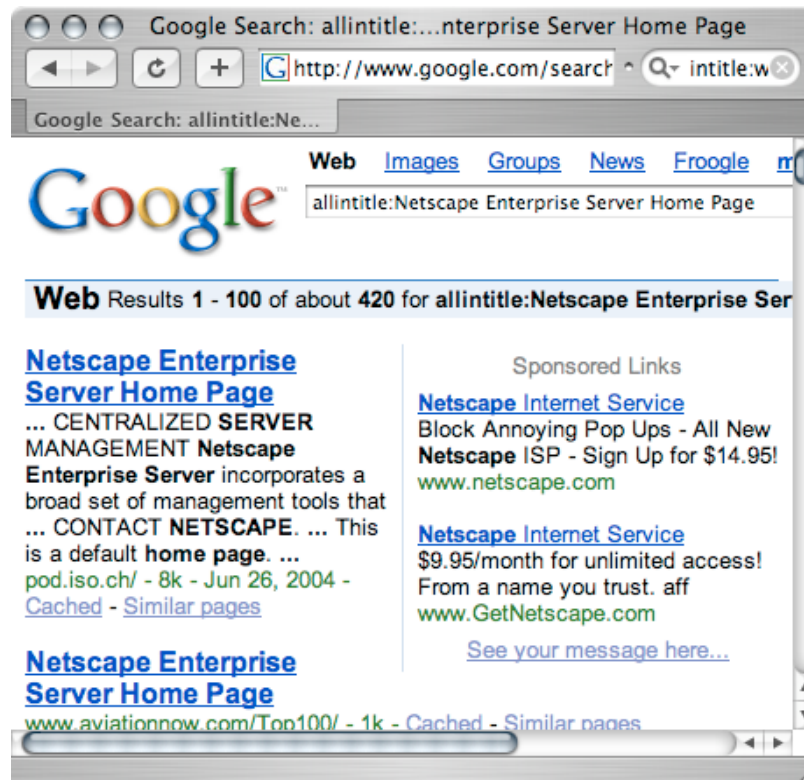
- Web servers with default pages can serve as juicy targets



**allintitle:"Welcome to Internet
Information Server"**

Default Server Pages

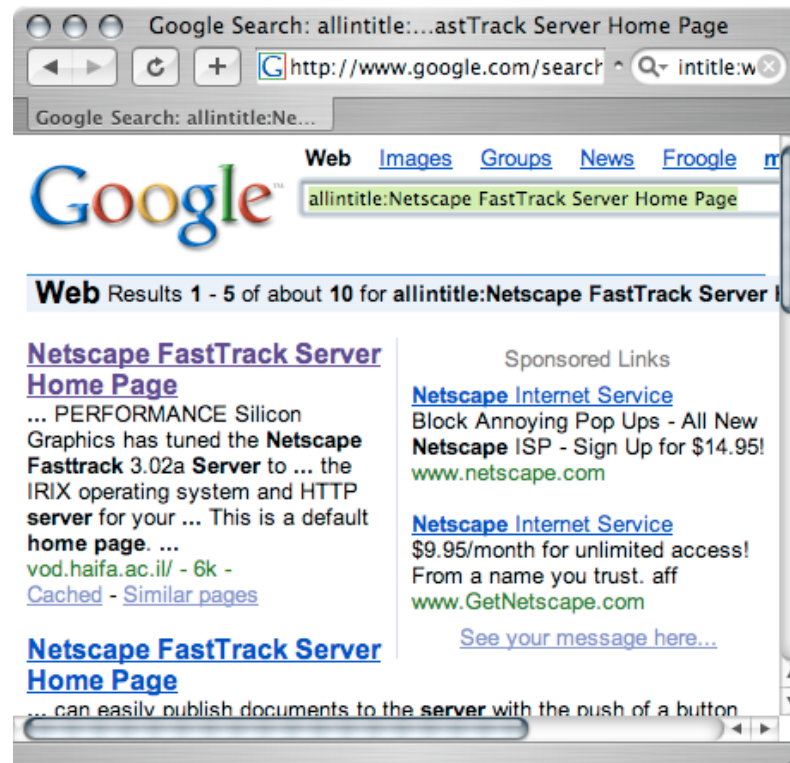
- Web servers with default pages can serve as juicy targets



allintitle:Netscape Enterprise Server Home Page

Default Server Pages

- Web servers with default pages can serve as juicy targets



**allintitle:Netscape FASTTRACK
Server Home Page**

Default Documents

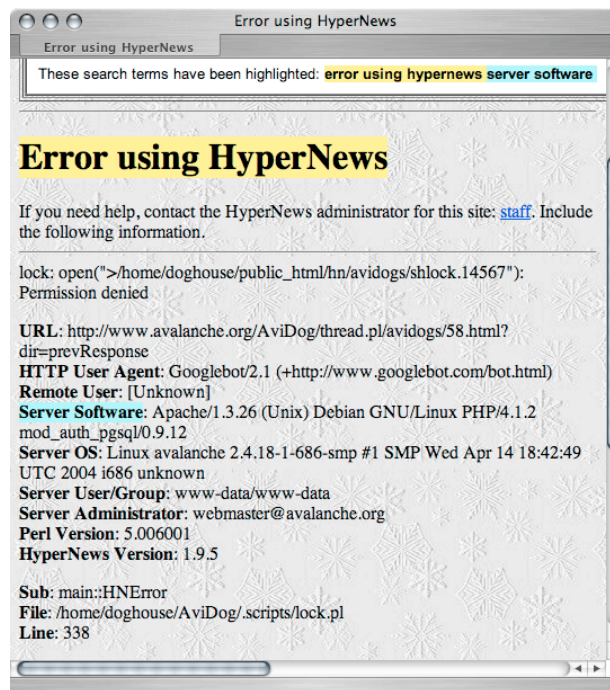
- Servers can also be profiled via default manuals and documentation



**intitle:"Apache HTTP Server"
intitle:"documentation"**

Error Messages

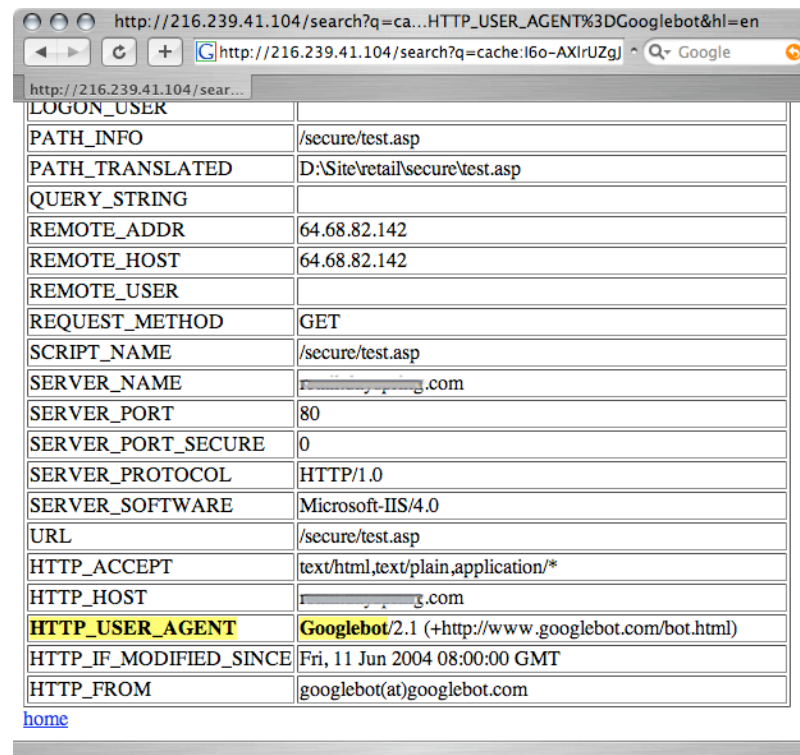
- Server profiling is easy with some error messages



**intitle:"Error using Hypernews"
"Server Software"**

Error Messages

- CGI environment vars provide a great deal of information
- The generic way to find these pages is by focusing on the trail left by the googlebot crawler



The screenshot shows a web browser window displaying a table of CGI environment variables. The browser's address bar shows the URL `http://216.239.41.104/search?q=ca...HTTP_USER_AGENT%3DGooglebot&hl=en`. The table lists various variables and their values, with the `HTTP_USER_AGENT` variable highlighted in yellow.

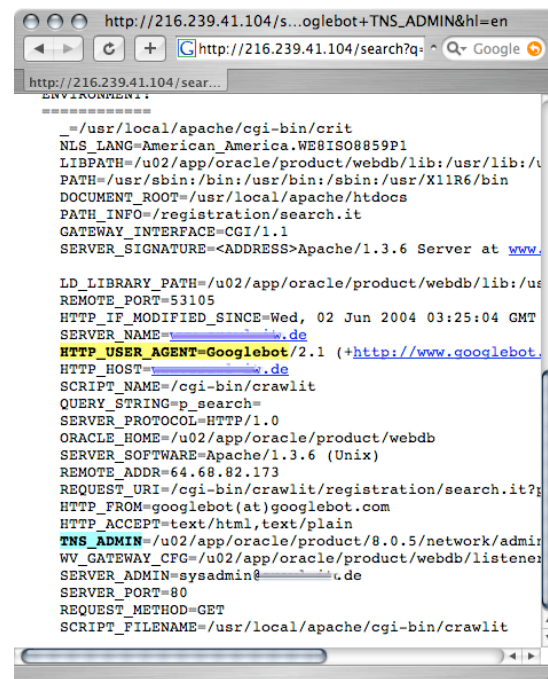
LOGON_USER	
PATH_INFO	/secure/test.asp
PATH_TRANSLATED	D:\Site\retail\secure\test.asp
QUERY_STRING	
REMOTE_ADDR	64.68.82.142
REMOTE_HOST	64.68.82.142
REMOTE_USER	
REQUEST_METHOD	GET
SCRIPT_NAME	/secure/test.asp
SERVER_NAME	retail.google.com
SERVER_PORT	80
SERVER_PORT_SECURE	0
SERVER_PROTOCOL	HTTP/1.0
SERVER_SOFTWARE	Microsoft-IIS/4.0
URL	/secure/test.asp
HTTP_ACCEPT	text/html,text/plain,application/*
HTTP_HOST	retail.google.com
HTTP_USER_AGENT	Googlebot/2.1 (+http://www.googlebot.com/bot.html)
HTTP_IF_MODIFIED_SINCE	Fri, 11 Jun 2004 08:00:00 GMT
HTTP_FROM	googlebot(at)googlebot.com

