Watching the Watchers

Target Exploitation via Public Search Engines

mail://johnny@ihackstuff.com
http://johnny.ihackstuff.com

what's this about?

- using search engines to do interesting (sometimes unintended) stuff
  - sp3ak l1ke l33to hax0rs
  - act as transparent proxy servers
  - sneak past security
  - find development sites

what's this about?

- using search engines to find exploitable targets on the web which
  - run certain operating systems
  - run certain web server software
  - harbor specific vulnerabilities
  - harbor sensitive data in public directories
  - harbor sensitive data in public files
- automating the process: googlescan

pick your poison

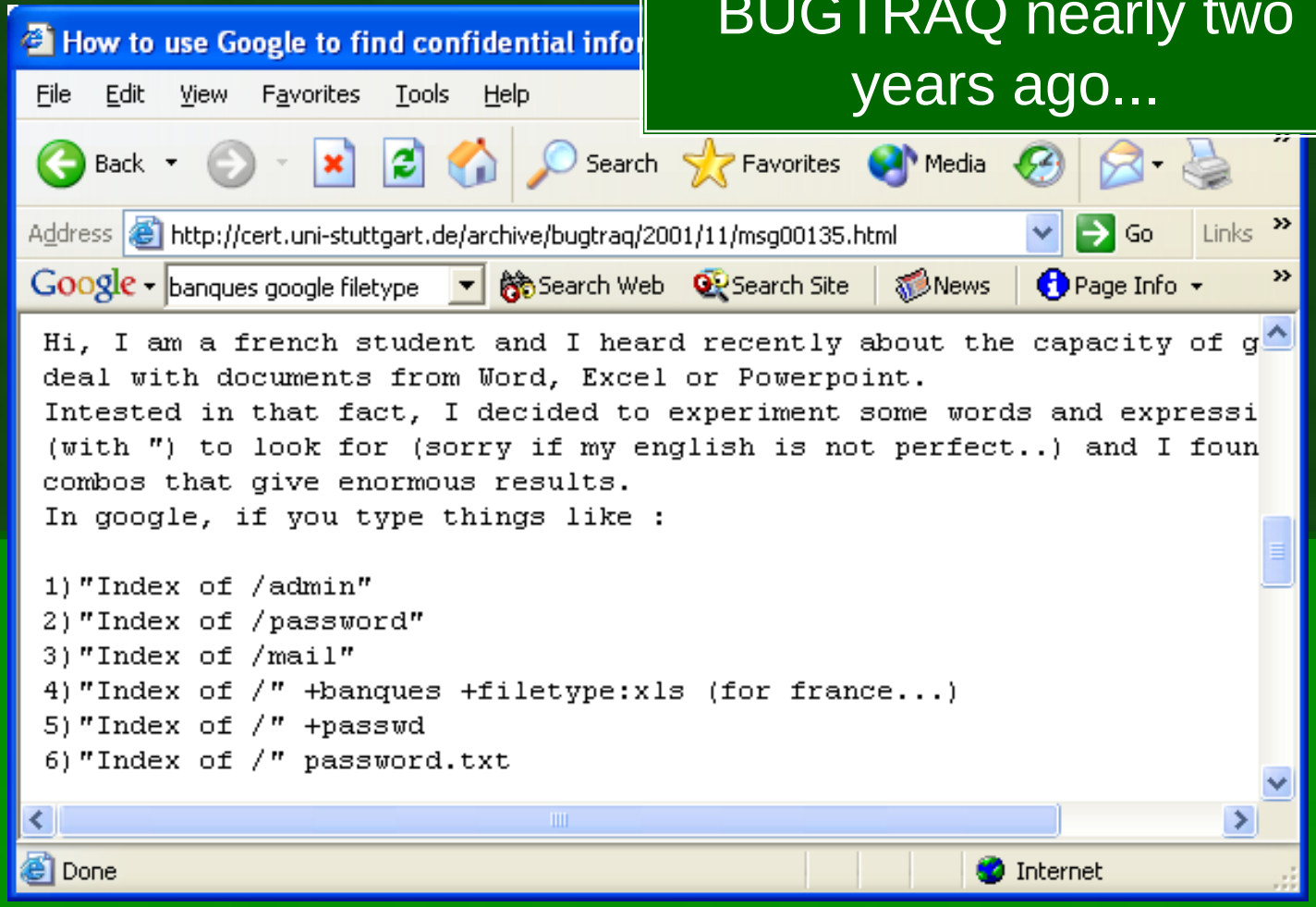
we have certain needs from a search engine:

- advanced search options (not just AND's and OR's)
- browsing down or changed pages (caching)
- instant response (zero-wait)
- document and language translations
- web, news, image and ftp searches

The obvious choice: ***Google***

not new...

Vincent GAILLOT <vgaillot@telecom.insa-lyon.fr> posted this to BUGTRAQ nearly two years ago...



doing interesting stuff

hax0r, “Google hacks,” proxy, auth bypass, finding development sites

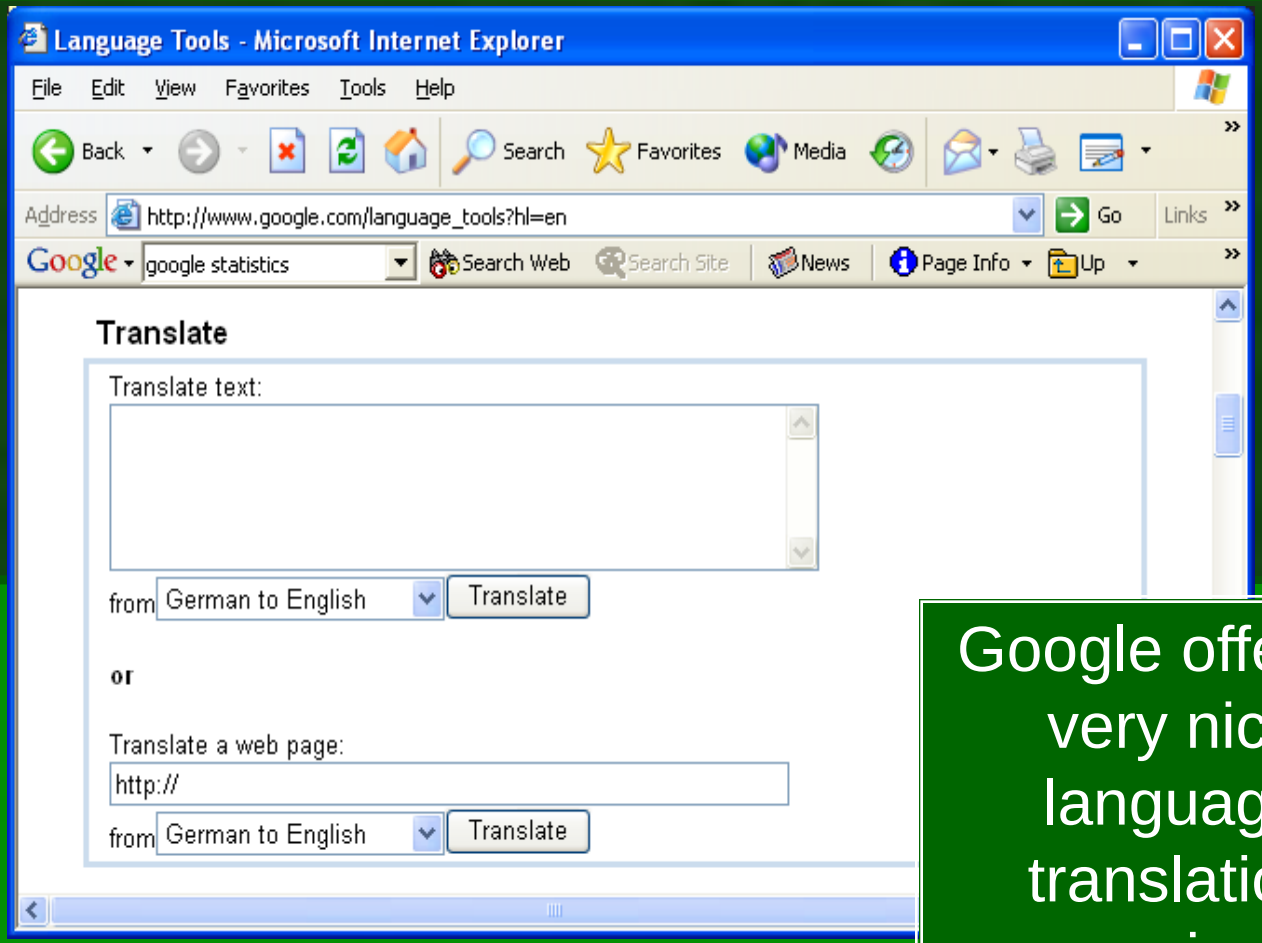
hax0r

for those of us spending way too much time spe@king hax0r...

/misc: “Google Hacks”

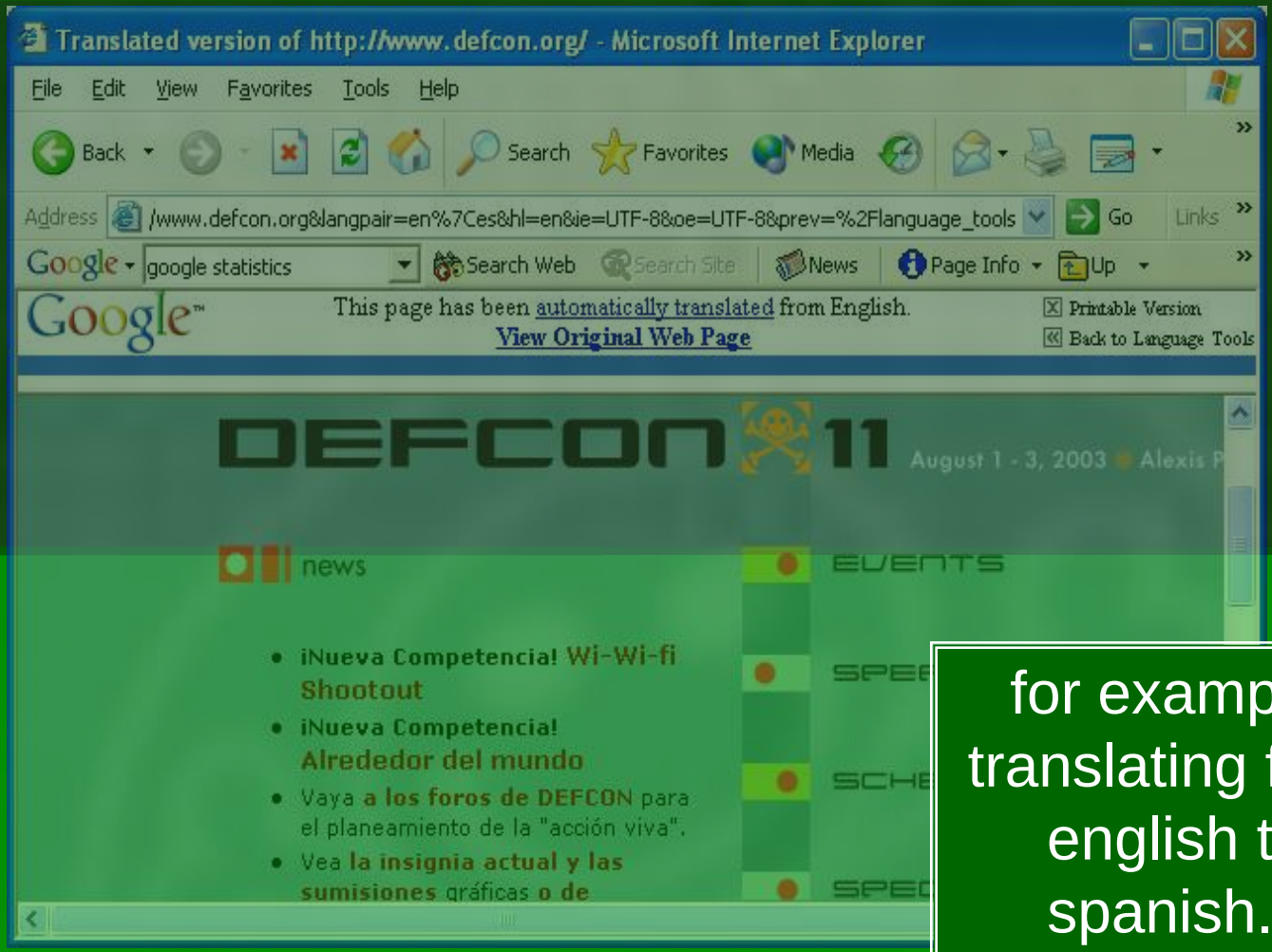
There *is* this book.
And it’s an O’REILLY book.
But it’s not about hacking.
It’s about searching.
I didn’t write it.
Because if I wrote it, it would really be about hacking using Google and that would get both Google and O’REILLY both really upset and then lawyers would get involved, which is never good unless of course the lawyer happens to be Jennifer Granick... =)

proxy



Google offers a very nice language translation service.

proxy



for example, translating from english to spanish...

proxy

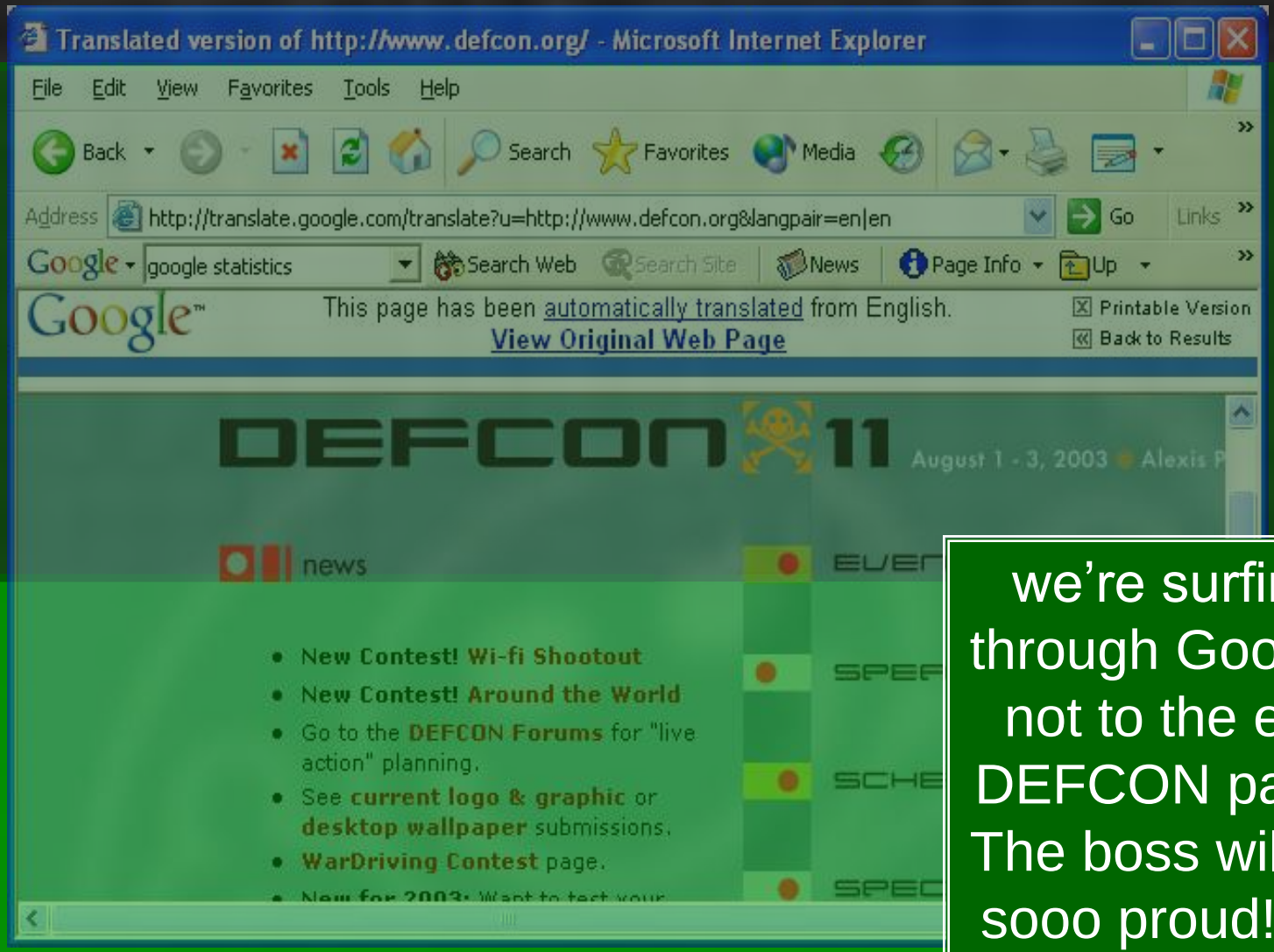
Our english-to-spanish translated Google page is:

http://translate.google.com/translate (main URL)
?u=http://www.defcon.org&langpair=en|es (options)

What happens if we play with the options a bit to provide an english-to-english translation, for example?

http://translate.google.com/translate (main URL)
?u=http://www.defcon.org&langpair=en|en (options)

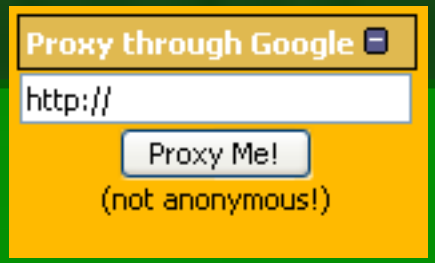
proxy



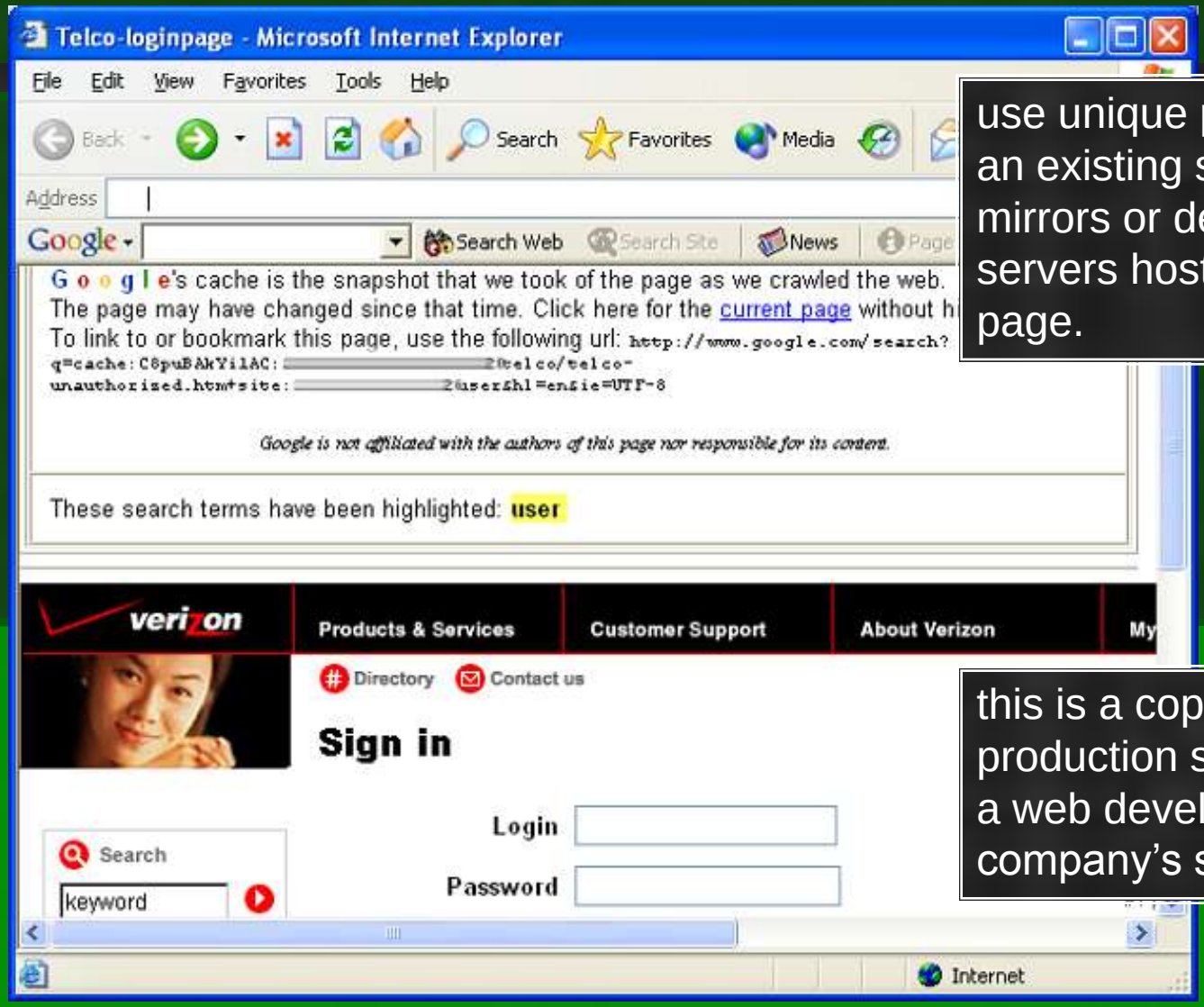
we're surfing through Google, not to the evil DEFCON page. The boss will be sooo proud! 8P

proxy

- Google proxy bouncers
  - http://exploit.wox.org/tools/googleproxy.html
  - http://johnny.ihackstuff.com



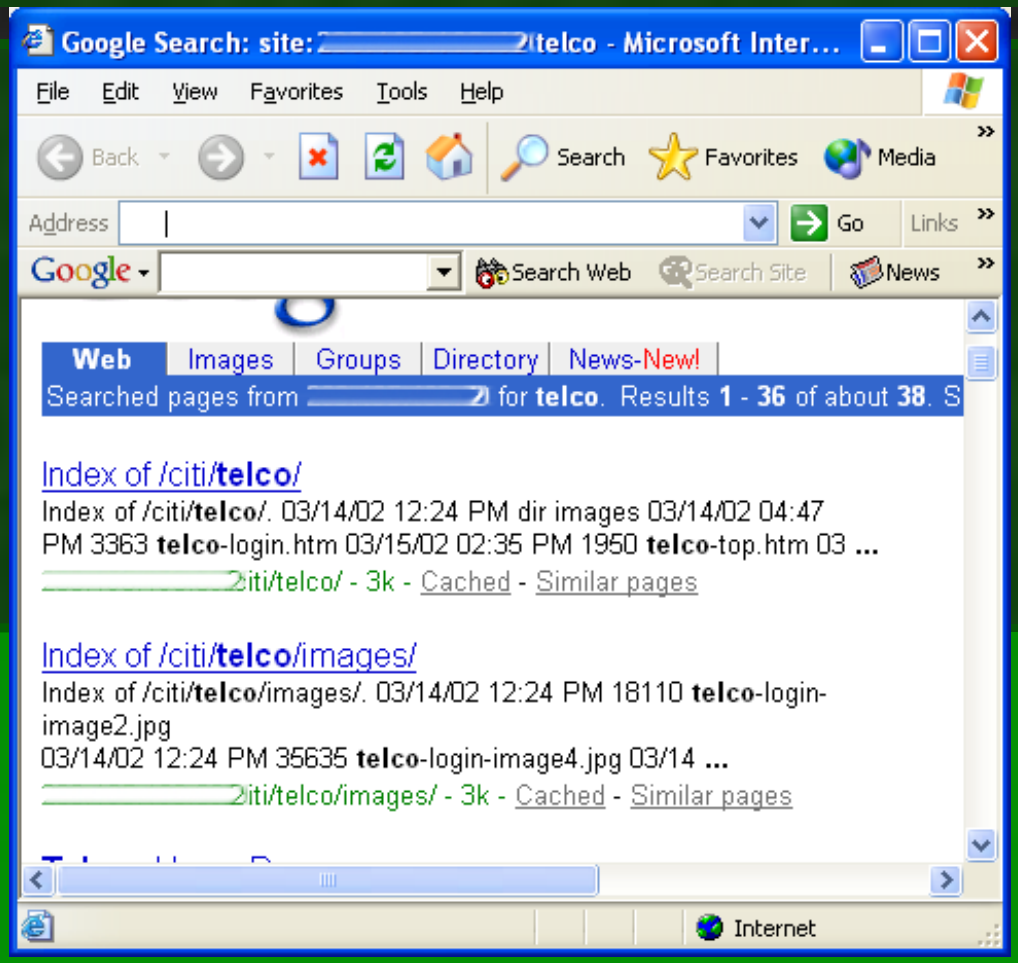
finding development sites



use unique phrases from an existing site to find mirrors or development servers hosting the same page.

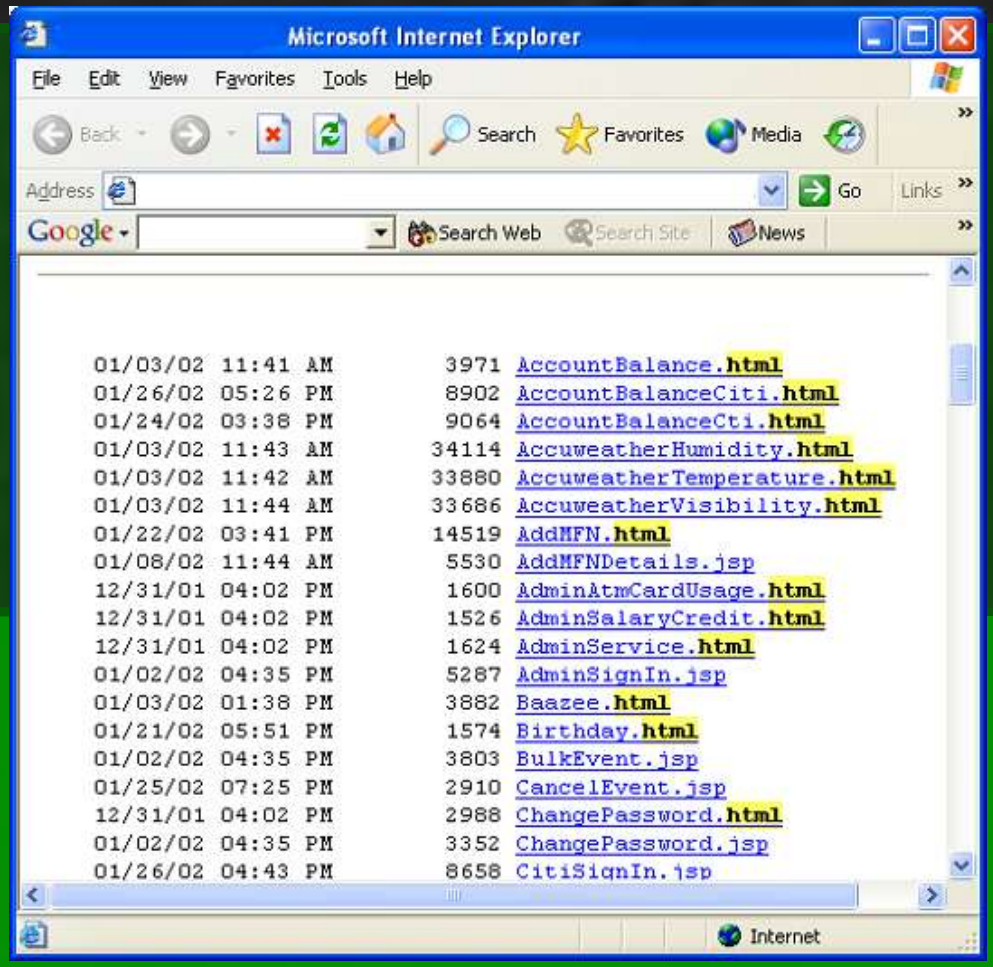
this is a copy of a production site found on a web development company's server...

finding development sites



- troll the development site with another search looking for more files on that server...

finding development sites



- eventually, creative searching can lead to pay dirt: a source code dump dir!

auth bypass

- Let's say an attacker is interested in what's behind www.thin-ice.com, a password protected page:

Authorization Required

auth bypass

- One search gives us insight into the structure of the site:

www.thin-ice.com/templates/search.php3?action=active
Similar pages

www.thin-ice.com/templates/search.php3
Similar pages

www.thin-ice.com/templates/member.php3?action=editprofile
Similar pages

www.thin-ice.com/templates/index.php3?frame=index

auth bypass

- Another search gives a cache link:

Index of /

Index of /. Name Last modified Size Description Parent Directory 03-Jan-2003 12:59 - 0001/ 09-Jan-2001 16:23 - 100/ 15-Oct-2002 19 ...

Google can show you the following information for this URL:

- Show Google's cache of www.thin-ice.com

auth bypass

- Another click takes us to the cached version of the page (no password needed!)

This is G o o g l e's cache of http://www.thin-ice.com/.
G o o g l e's cache is the snapshot that we took of the page as we crawled the web.
The page may have changed since that time. Click here for the current page without highlighting.
To link to or bookmark this page, use the following url: http://www.google.com/search?q=cache:040umZuWia0C:www.thin-ice.com/+&hl=en&ie=UTF-8

Google is not affiliated with the authors of this page nor responsible for its content.