[home](#)

“HTTP_USER_AGENT=Googlebot”

Error Messages

- after a generic search, we can narrow down to the fields we find more interesting



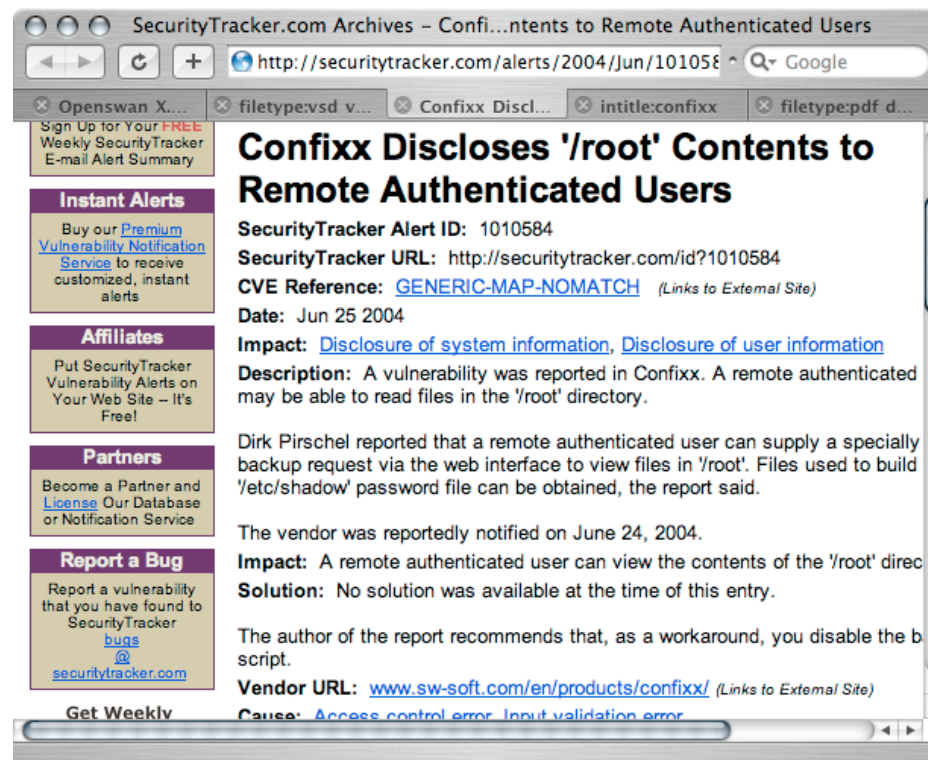
```
ENVIRONMENT:  
=====
```

```
  _=/usr/local/apache/cgi-bin/crit  
  NLS_LANG=American_America.WE8ISO8859P1  
  LIBPATH=/u02/app/oracle/product/webdb/lib:/usr/lib:/u02/app/oracle/product/webdb/lib:/usr/lib:/u02/app/oracle/product/webdb/lib:/usr/lib:/u02/app/oracle/product/webdb/lib:/usr/lib  
  PATH=/usr/sbin:/bin:/usr/bin:/sbin:/usr/X11R6/bin  
  DOCUMENT_ROOT=/usr/local/apache/htdocs  
  PATH_INFO=/registration/search.it  
  GATEWAY_INTERFACE=CGI/1.1  
  SERVER_SIGNATURE=<ADDRESS>Apache/1.3.6 Server at www.  
  
  LD_LIBRARY_PATH=/u02/app/oracle/product/webdb/lib:/usr/lib:/u02/app/oracle/product/webdb/lib:/usr/lib:/u02/app/oracle/product/webdb/lib:/usr/lib  
  REMOTE_PORT=53105  
  HTTP_IF_MODIFIED_SINCE=Wed, 02 Jun 2004 03:25:04 GMT  
  SERVER_NAME=www.de  
  HTTP_USER_AGENT=Googlebot/2.1 (+http://www.googlebot.com)  
  HTTP_HOST=www.de  
  SCRIPT_NAME=/cgi-bin/crawlit  
  QUERY_STRING=p_search=  
  SERVER_PROTOCOL=HTTP/1.0  
  ORACLE_HOME=/u02/app/oracle/product/webdb  
  SERVER_SOFTWARE=Apache/1.3.6 (Unix)  
  REMOTE_ADDR=64.68.82.173  
  REQUEST_URI=/cgi-bin/crawlit/registration/search.it?  
  HTTP_FROM=googlebot(at)googlebot.com  
  HTTP_ACCEPT=text/html,text/plain  
  TNS_ADMIN=/u02/app/oracle/product/8.0.5/network/admin  
  WV_GATEWAY_CFG=/u02/app/oracle/product/webdb/listener  
  SERVER_ADMIN=sysadmin@www.de  
  SERVER_PORT=80  
  REQUEST_METHOD=GET  
  SCRIPT_FILENAME=/usr/local/apache/cgi-bin/crawlit
```

“HTTP_USER_AGENT=Googlebot”
TNS_ADMIN

Vulnerability Trolling

- Many attackers find vulnerable targets via Google
- A typical security advisory may look like this:



The screenshot shows a web browser window with the address bar displaying "http://securitytracker.com/alerts/2004/Jun/1010584". The page title is "SecurityTracker.com Archives - Confi...ntents to Remote Authenticated Users". The main content area features a large heading: "Confixx Discloses '/root' Contents to Remote Authenticated Users". Below the heading, the alert ID is "1010584" and the URL is "http://securitytracker.com/id?1010584". The CVE reference is "GENERIC-MAP-NOMATCH". The date is "Jun 25 2004". The impact is described as "Disclosure of system information, Disclosure of user information". The description states: "A vulnerability was reported in Confixx. A remote authenticated user may be able to read files in the '/root' directory." A paragraph follows: "Dirk Pirschel reported that a remote authenticated user can supply a specially backup request via the web interface to view files in '/root'. Files used to build '/etc/shadow' password file can be obtained, the report said." The vendor was notified on June 24, 2004. The impact is: "A remote authenticated user can view the contents of the '/root' directory." The solution is: "No solution was available at the time of this entry." A recommendation is given: "The author of the report recommends that, as a workaround, you disable the backup script." The vendor URL is "www.sw-soft.com/en/products/confixx/". The cause is "Access control error, Input validation error".

Confixx Discloses '/root' Contents to Remote Authenticated Users

SecurityTracker Alert ID: 1010584
SecurityTracker URL: <http://securitytracker.com/id?1010584>
CVE Reference: [GENERIC-MAP-NOMATCH](#) (Links to External Site)
Date: Jun 25 2004

Impact: [Disclosure of system information](#), [Disclosure of user information](#)

Description: A vulnerability was reported in Confixx. A remote authenticated user may be able to read files in the '/root' directory.

Dirk Pirschel reported that a remote authenticated user can supply a specially backup request via the web interface to view files in '/root'. Files used to build '/etc/shadow' password file can be obtained, the report said.

The vendor was reportedly notified on June 24, 2004.

Impact: A remote authenticated user can view the contents of the '/root' directory.

Solution: No solution was available at the time of this entry.

The author of the report recommends that, as a workaround, you disable the backup script.

Vendor URL: www.sw-soft.com/en/products/confixx/ (Links to External Site)

Cause: [Access control error](#), [Input validation error](#)