Index of /

Name	Last modified	Size	Description
Parent Directory	03-Jan-2003 12:59	-	
0001/	09-Jan-2001 16:23	-	

auth bypass

- One more click to the really interesting stuff... site source code!

```
-rw-r--r--    1 root             wheel        0 Nov 27 15:46 allphpfiles
-rw-rw-r--    1 root             www       7695 Sep  4  2001 announcement.php3
-rw-rw-r--    1 root             www     993280 Sep  4  2001 blazeboard-php3_0.98.tar
-rwxrwxr-x    1 root             www        881 Jul 17  2001 configuration.php
-rwxrwxr-x    1 root             www       2870 Jul 17  2001 countReplies.php
-rw-rw-r--    1 root             www        819 Sep  4  2001 counter.php3
-rwxrwxr-x    1 root             www       4051 Jul 17  2001 dbSetup.php
-rwxrwxr-x    1 root             www       4852 Jul 17  2001 default.php
-rwxrwxr-x    1 root             www       4053 Jul 17  2001 expand.php
-rw-rw-r--    1 root             www       1742 Sep  4  2001 faq.php3
```

*this site was notified and secured before making this public. sorry, kids ;-)

evil searching: the basics

tools of the trade

Google search syntax

Tossing Google around requires a firm grasp of the basics.

Many of the details can be found here:

http://www.google.com/apis/reference.html

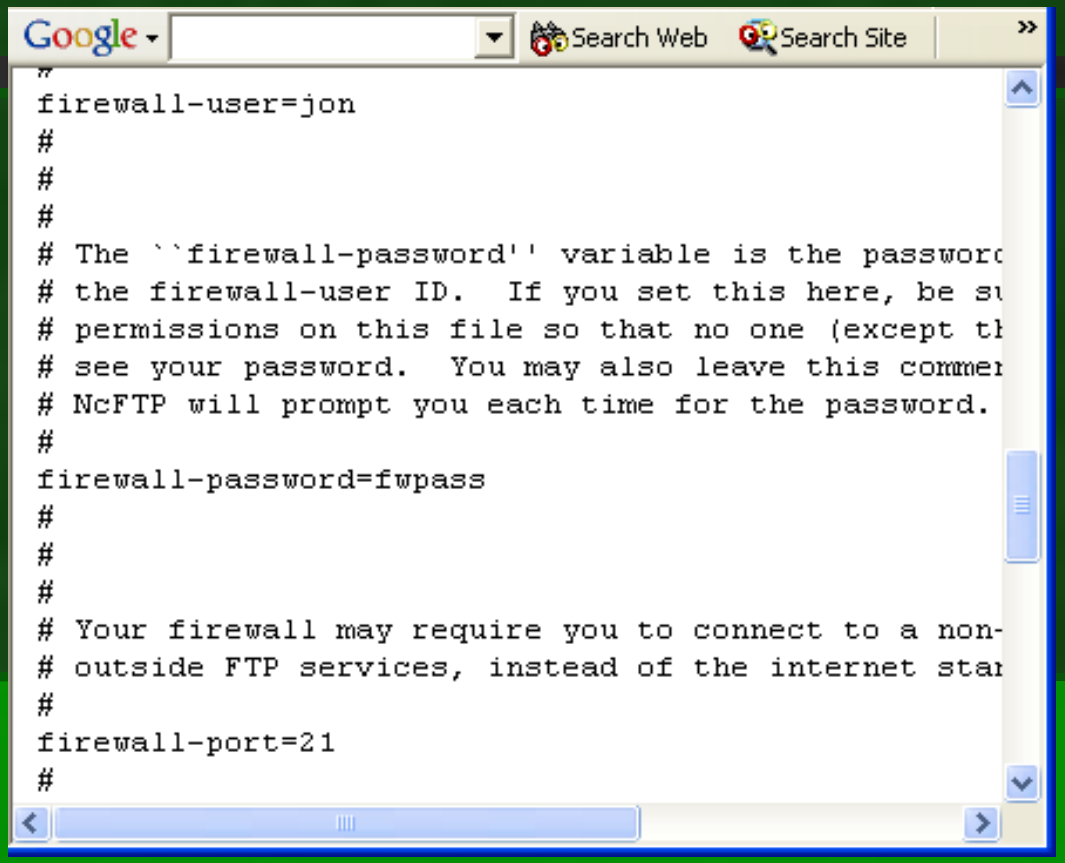
simple word search

A simple search...

simple word search

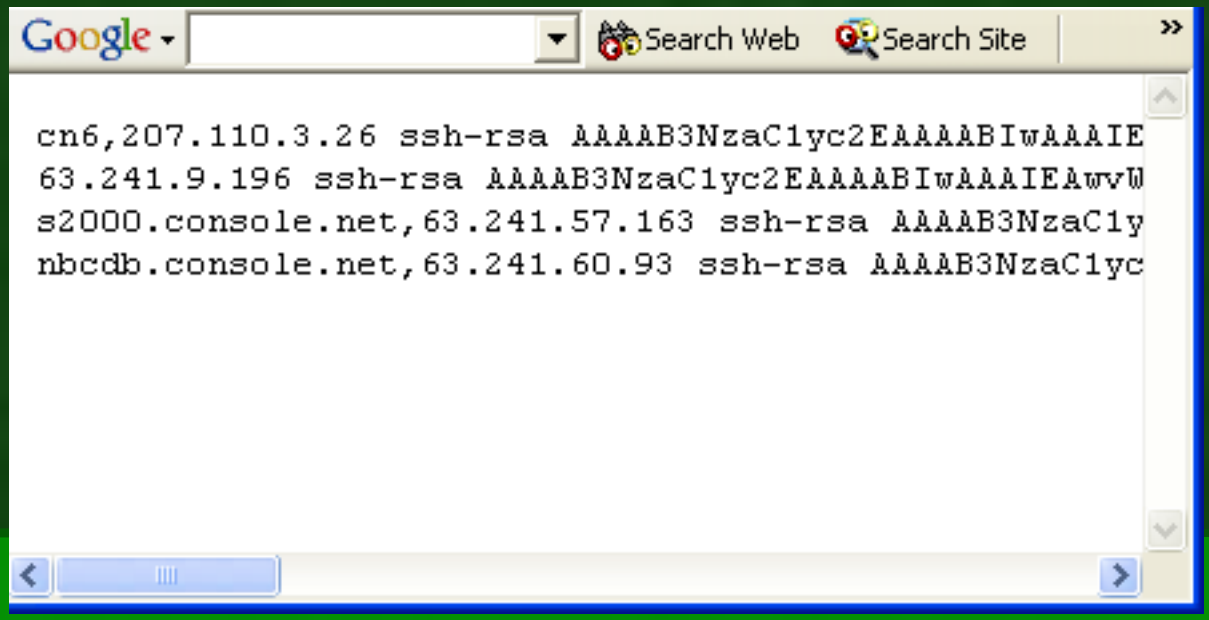
...can return amazing results. This is the contents of a live .bash_history file!

simple word search



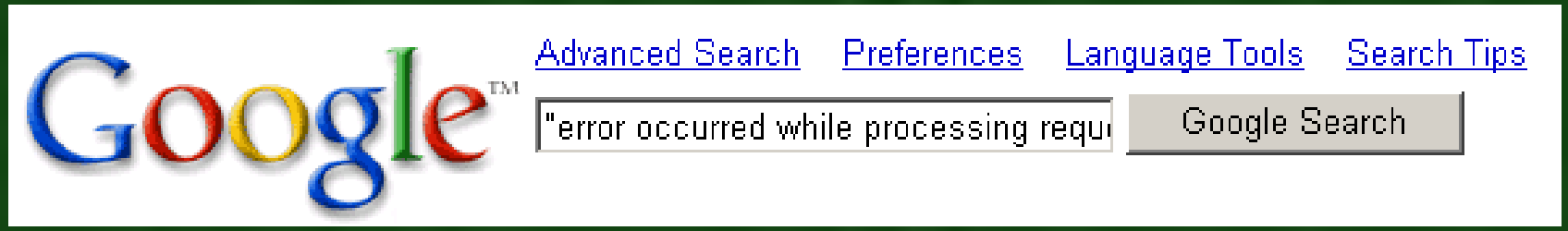
Crawling around on the same web site reveals a firewall configuration file complete with a username and password...

simple word search



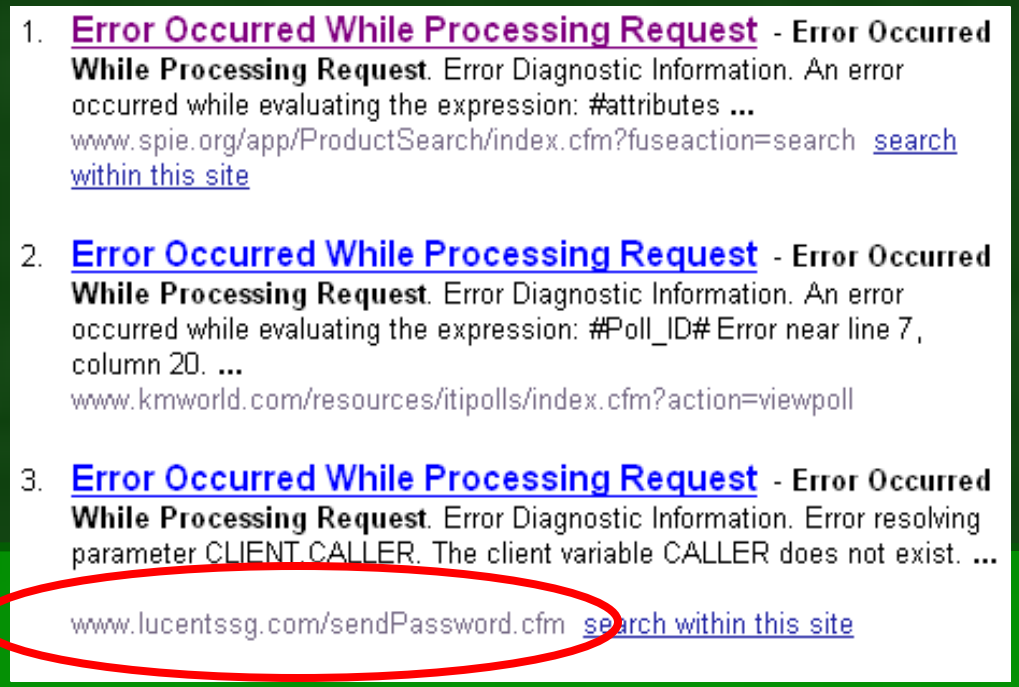
...as well as an ssh known hosts file!

simple phrase search



Creativity with search phrases (note the use of quotes)…

simple phrase search



1. **Error Occurred While Processing Request** - **Error Occurred While Processing Request**. Error Diagnostic Information. An error occurred while evaluating the expression: #attributes ...
   www.spie.org/app/ProductSearch/index.cfm?fuseaction=search search within this site

2. **Error Occurred While Processing Request** - **Error Occurred While Processing Request**. Error Diagnostic Information. An error occurred while evaluating the expression: #Poll_ID# Error near line 7, column 20. ...
   www.kmworld.com/resources/itipolls/index.cfm?action=viewpoll

3. **Error Occurred While Processing Request** - **Error Occurred While Processing Request**. Error Diagnostic Information. Error resolving parameter CLIENT.CALLER. The client variable CALLER does not exist. ...
   www.lucentssg.com/sendPassword.cfm search within this site

...can reveal interesting tidbits like this Cold Fusion error message.

simple phrase search

(Error messages can be very revealing.)

Error Diagnostic Information

Error resolving parameter **CLIENT.CALLER**

The client variable CALLER does not exist. The cause of this error is very likely on

1. The name of the client variable has been misspelled.
2. The client variable has not yet been created or has timed out.

To set default values for client variables you should use the CFPARAM tag (e.g. <CFPARAM NAME="Client.Color" DEFAULT="Red">)

The error occurred while evaluating the expression:

```
"#client.caller#"
```

Template: C:\InetPub\wwwroot\lucent production\sendPassword.cfm

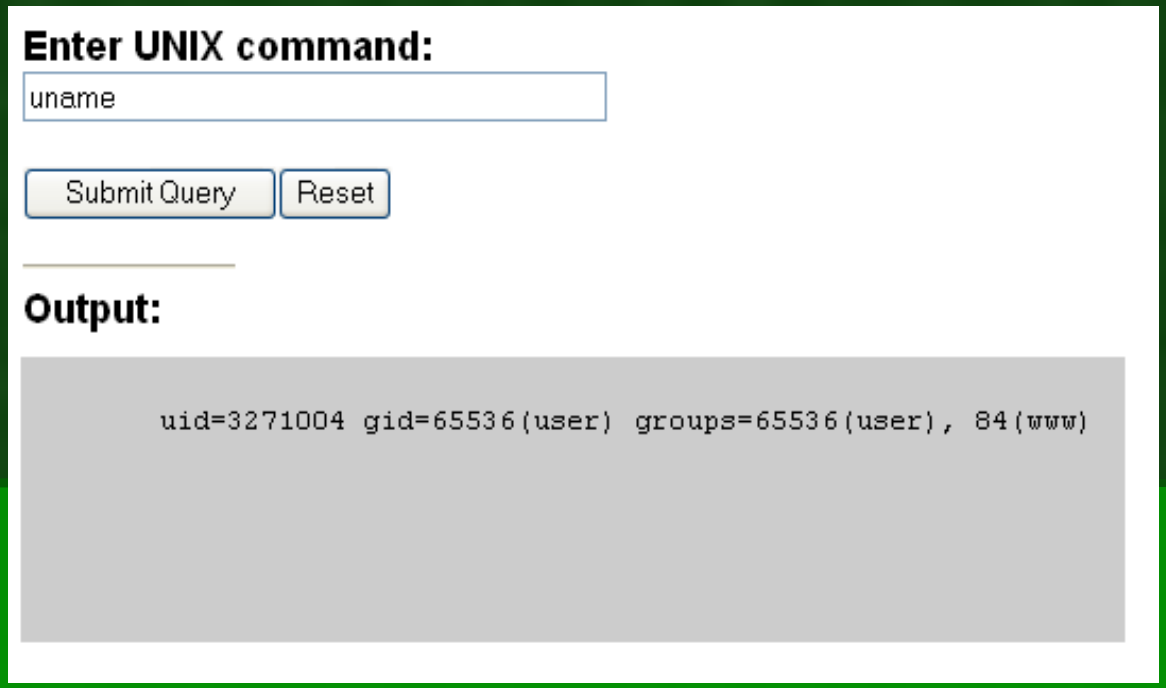
The error occurred while processing an element with a general identifier of (CFLOCATION), occupying document position (22:1) to (22:48).

Date/Time: 07/02/03 11:11:09
Browser: Mozilla/4.0 (compatib IE 6.0; Windows NT 5.1; .NET CLR 1.0.3705)
Remote Address:
HTTP Referer: http://www.g /search?hl=en&lr=&ie=UTF-8&oe=UTF-8&q=%22Error+Occurred+While+Proc g+Request%22+site%3Alucentssg.com
Template: C:\InetPub\wwwroot\lucent production\sendPassword.cfm

simple phrase search II

Sometimes the most idiotic searches ("enter UNIX command")...

simple phrase search II



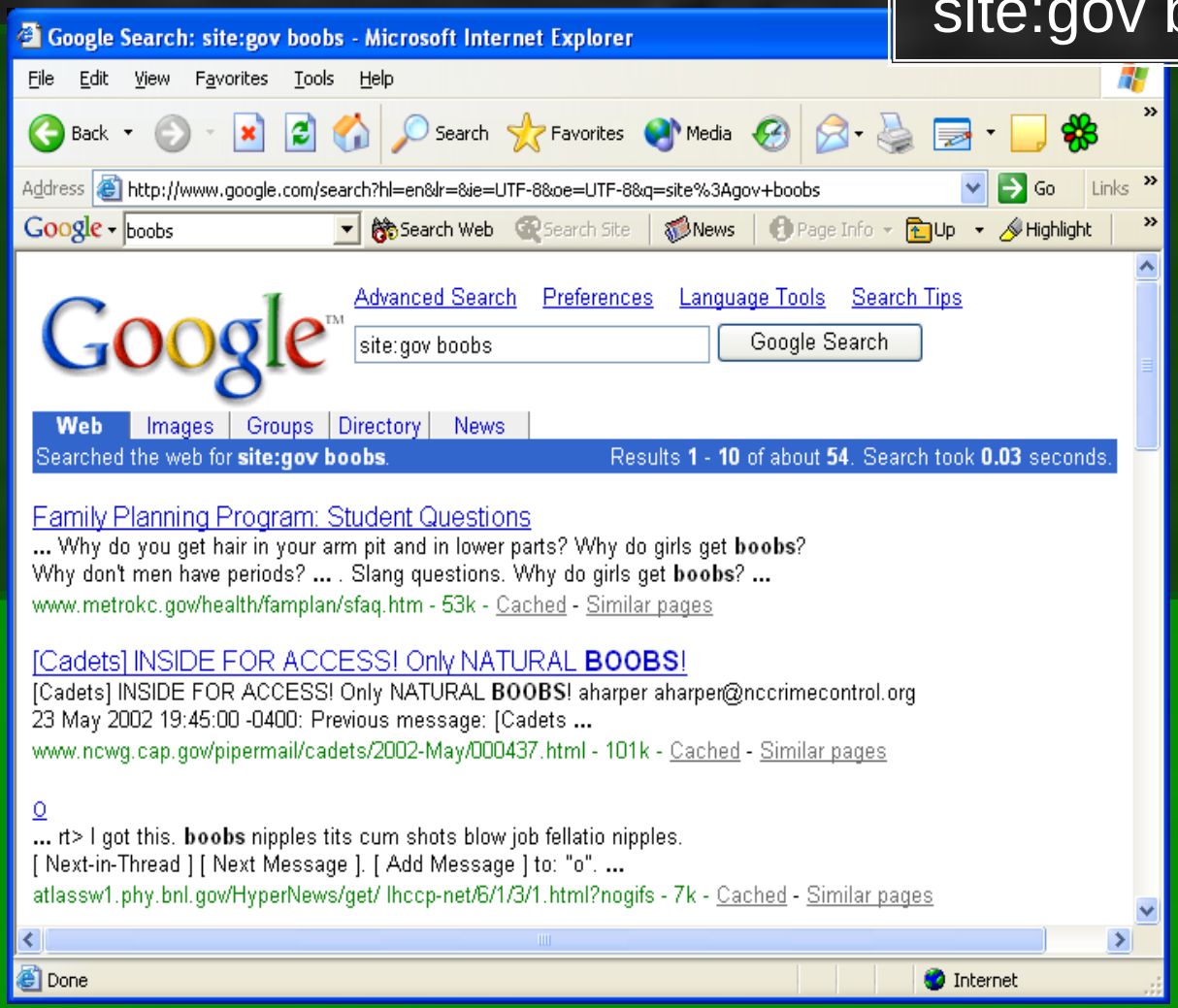
...can be the most rewarding!

special characters

symbol	use
+ (plus)	AND, force use
- (dash)	NOT (when used outside quotes)
. (period)	any character
- (dash)	space (when used in quotes)
* (asterisk)	wildcard word (when used in quotes)

site: site-specific search

site:gov boobs

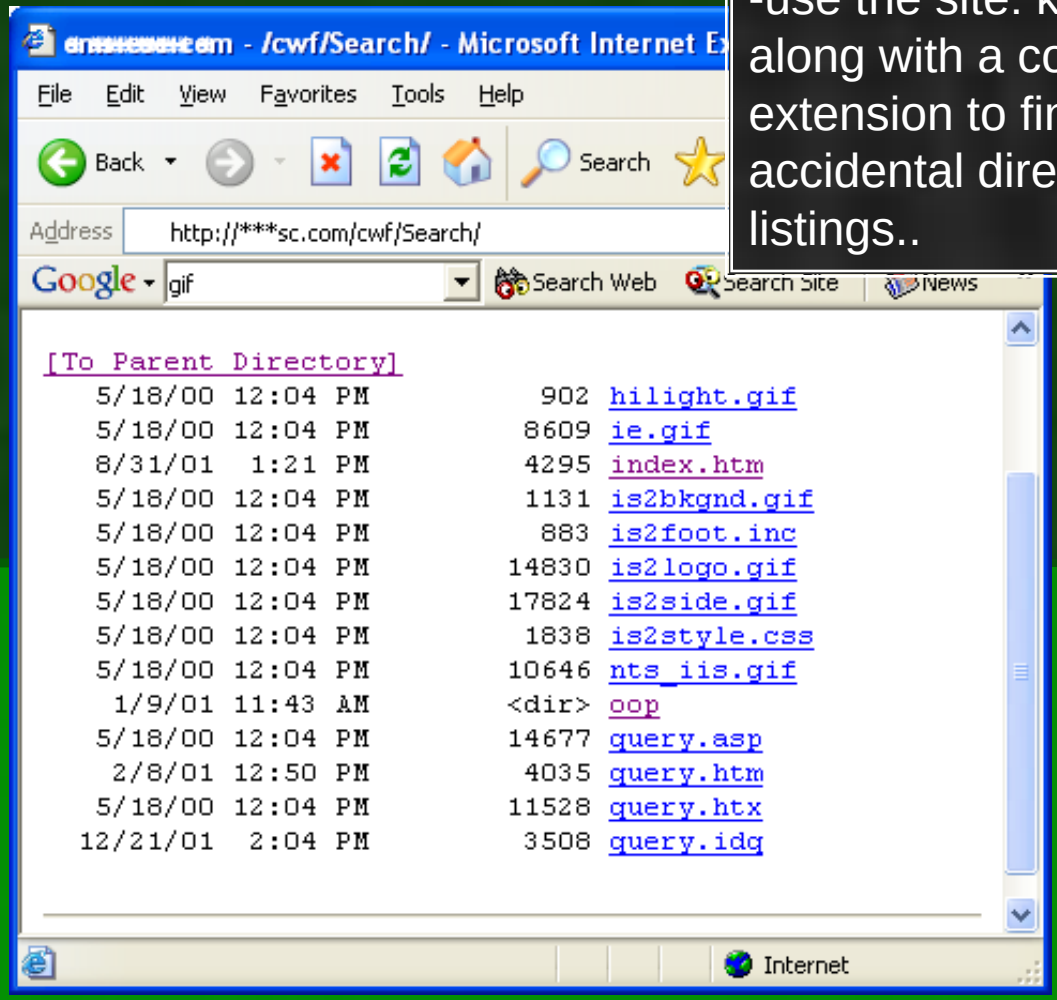


site: crawling

site:defcon.org defcon

-use the site: keyword along with the site name for a quick list of potential servers and directories

site: crawling



-use the site: keyword along with a common file extension to find accidental directory listings..

Date Searching

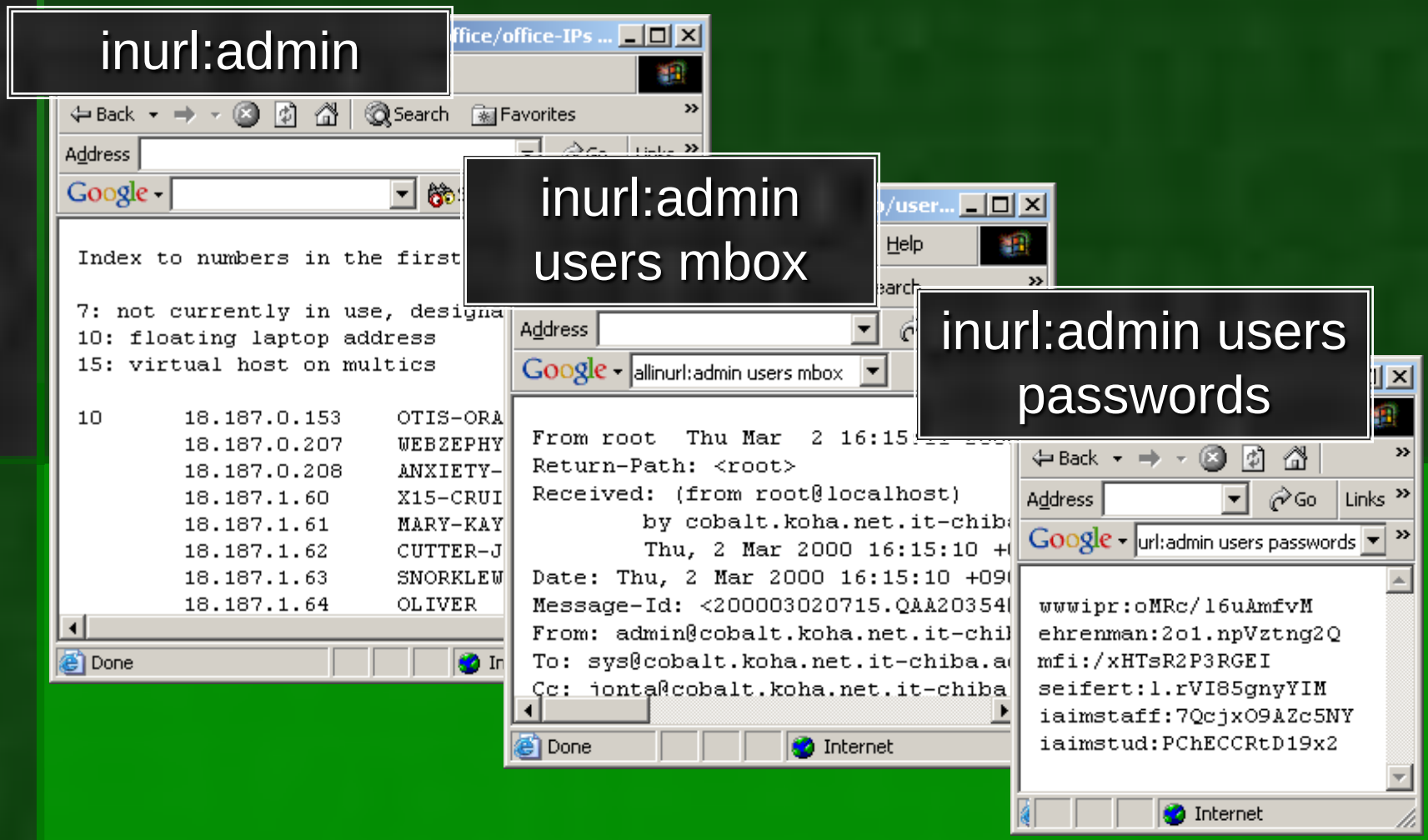
• Date Restricted Search	• Star Wars daterange:2452122-2452234	• If you want to limit your results to documents that were published within a specific date range, then you can use the “daterange: “ query term to accomplish this. The “daterange:” query term must be in the following format: • daterange:<start_date>-<end date> where • <start_date> = Julian date indicating the start of the date range <end_date> = Julian date indicating the end of the date range • The Julian date is calculated by the number of days since January 1, 4713 BC. For example, the Julian date for August 1, 2001 is 2452122.

Title searching

Title Search (term)	intitle:Google search	If you prepend "intitle:" to a query term, Google search restricts the results to documents containing that word in the title. Note there can be no space between the "intitle:" and the following word. Note: Putting "intitle:" in front of every word in your query is equivalent to putting "allintitle:" at the front of your query.
Title Search (all)	allintitle: Google search	Starting a query with the term "allintitle:" restricts the results to those with all of the query words in the title.