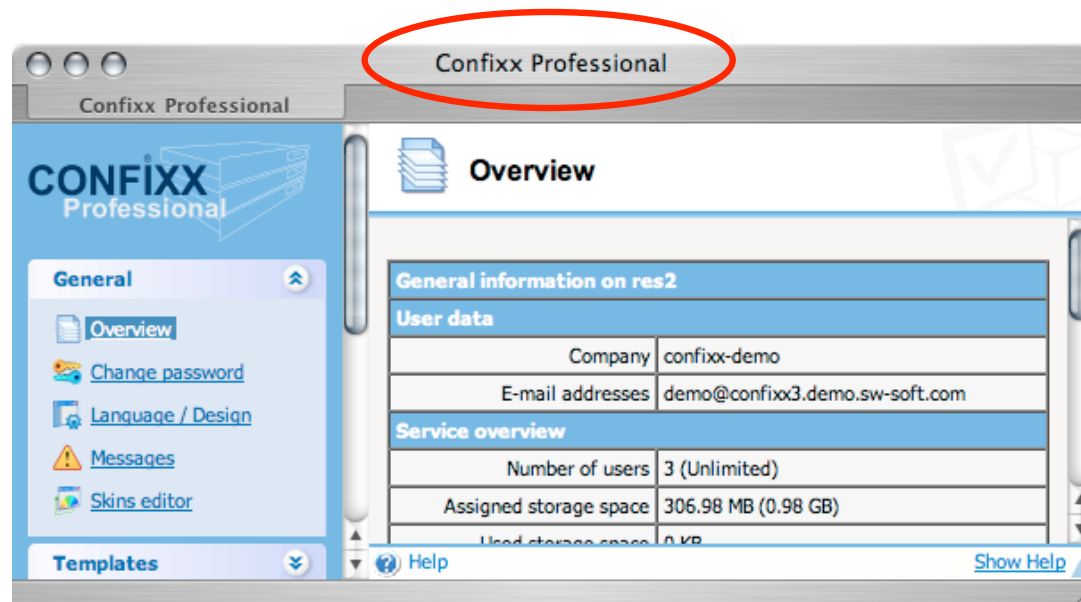
Vulnerability Trolling

- A quick browse of the vendor's website reveals a demo of the product



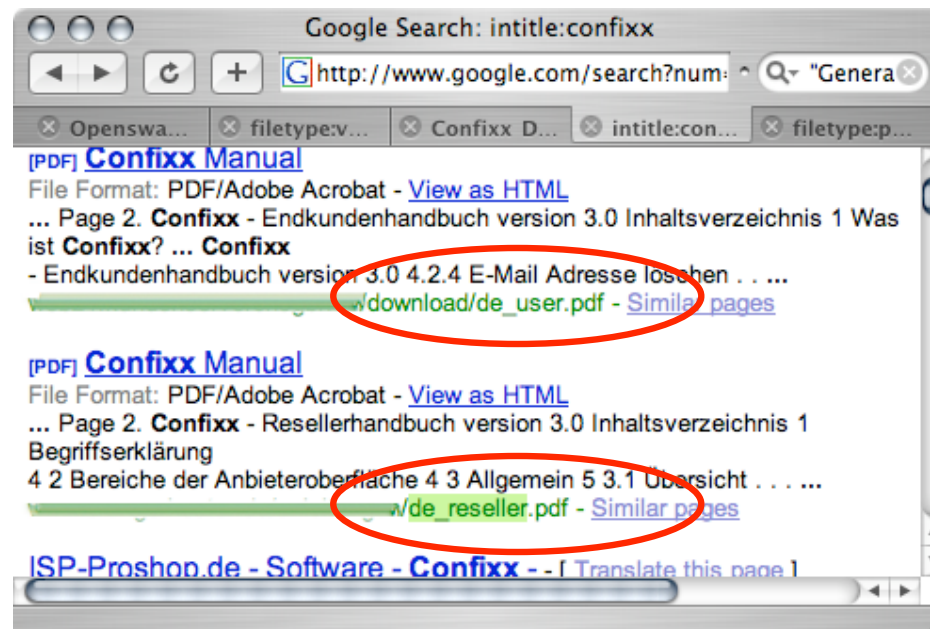
Vulnerability Trolling

- The demo page suggests one method for finding targets



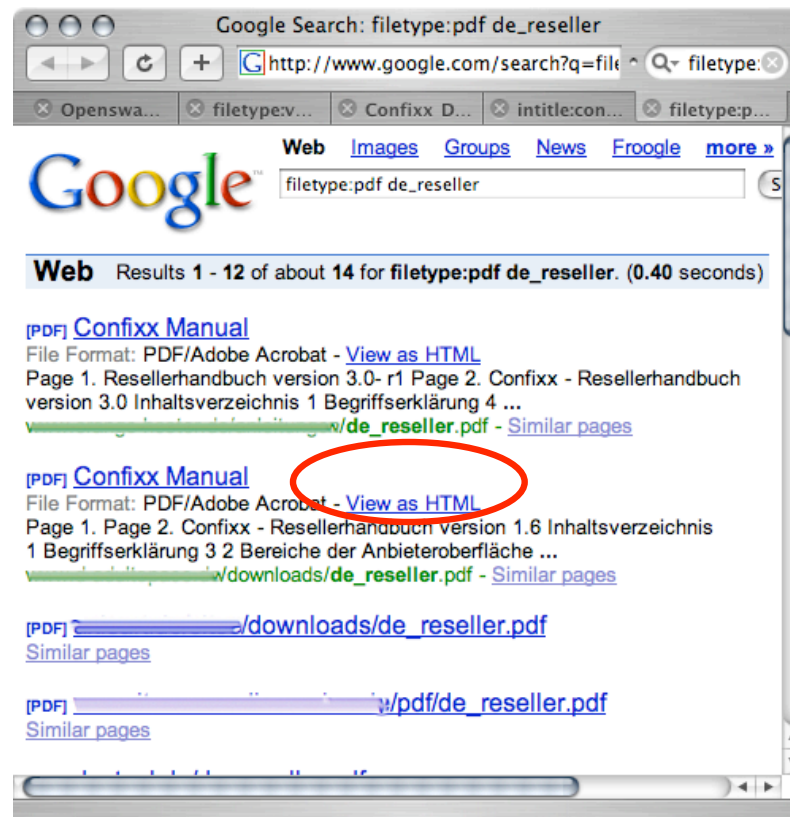
Vulnerability Trolling

- A quick intitle: search suggests more vectors...



Vulnerability Trolling

- This search finds the documentation included with the product
- These sites are probably poorly configured



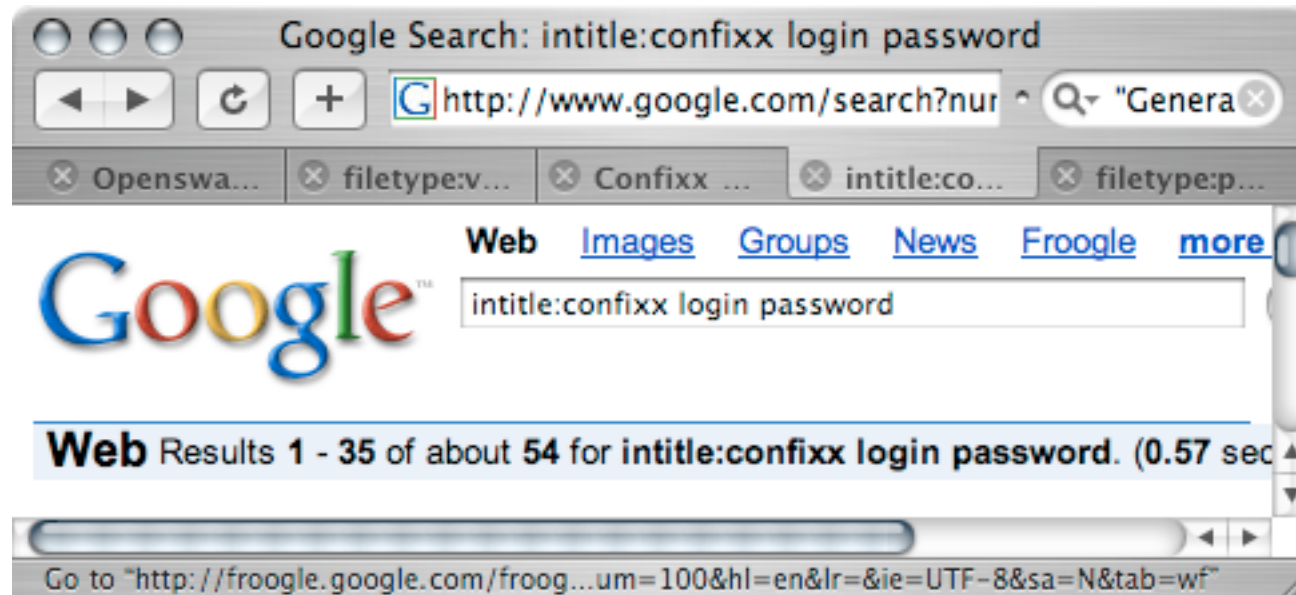
Vulnerability Trolling

- Other searches are easy to discover as well...



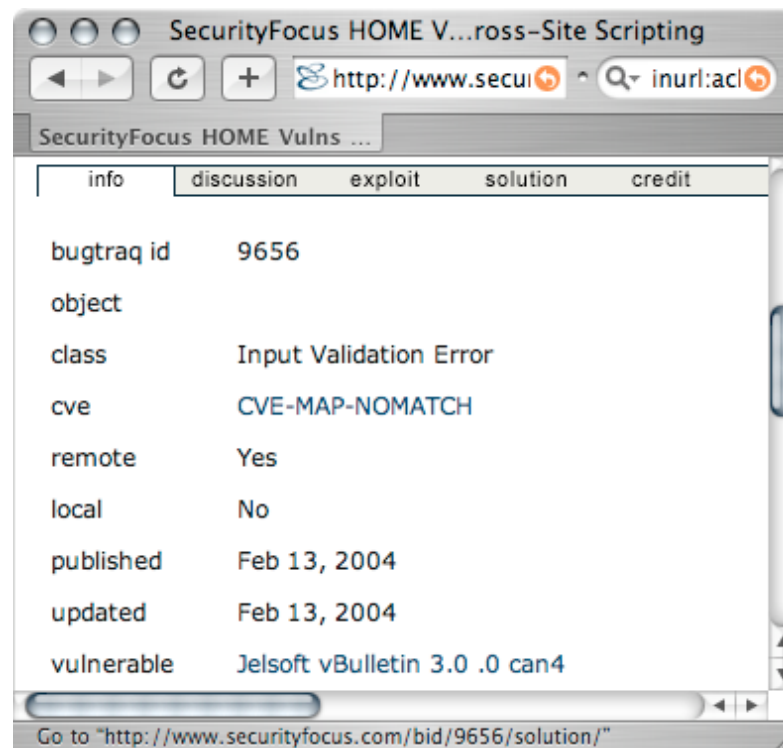
Vulnerability Trolling

- Other searches are easy to discover as well...



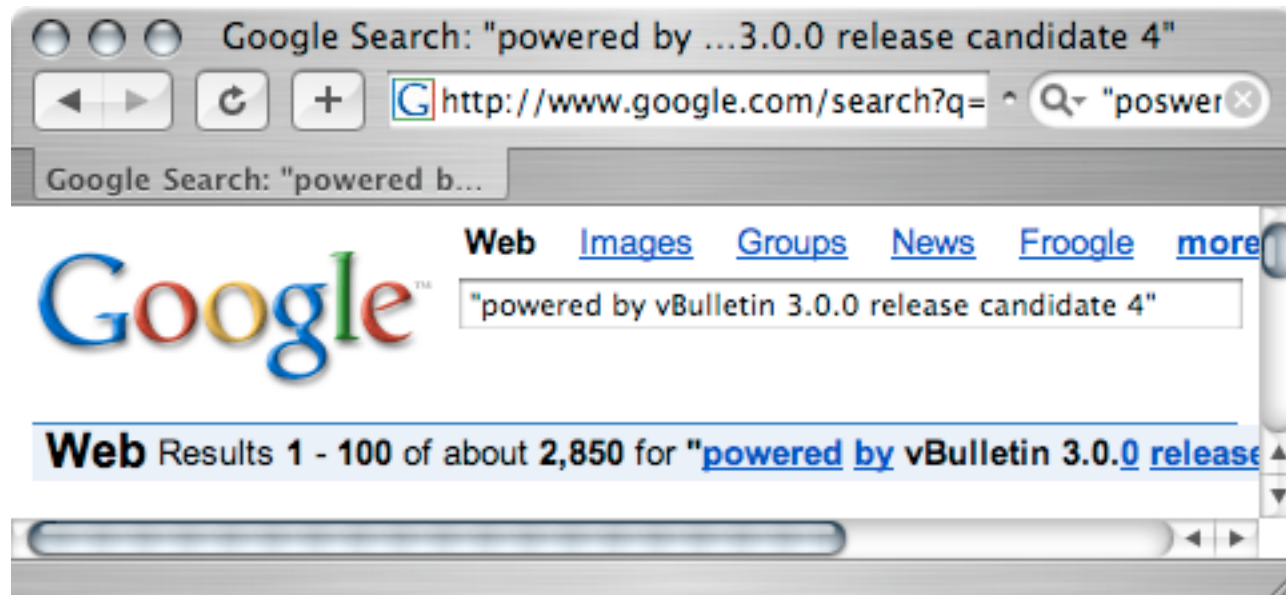
Vulnerability Trolling

- Many times, a good search string is much simpler to come up with
- Consider this advisory:



Vulnerability Trolling

- A creative search finds vulnerable targets



CGI Scanning

- In order to locate web vulnerabilities on a larger scale, many attacker will use a 'CGI' scanner
- Most scanners read a data file and query target web servers looking for the vulnerable files

CGI Scanning

- A CGI scanner's vulnerability file...

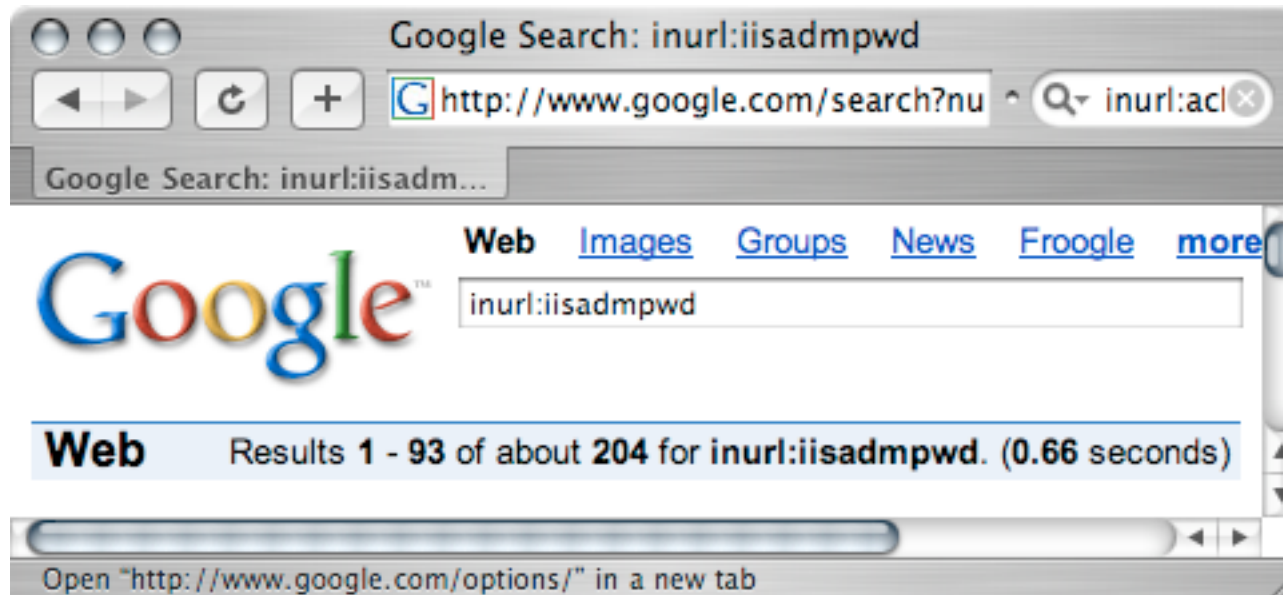
```
/iisadmpwd/  
/iisadmpwd/achg.htr  
/iisadmpwd/aexp.htr  
/iisadmpwd/aexp2.htr  
/iisadmpwd/aexp2b.htr
```

- can be converted to Google queries in a number of different ways:

```
inurl;/iisadmpwd/  
inurl;/iisadmpwd/achg.htr  
inurl;/iisadmpwd/aexp.htr  
inurl;/iisadmpwd/aexp2.htr  
inurl;/iisadmpwd/aexp2b.htr
```

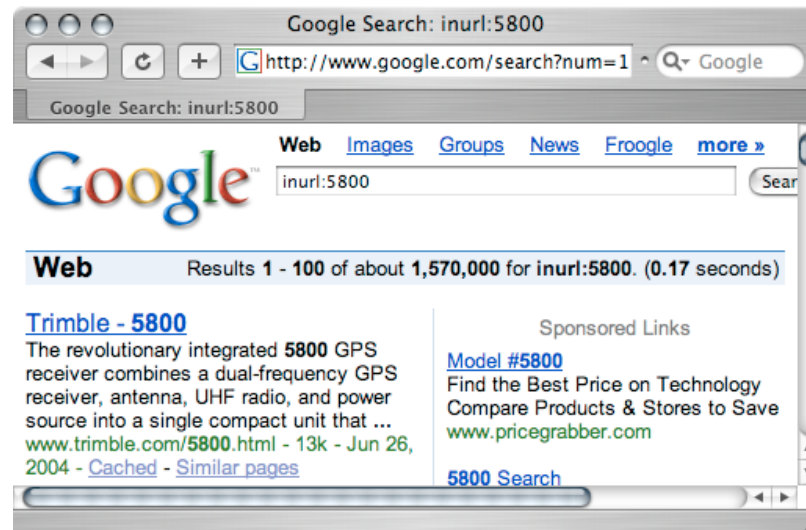
Vulnerability Trolling

- Regardless of the age of the vulnerability, there are usually vulnerable targets



Port Scanning

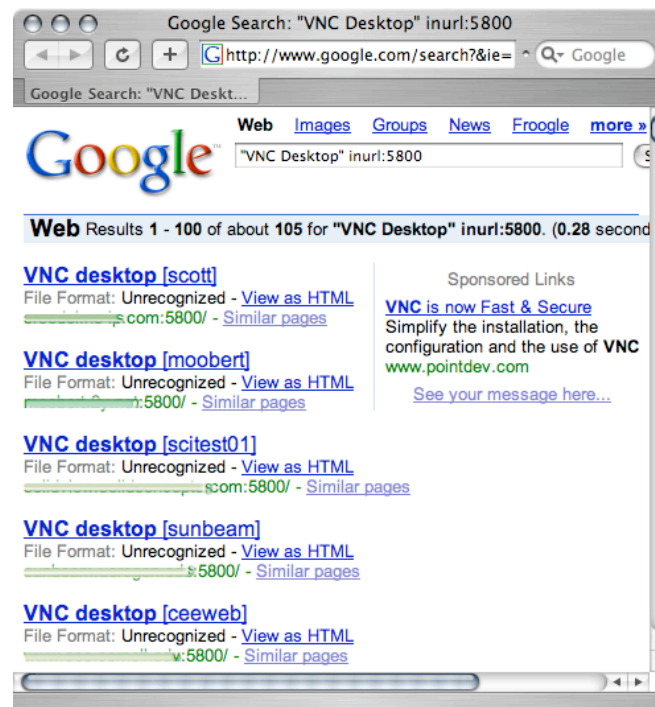
- Although port numbers are sometimes found in the url, there's no easy way to scan just for a port number... the results are much too copious



inurl:5800

Port Scanning

- We can use creative queries to sniff out services that may be listening on particular ports
- VNC Desktop, port 5800



"VNC Desktop" inurl:5800

Port Scanning

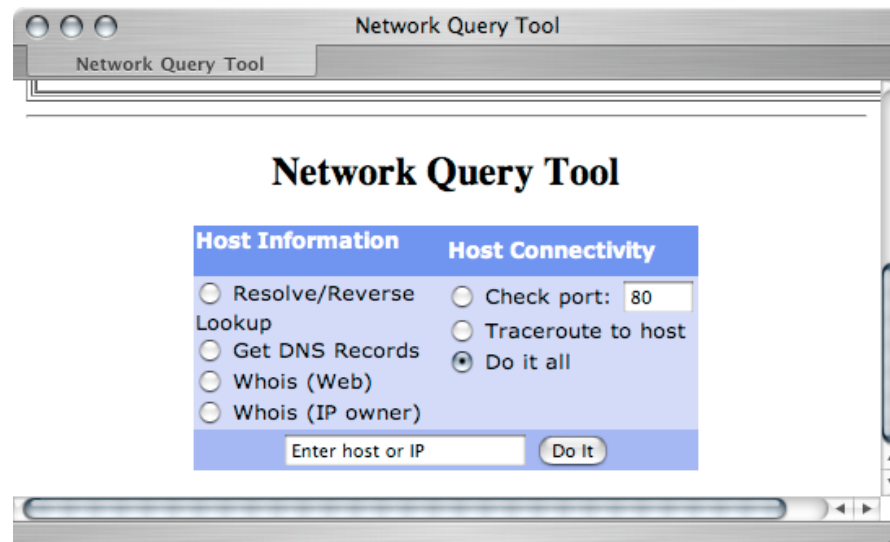
- Webmin, port 10000



inurl:webmin inurl:10000

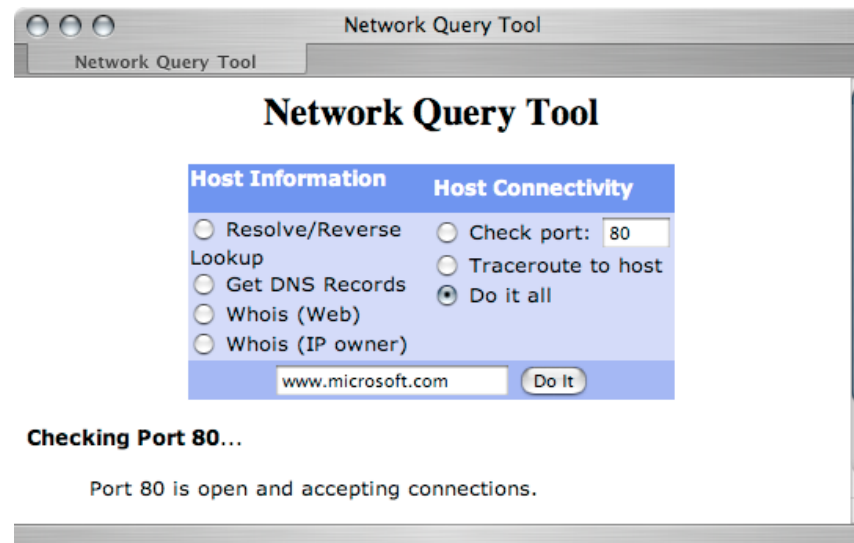
Port Scanning

- Google can be used to find sites to do the portscanning for you
- Consider the Network Query Tool