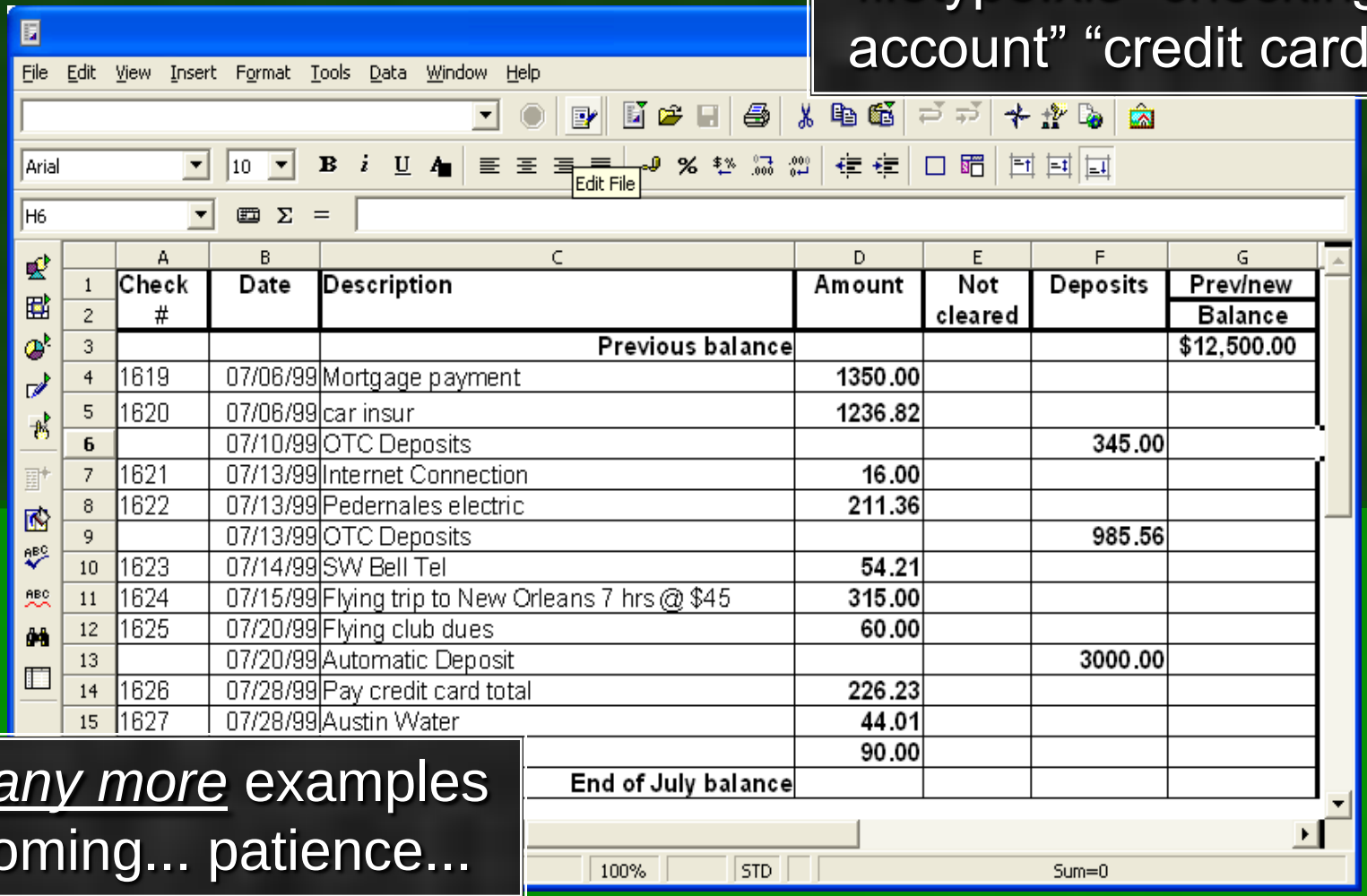
INURL: URL Searches

inurl: find the search term within the URL



filetype:

filetype:xls “checking account” “credit card”



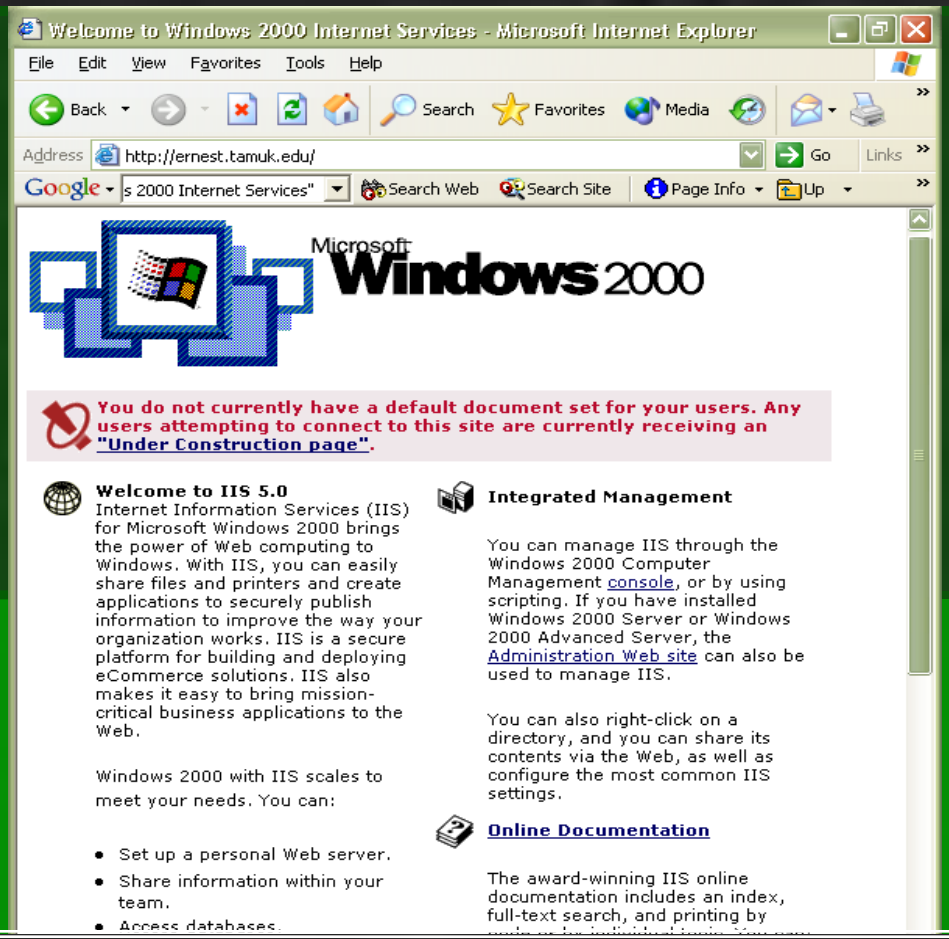
	A	B	C	D	E	F	G
1	Check #	Date	Description	Amount	Not cleared	Deposits	Prev/new Balance
3			Previous balance				$12,500.00
4	1619	07/06/99	Mortgage payment	1350.00			
5	1620	07/06/99	car insur	1236.82			
6		07/10/99	OTC Deposits			345.00	
7	1621	07/13/99	Internet Connection	16.00			
8	1622	07/13/99	Pedernales electric	211.36			
9		07/13/99	OTC Deposits			985.56	
10	1623	07/14/99	SW Bell Tel	54.21			
11	1624	07/15/99	Flying trip to New Orleans 7 hrs @ $45	315.00			
12	1625	07/20/99	Flying club dues	60.00			
13		07/20/99	Automatic Deposit			3000.00	
14	1626	07/28/99	Pay credit card total	226.23			
15	1627	07/28/99	Austin Water	44.01			
				90.00			
			End of July balance				

many more examples coming... patience...

finding interesting stuff

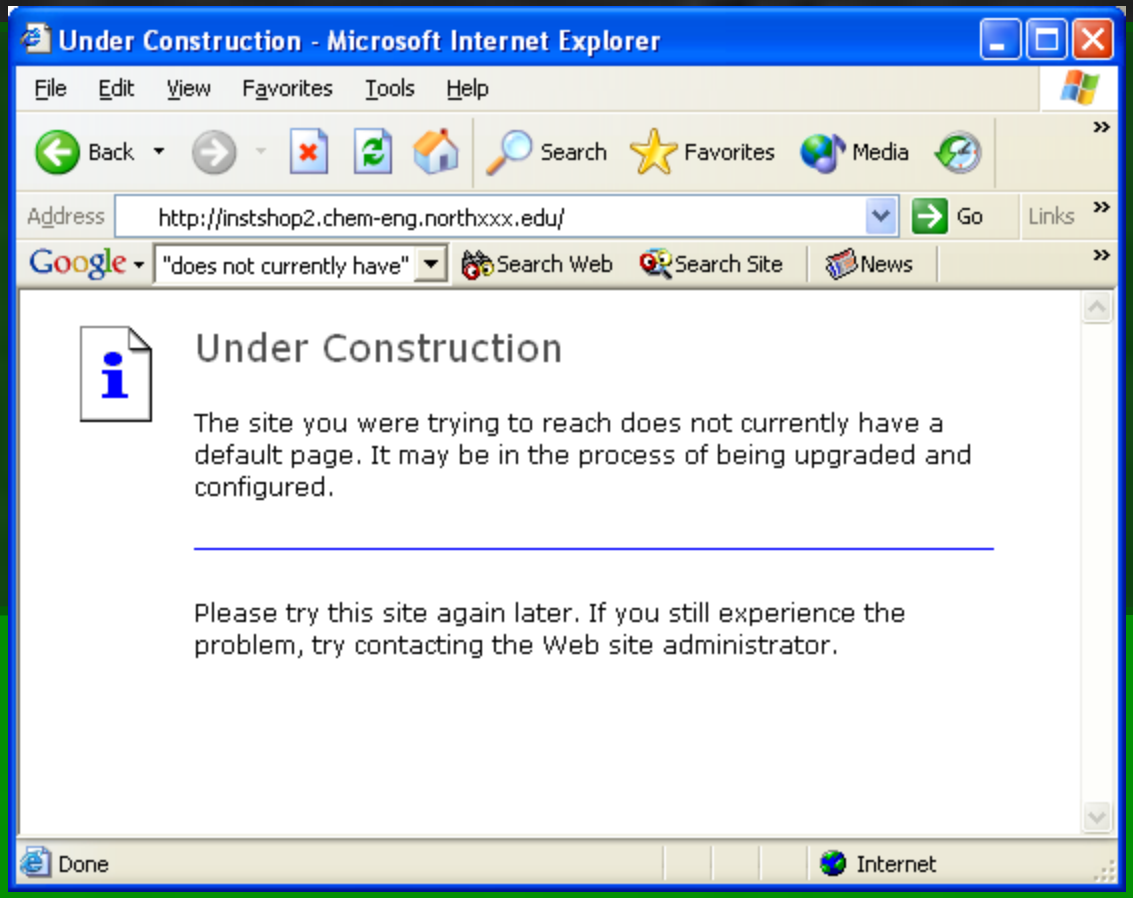
finding OS and web server versions

Windows-based default server



intitle:"Welcome to Windows 2000 Internet Services"