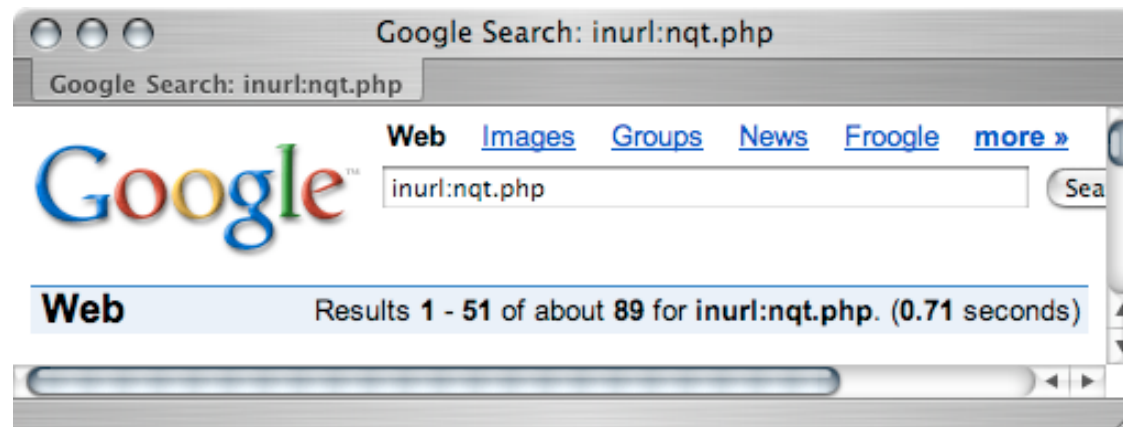
Port Scanning

- NQT allows web users to perform traceroutes, rdns lookups and port scans.
- This is the NQT program checking port 80 on www.microsoft.com:



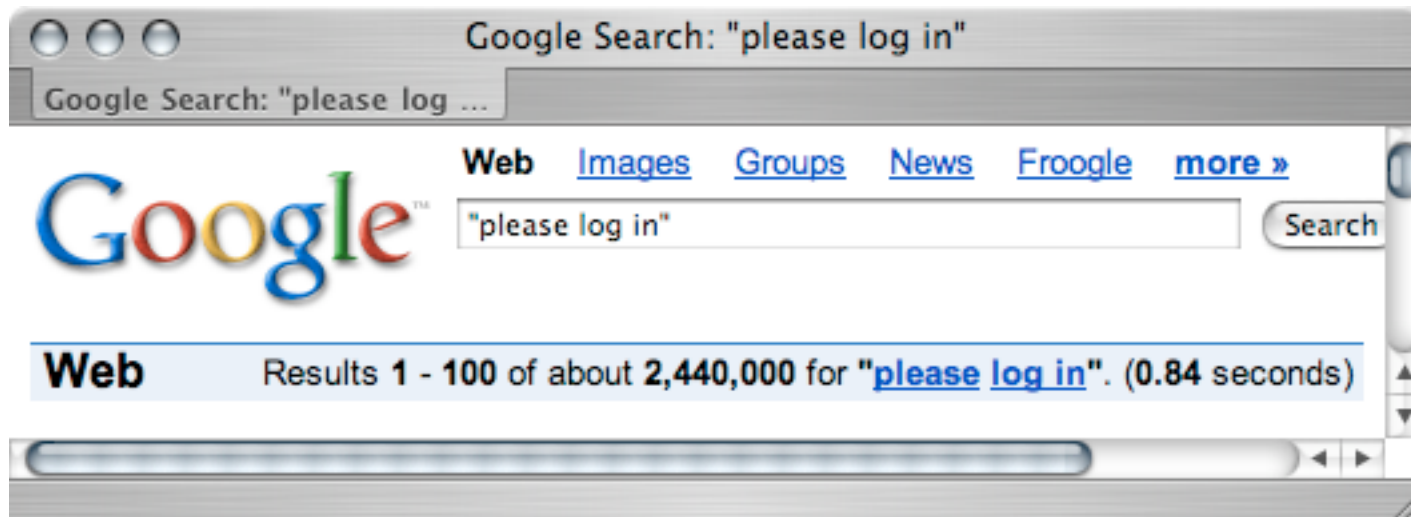
Port Scanning

- Google can be used to locate servers running the NQT program, nqt.php
- Once servers are harvested, they can be used to perform port scans (usually through a web proxy)
- NQT also allows remote posts, so that more than one port can be checked at a time