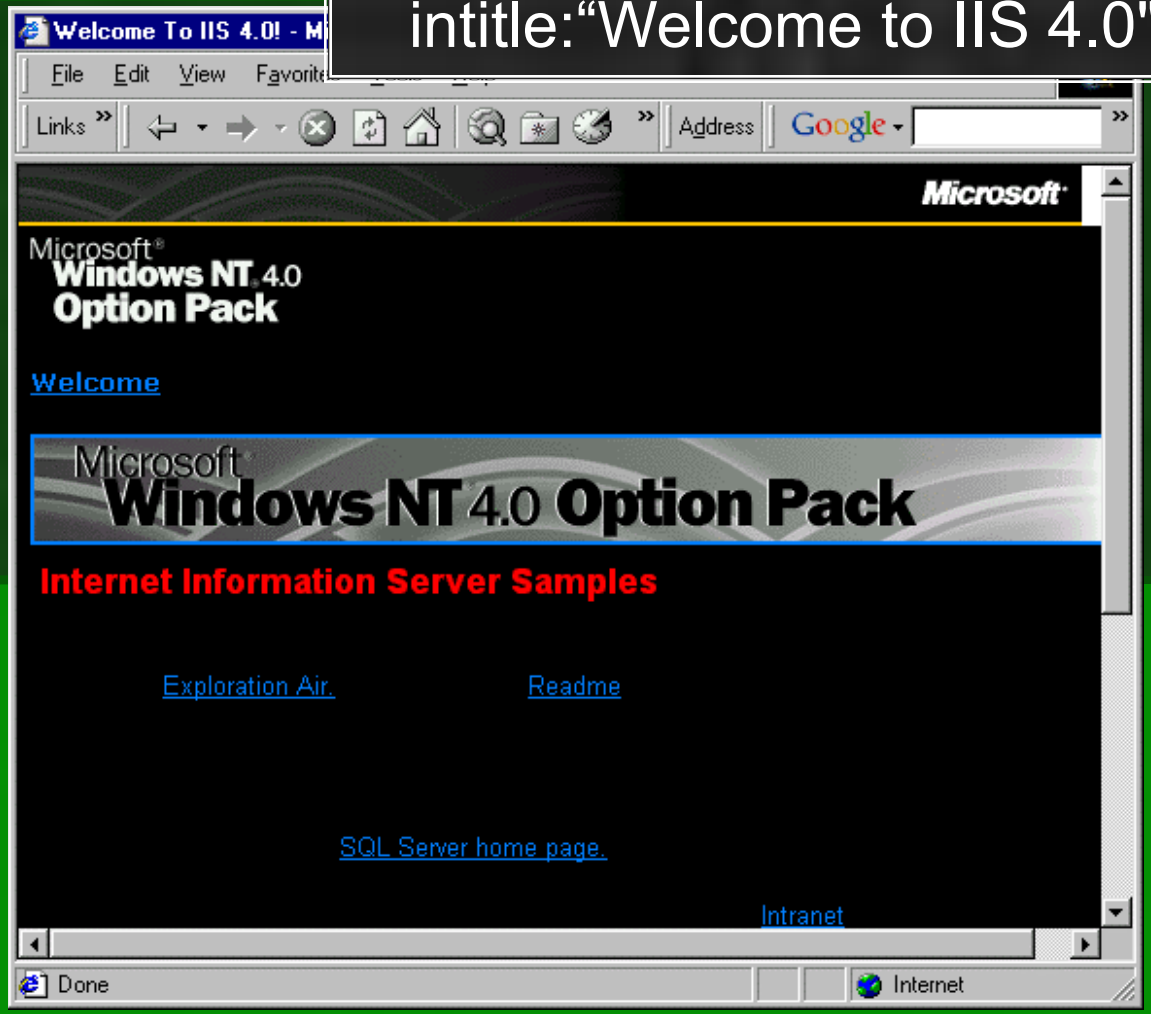
Windows-based default server



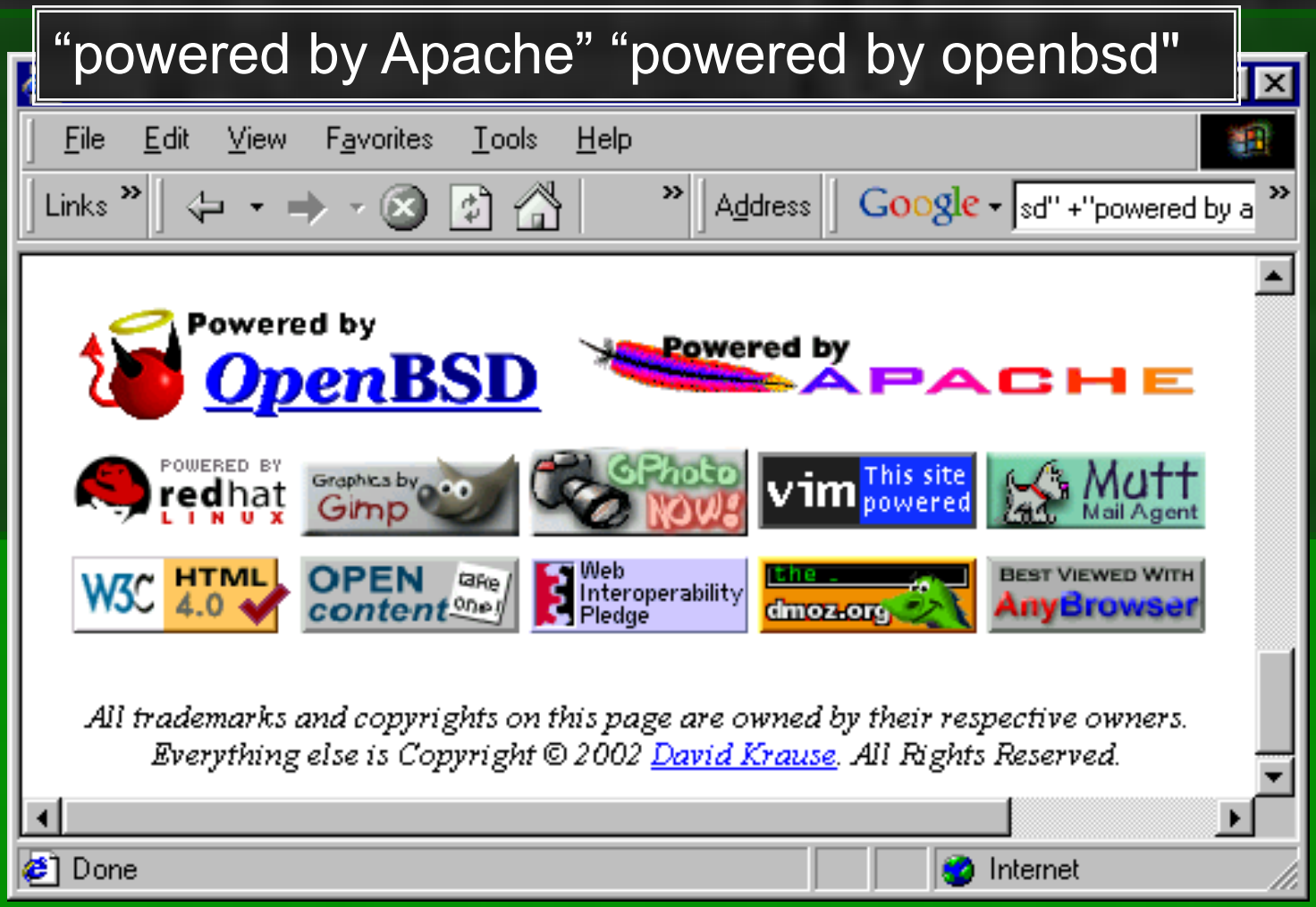
intitle:"Under construction" "does not currently have"

Windows NT 4.0

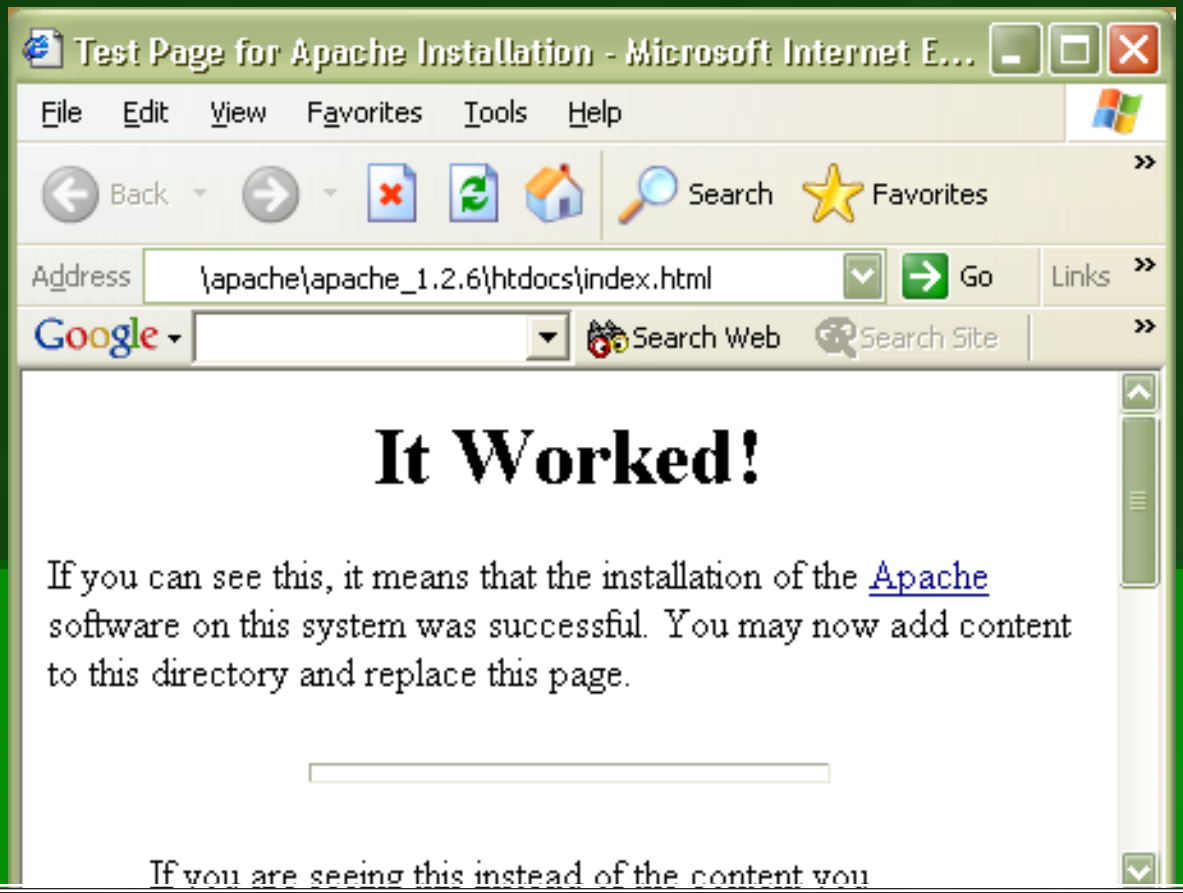
intitle:“Welcome to IIS 4.0"



OpenBSD/Apache (scalp=)

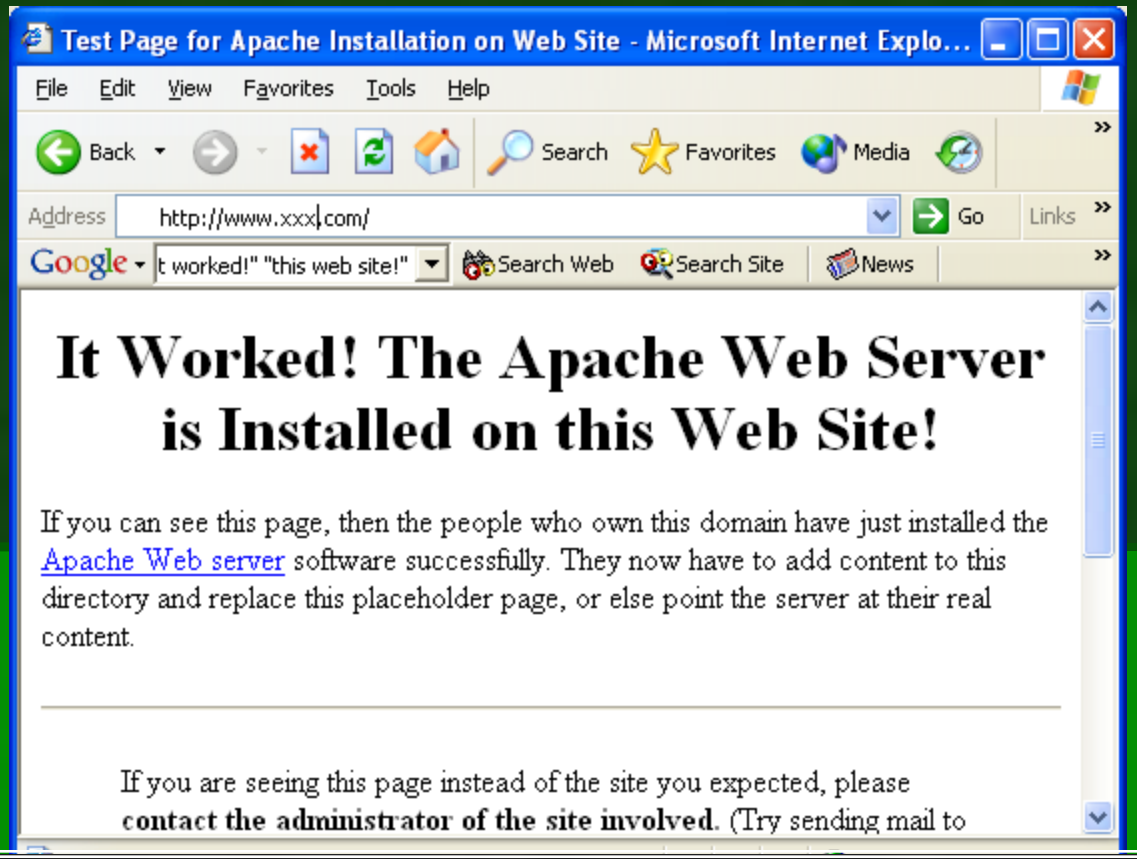


Apache 1.2.6



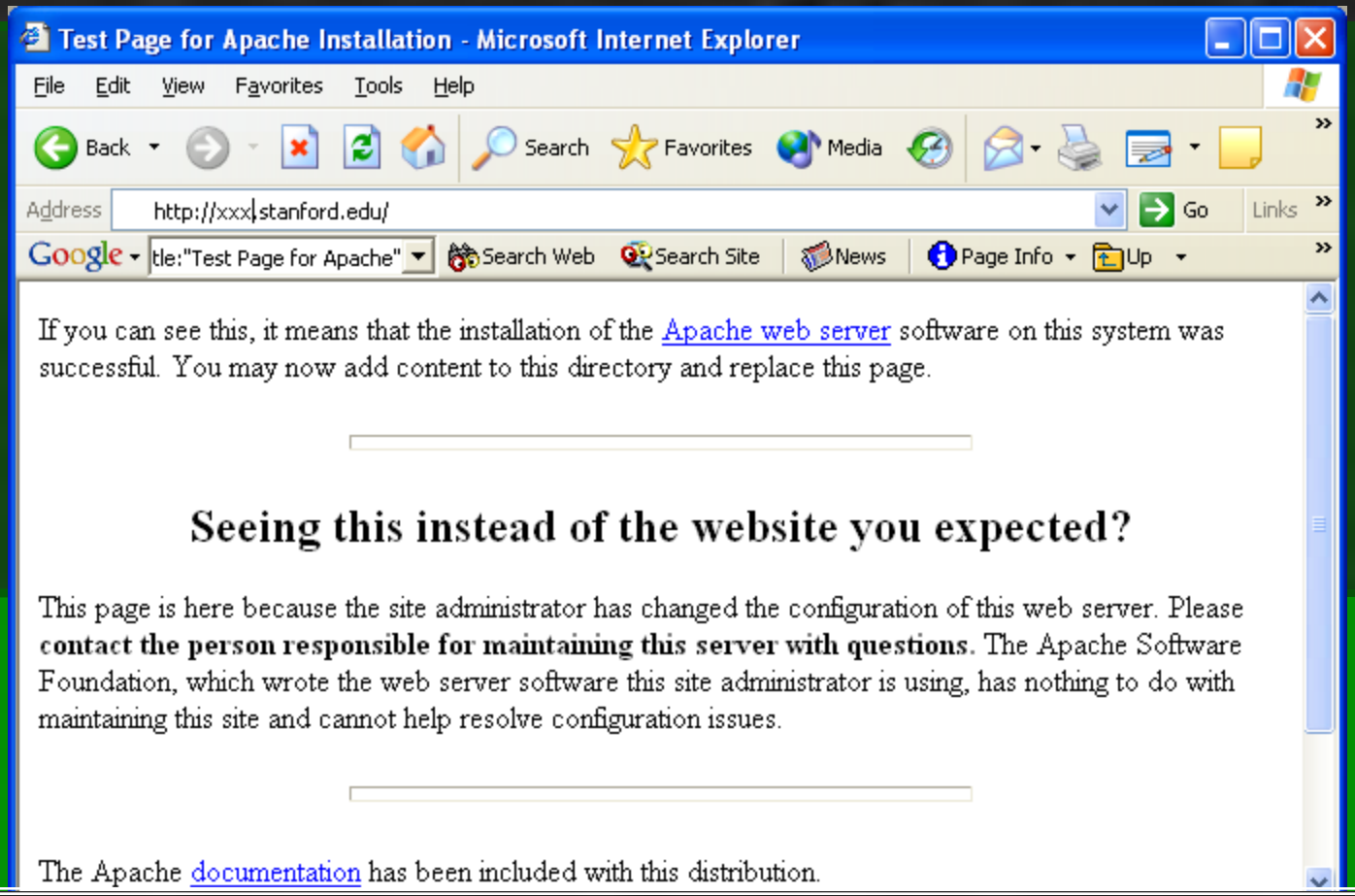
Intitle:"Test Page for Apache" "It Worked!"

Apache 1.3.0 – 1.3.9



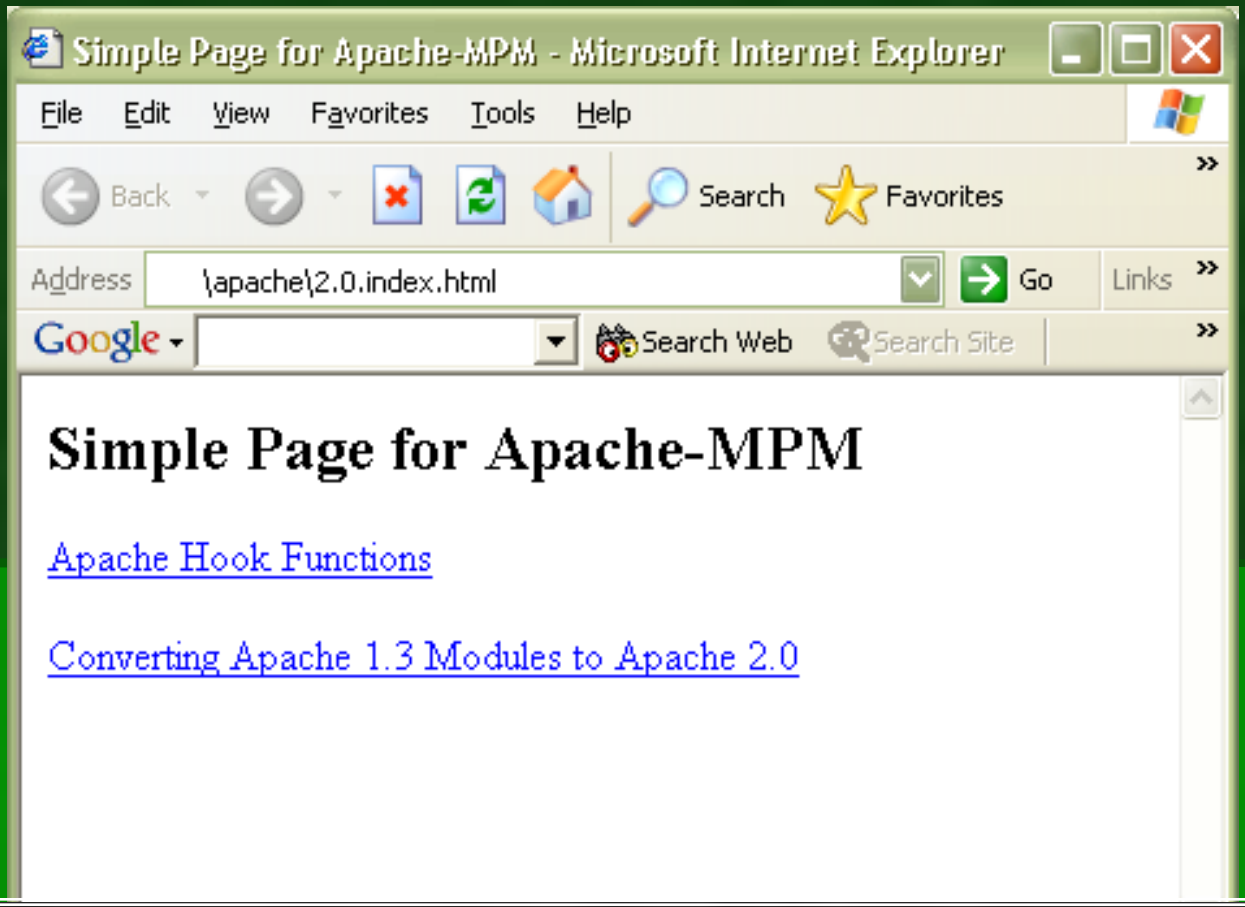
Intitle:"Test Page for Apache" "It worked!" "this web site!"

Apache 1.3.11 - 1.3.26



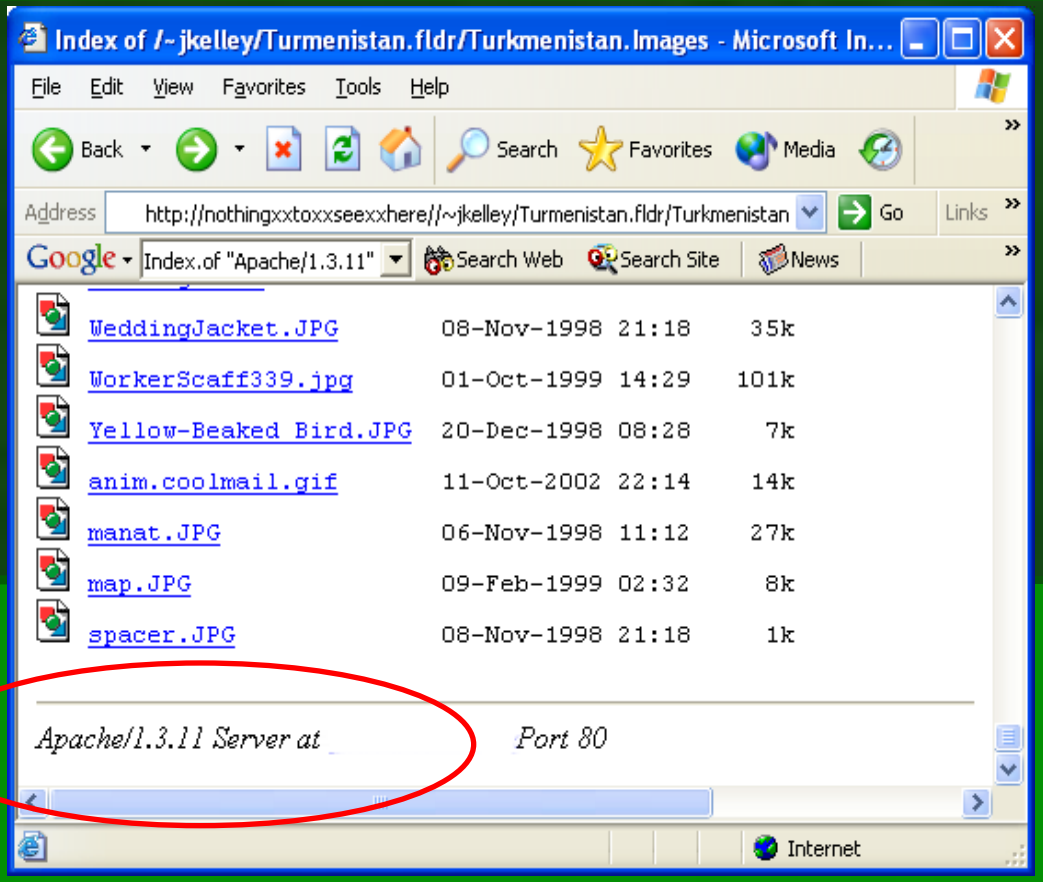
"seeing this instead" intitle:"Test Page for Apache"

Apache 2.0



Intitle:"Simple page for Apache" "Apache Hook Functions"

Directory Info Gathering



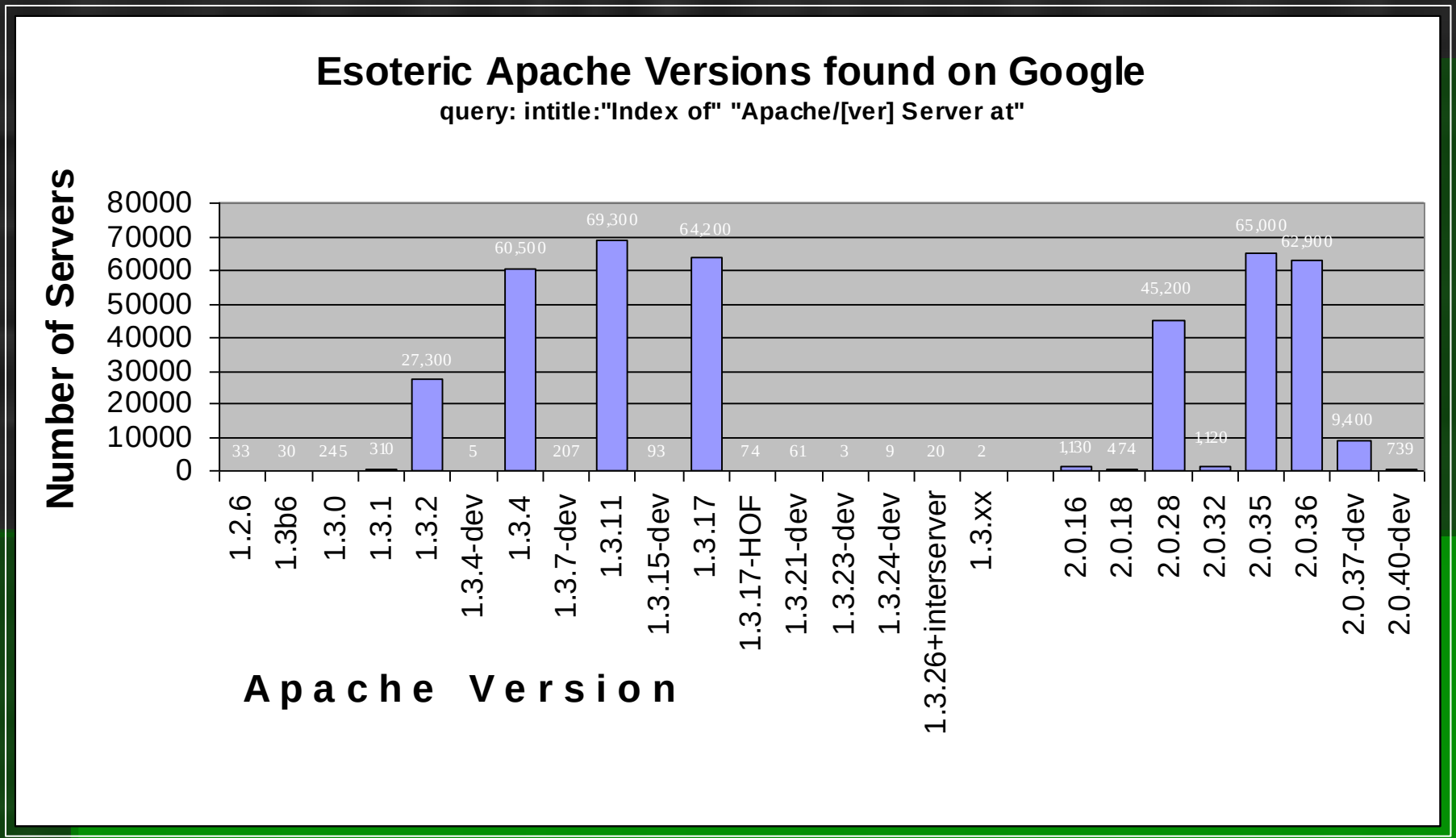
- Some servers, like Apache, generate a server version tag...

Apache Version Info

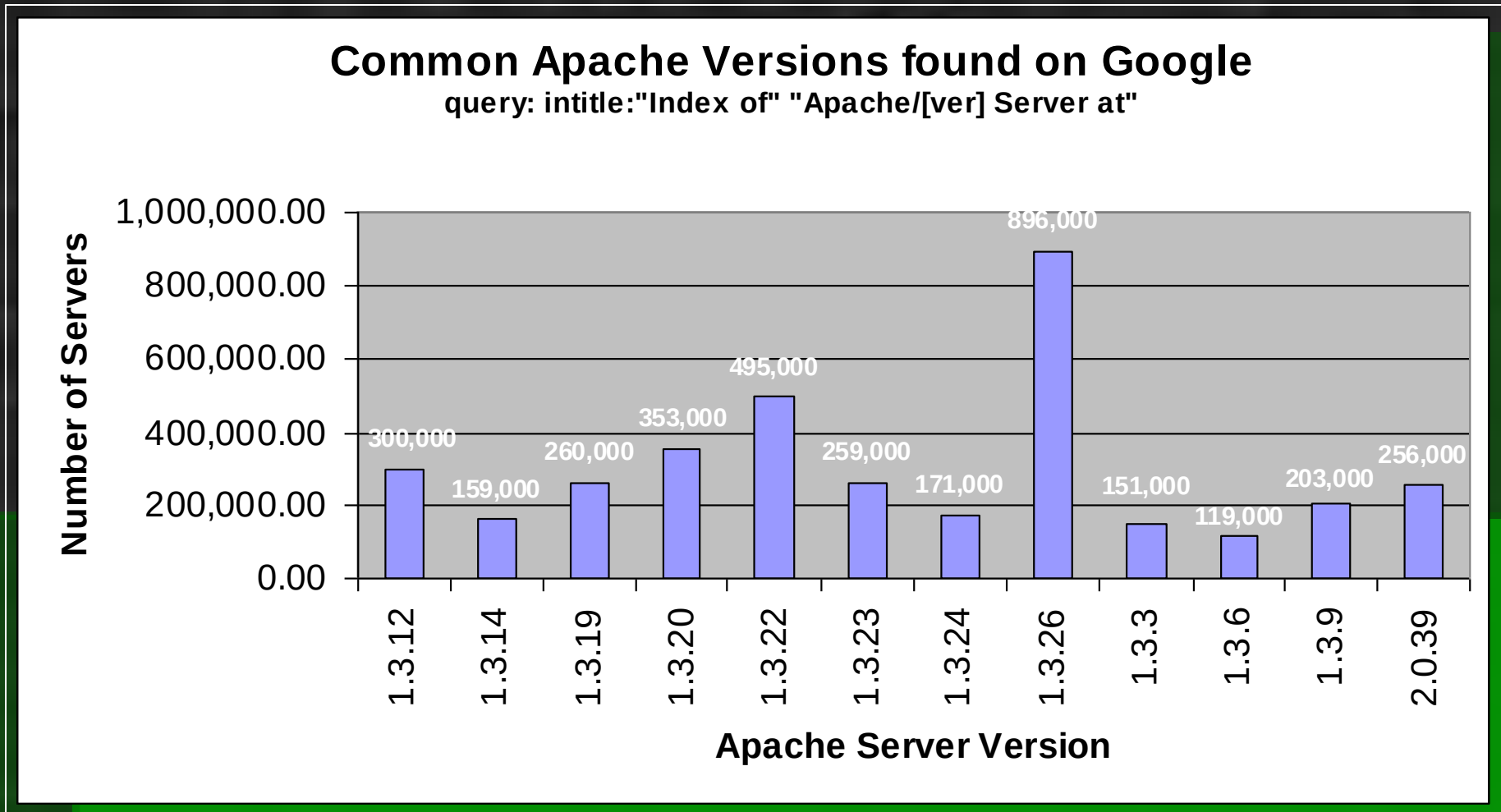
•Apache Version	•Number of Servers
• 1.3.6	• 119,000.00
• 1.3.3	• 151,000.00
• 1.3.14	• 159,000.00
• 1.3.24	• 171,000.00
• 1.3.9	• 203,000.00
• 2.0.39	• 256,000.00
• 1.3.23	• 259,000.00
• 1.3.19	• 260,000.00
• 1.3.12	• 300,000.00
	• 353,00

...which we can harvest for some quick stats...