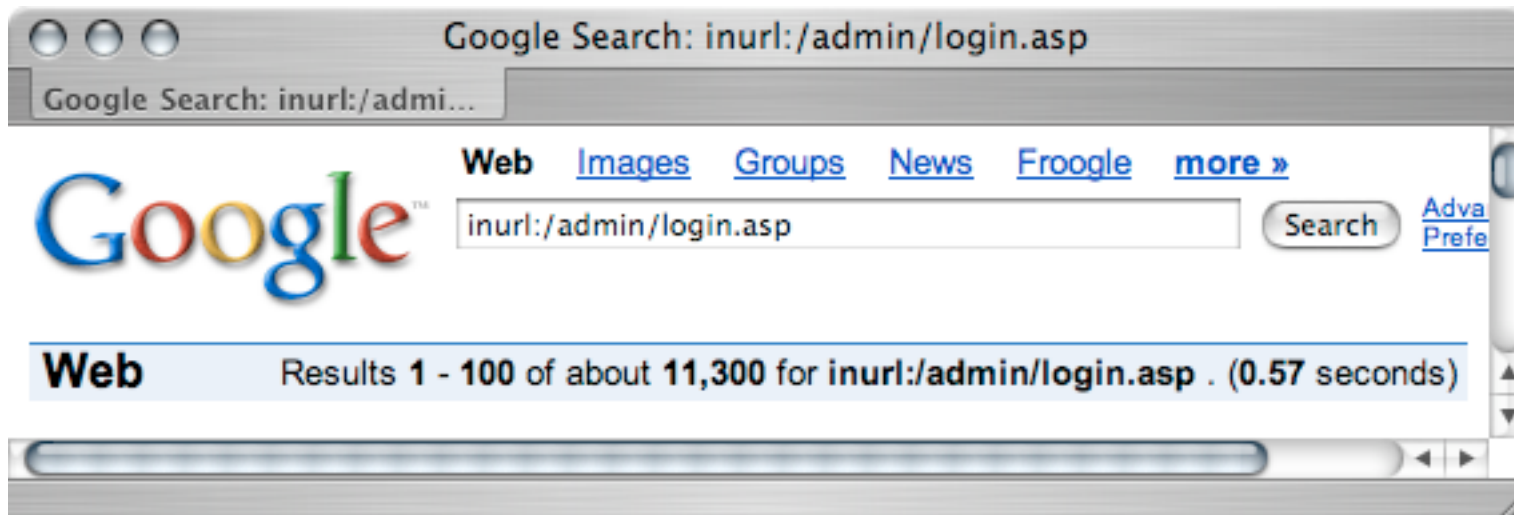
Login Portals

- The most generic of login portals



Login Portals

- Another very generic portal



Login Portals

- Microsoft Outlook Web Access



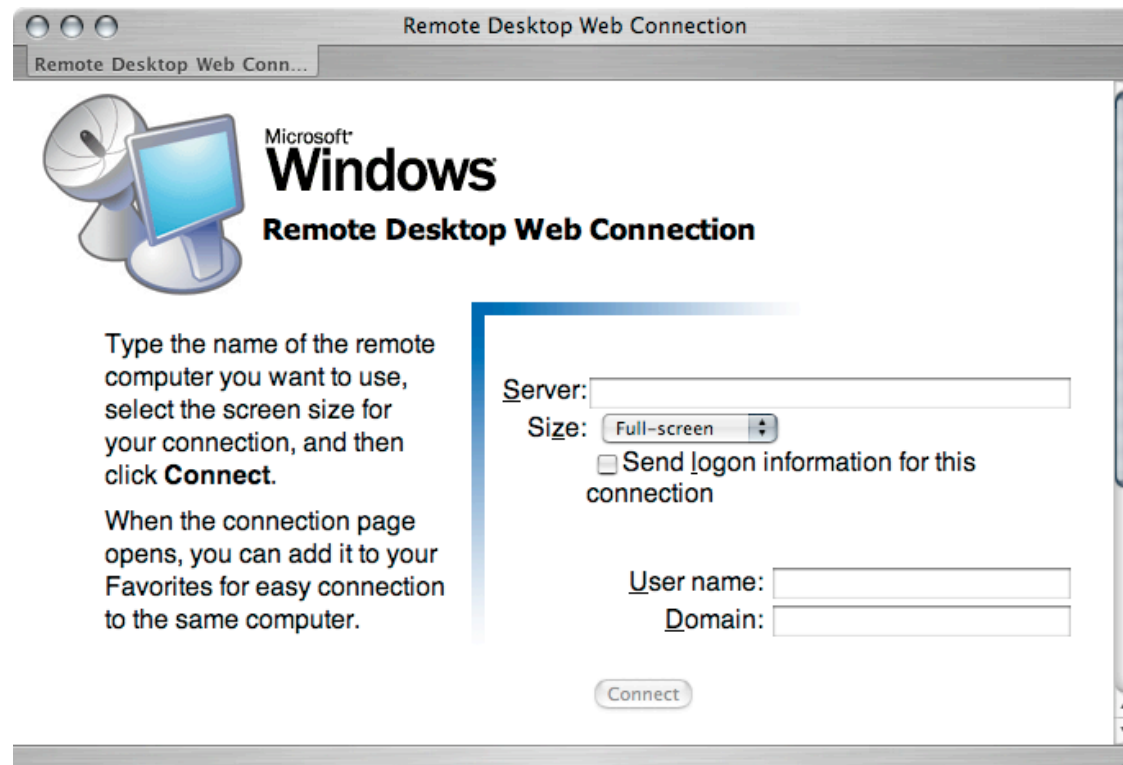
Login Portals

- Coldfusion Admin Page



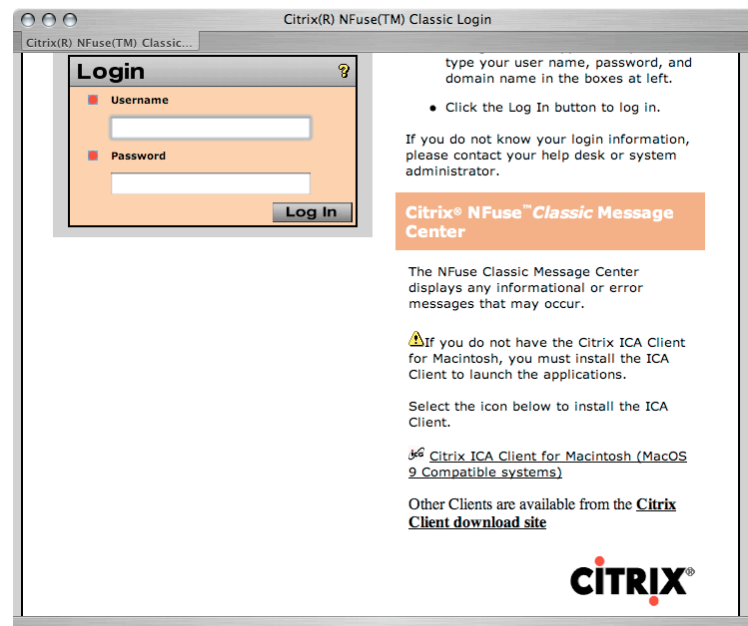
Login Portals

- Windows Remote Desktop



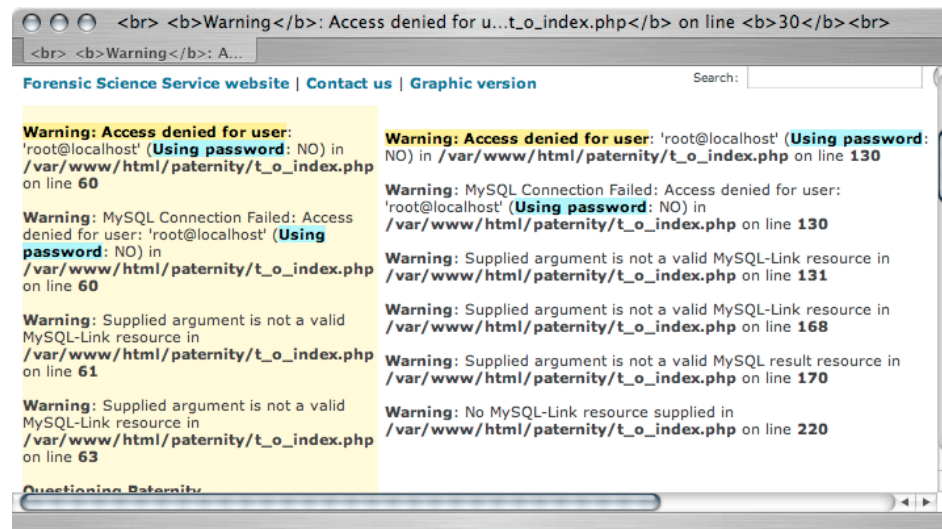
Login Portals

- Citrix Metaframe



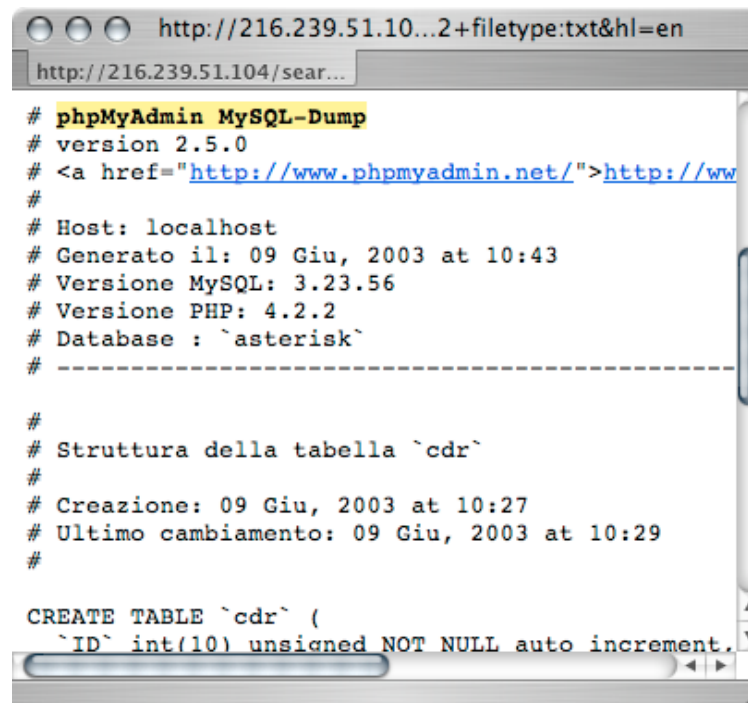
SQL Information

- Gathering SQL usernames is simple with this search



SQL Information

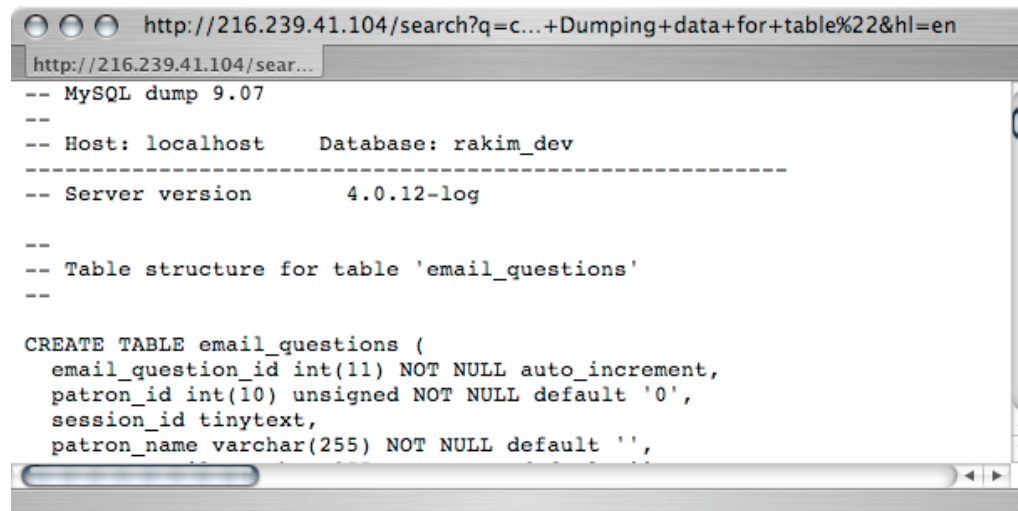
- This is an SQL dump made by phpmyadmin

A screenshot of a web browser window showing an SQL dump file. The address bar contains the URL 'http://216.239.51.104/sear...'. The main content area displays the following text:

```
# phpMyAdmin MySQL-Dump
# version 2.5.0
# <a href="http://www.phpmyadmin.net/">http://ww
#
# Host: localhost
# Generato il: 09 Giu, 2003 at 10:43
# Versione MySQL: 3.23.56
# Versione PHP: 4.2.2
# Database : `asterisk`
# -----
#
# Struttura della tabella `cdr`
#
# Creazione: 09 Giu, 2003 at 10:27
# Ultimo cambiamento: 09 Giu, 2003 at 10:29
#
CREATE TABLE `cdr` (
  `ID` int(10) unsigned NOT NULL auto increment,
```

SQL Information

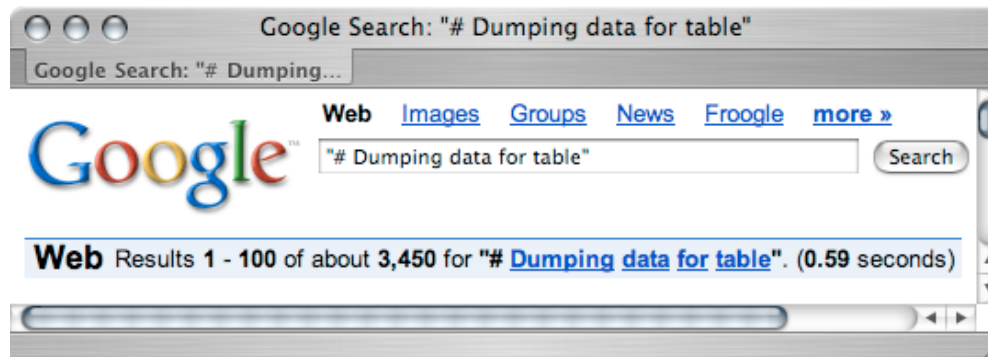
- This is a complete database schema dump, essentially a complete database backup



```
http://216.239.41.104/search?q=c...+Dumping+data+for+table%22&hl=en
http://216.239.41.104/sear...
-- MySQL dump 9.07
--
-- Host: localhost    Database: rakim_dev
-----
-- Server version    4.0.12-log
--
-- Table structure for table 'email_questions'
--
CREATE TABLE email_questions (
  email_question_id int(11) NOT NULL auto_increment,
  patron_id int(10) unsigned NOT NULL default '0',
  session_id tinytext,
  patron_name varchar(255) NOT NULL default '',
```

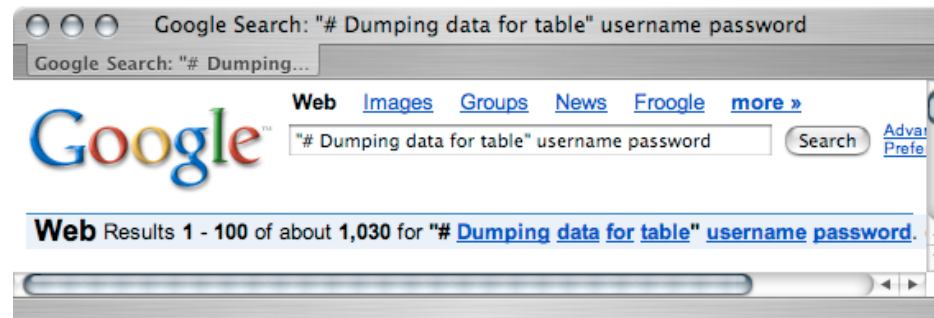
SQL Information

- This query will locate SQL schemas on the web



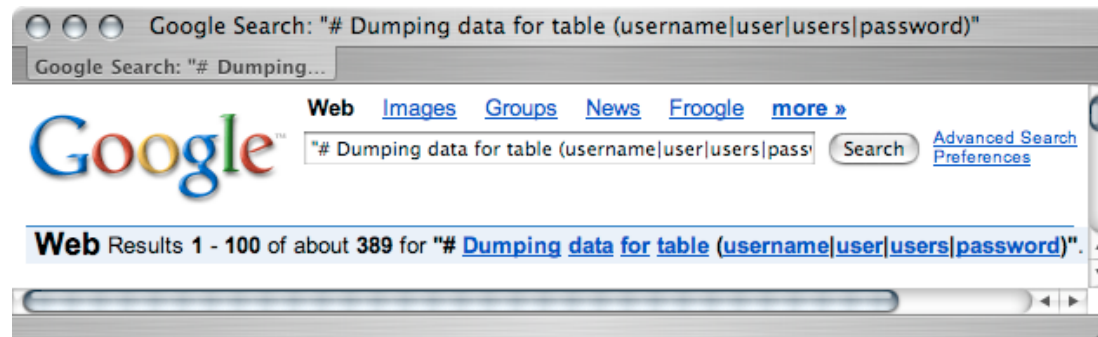
SQL Information

- In addition, this query finds the words username and password inside the SQL dump



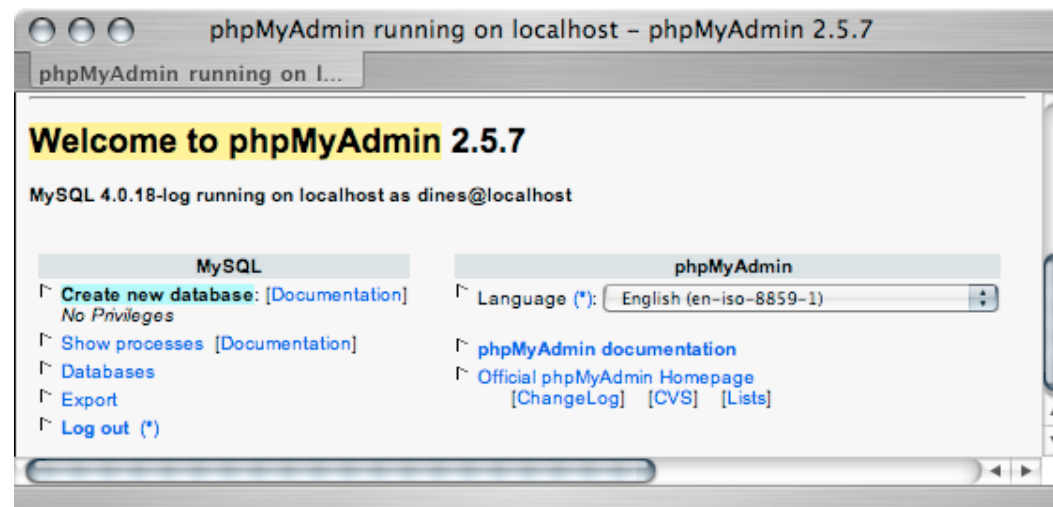
SQL Information

- This potent query finds SQL dumps wither username, user, users or password as a table name



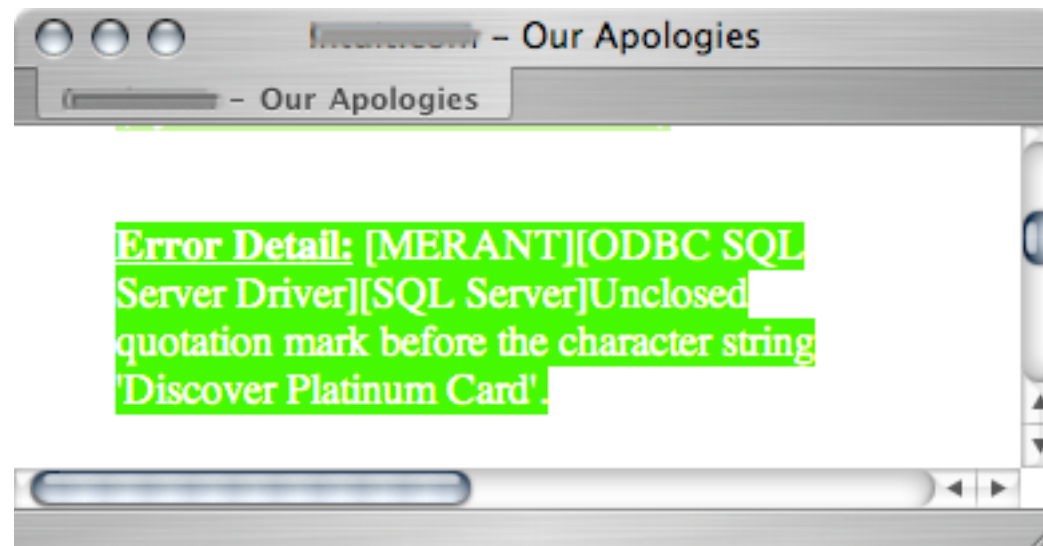
SQL Information

- This graphical front-end to SQL is mis-configured to allow anyone admin access



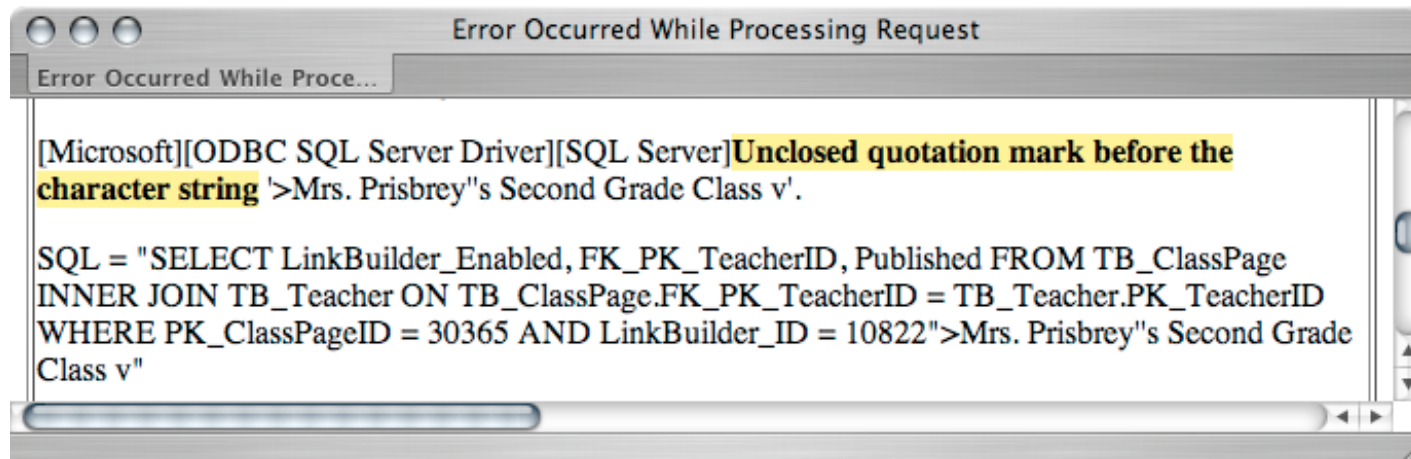
SQL Information

- ...another SQL injection target...



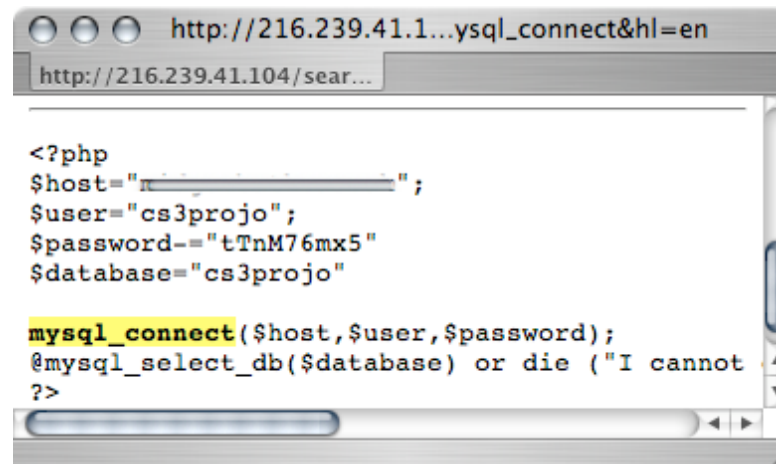
SQL Information

- ..and another...



SQL Information

- the `mysql_connect` function makes a database query with a supplied username and password
- This file should not be on the web



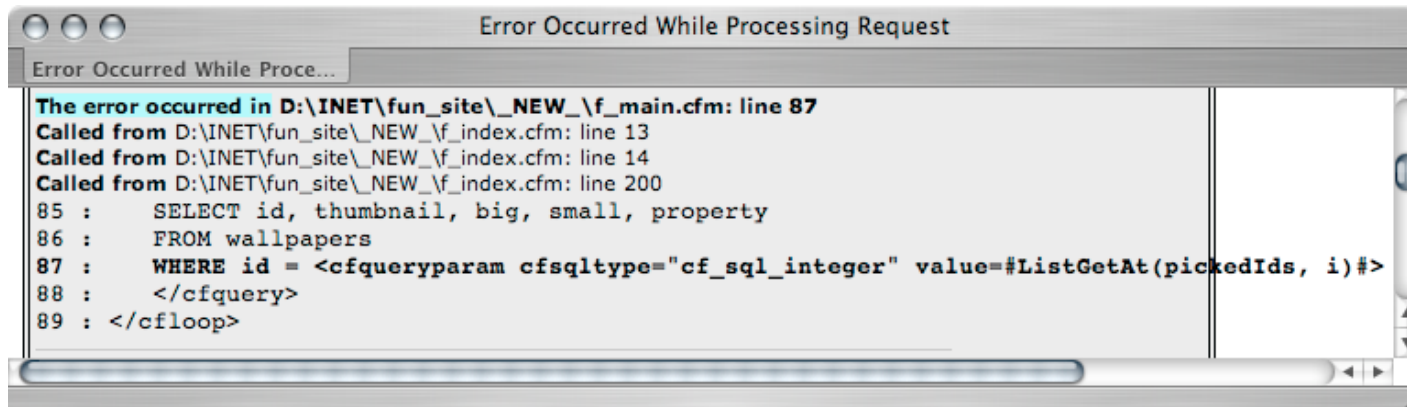
A screenshot of a web browser window. The address bar shows the URL `http://216.239.41.104/ysql_connect&hl=en`. The page content displays PHP code for connecting to a MySQL database. The code is as follows:

```
<?php
$host="";
$user="cs3projo";
$password="tTnM76mx5";
$database="cs3projo"

mysql_connect($host,$user,$password);
@mysql_select_db($database) or die ("I cannot
?>
```

SQL Information

- In most cases, there's nothing better for an SQL injector than a complete line of SQL source code...

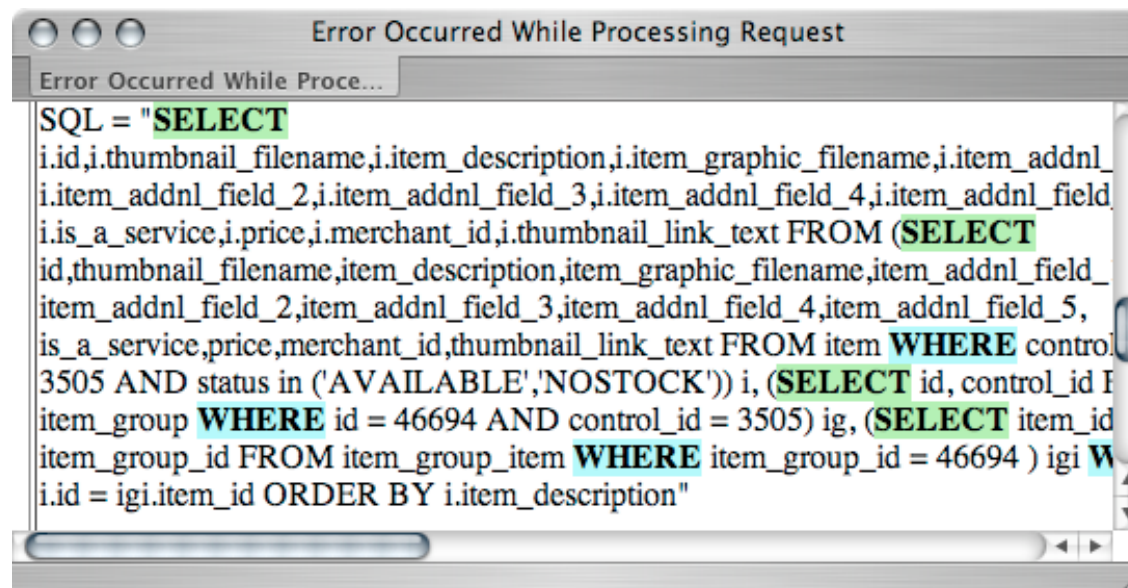


The screenshot shows a window titled "Error Occurred While Processing Request". The error message indicates that the error occurred in `D:\INET\fun_site_NEW_f_main.cfm` at line 87. The stack trace shows the error was called from `D:\INET\fun_site_NEW_f_index.cfm` at lines 13, 14, and 200. The SQL query snippet is as follows:

```
85 :   SELECT id, thumbnail, big, small, property
86 :   FROM wallpapers
87 :   WHERE id = <cfqueryparam cfsqltype="cf_sql_integer" value=#ListGetAt(pickedIds, i)#>
88 :   </cfquery>
89 : </cfloop>
```

SQL Information

- ...except for really long lines of SQL code...



The image shows a screenshot of a web browser window with a title bar that reads "Error Occurred While Processing Request". The main content area of the browser displays a long SQL query. The query is wrapped across multiple lines and contains several keywords highlighted in green: **SELECT**, **WHERE**, and **ORDER BY**. The query is as follows:

```
SQL = "SELECT  
i.id,i.thumbnail_filename,i.item_description,i.item_graphic_filename,i.item_addnl_  
i.item_addnl_field_2,i.item_addnl_field_3,i.item_addnl_field_4,i.item_addnl_field_  
i.is_a_service,i.price,i.merchant_id,i.thumbnail_link_text FROM (SELECT  
id,thumbnail_filename,item_description,item_graphic_filename,item_addnl_field_  
item_addnl_field_2,item_addnl_field_3,item_addnl_field_4,item_addnl_field_5,  
is_a_service,price,merchant_id,thumbnail_link_text FROM item WHERE control  
3505 AND status in ('AVAILABLE','NOSTOCK')) i, (SELECT id, control_id F  
item_group WHERE id = 46694 AND control_id = 3505) ig, (SELECT item_id  
item_group_id FROM item_group_item WHERE item_group_id = 46694 ) igi W  
i.id = igi.item_id ORDER BY i.item_description"
```

Examples

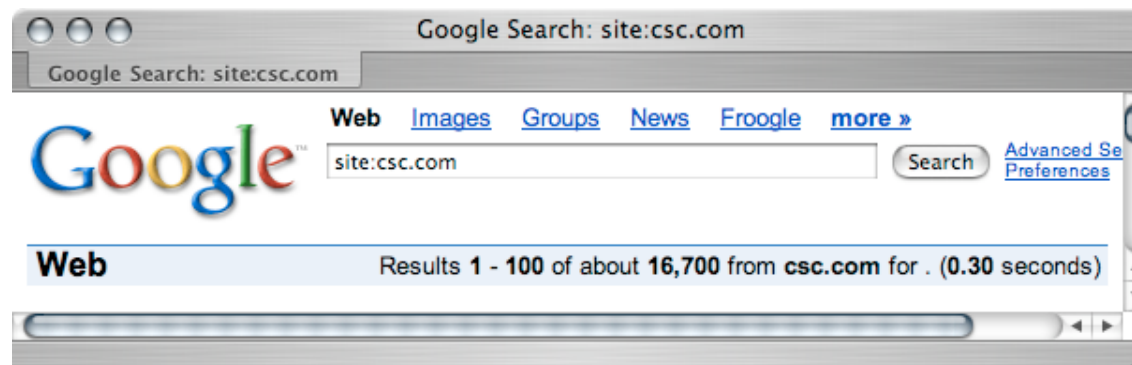
- *** LIVE EXAMPLES REMOVED FROM PRINT VERSION***

Prevention

- Do not put sensitive data on your web site, even temporarily
- Proactively check your web presence with Google on a regular basis
- Use sites like <http://johnny.ihackstuff.com> to keep up on the latest “Google Hacks”

Prevention

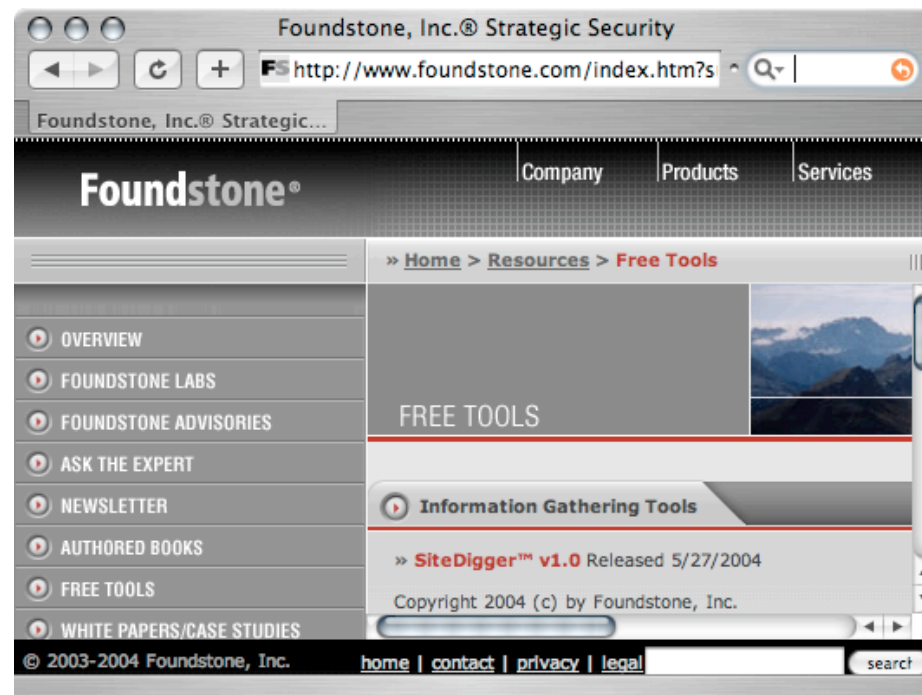
- Use site: queries against each of your web hosts
- Don't forget about hosts that do not have DNS names



- Scan each result page, ensuring that each and every page it supposed to be in Google's database

Prevention

- Automate your scans with tools like sitedigger by Foundstone



Presentation Materials

- This is a condensed version of the actual presentation given at the event
- For more information, please see:
<http://johnny.ihackstuff.com>
- e-mail: johnny@ihackstuff.com

Thanks

- Thanks to God for the gift of life.
- Thanks to my wife for the gift of love.
- Thanks to my kids for the gift of laughter.
- Thanks to my friends for filling in the blanks.