Weird Apache Versions



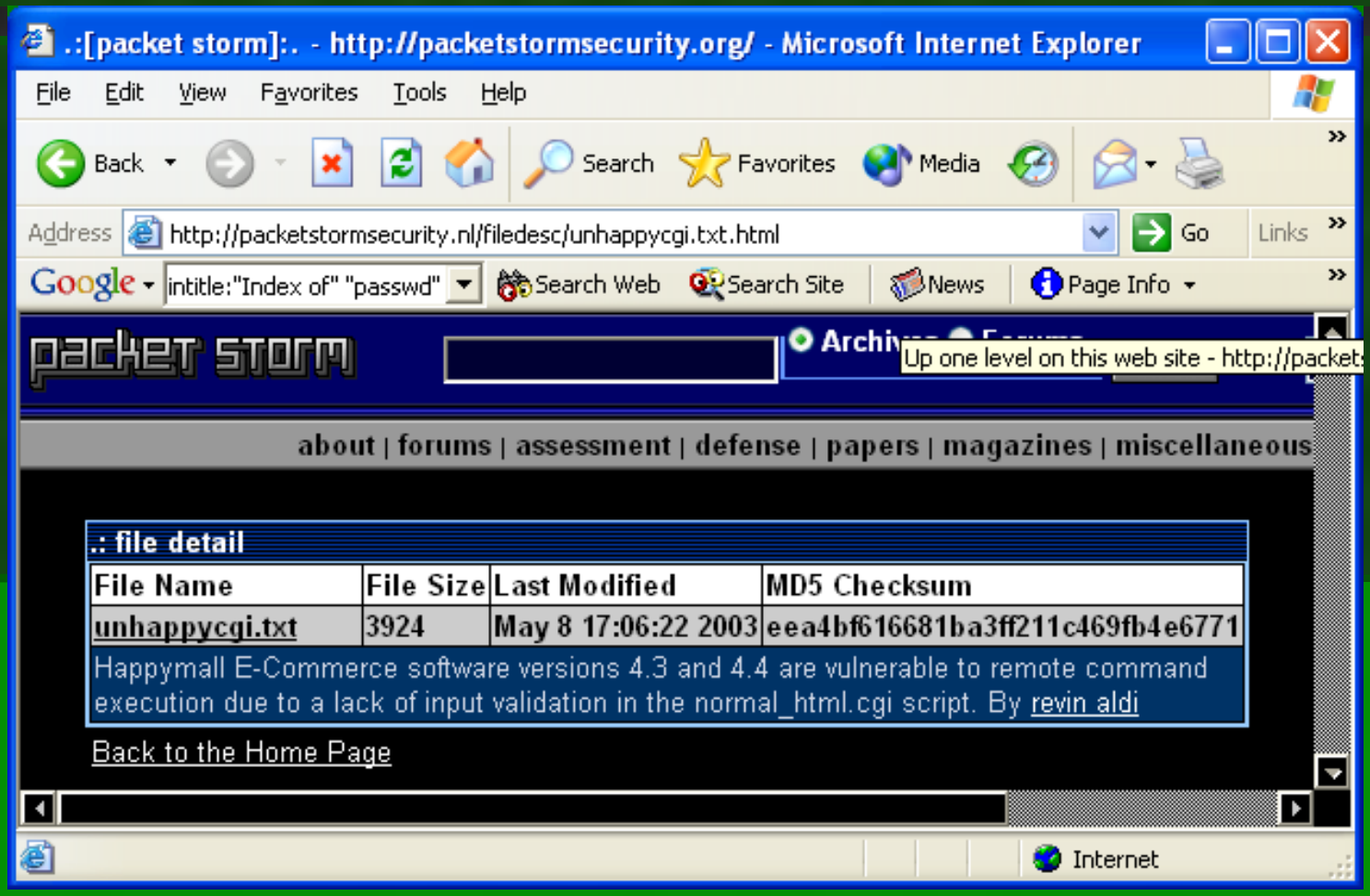
Common Apache Versions



vulnerability trolling

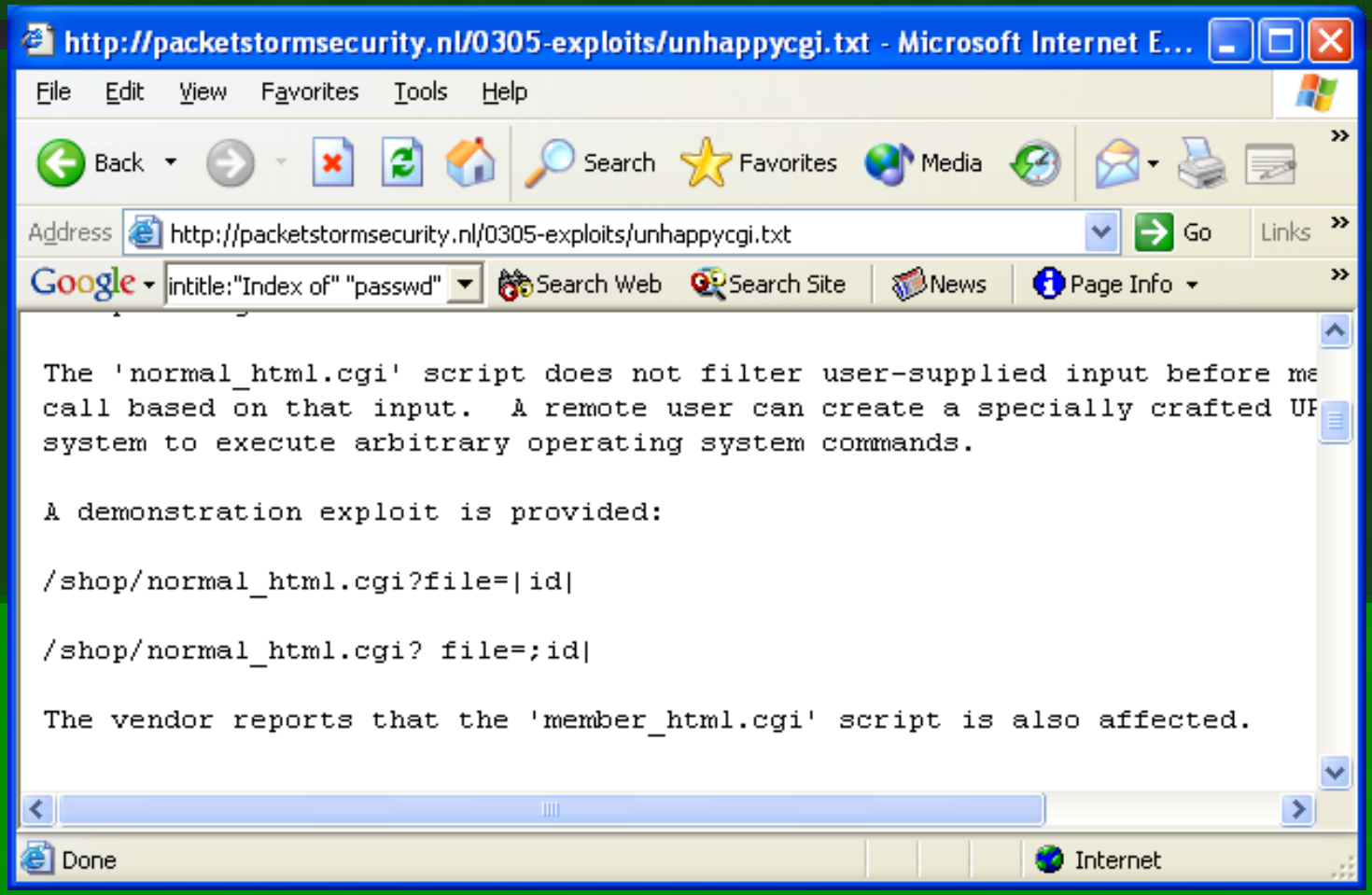
finding 0day targets...

vulnerability trolling



A new vulnerability hits the streets...

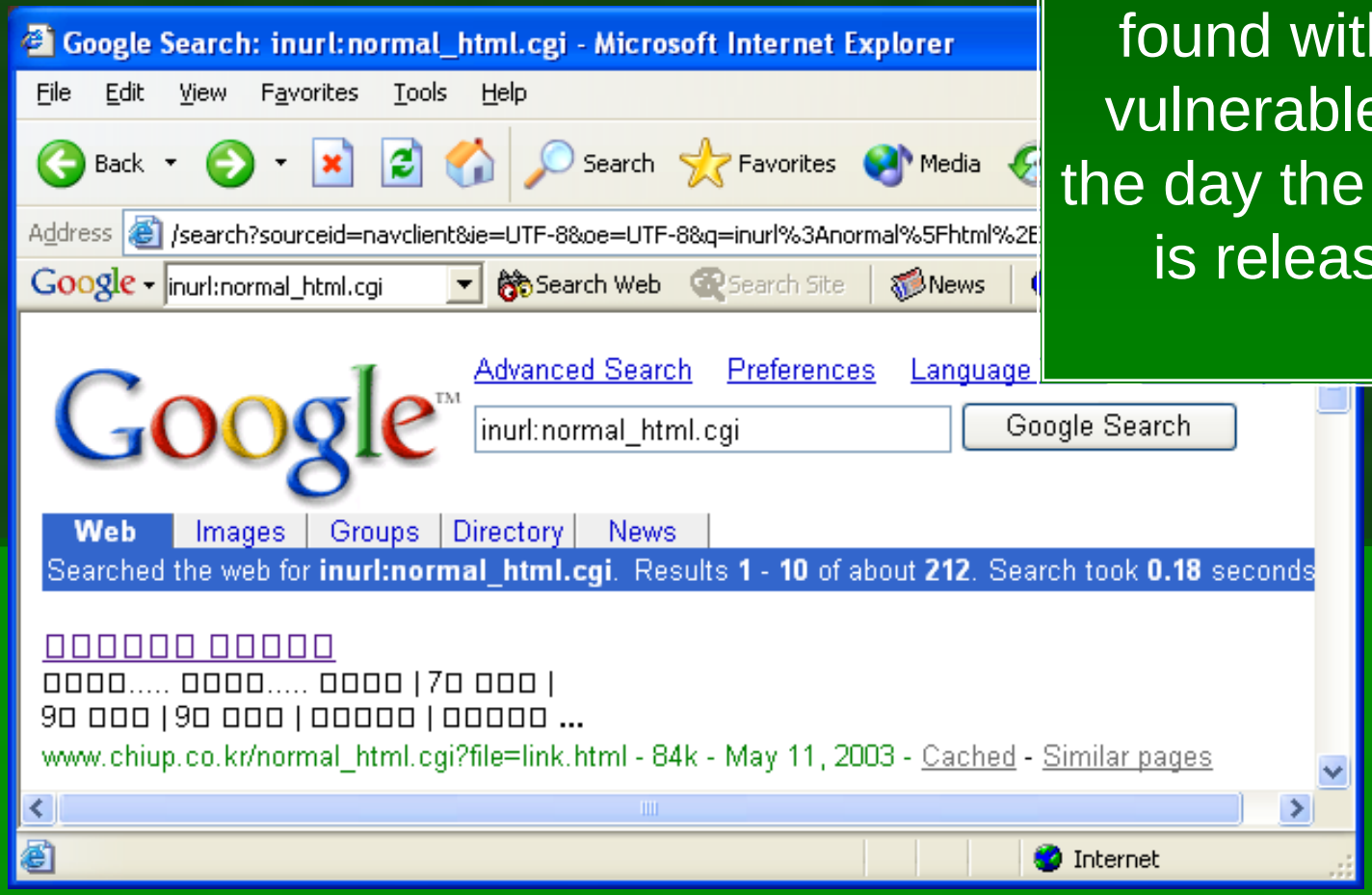
vulnerability trolling



The vulnerability lies in a cgi script called “normal_html.cgi”

vulnerability trolling

212 sites are found with the vulnerable CGI the day the exploit is released.



more interesting stuff...

finding sensitive data in directories and files

Directory Listings

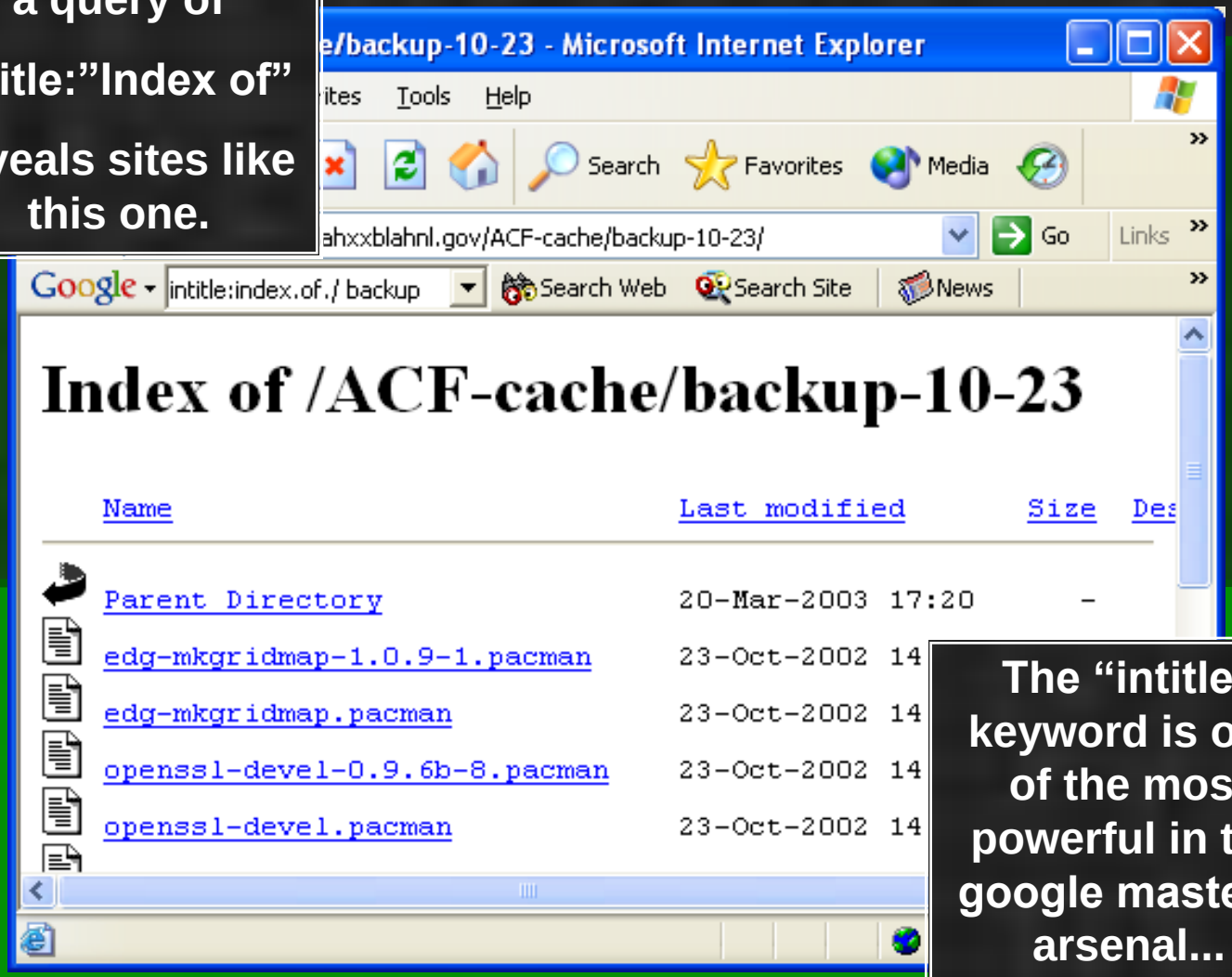
- Directory listings are often misconfigurations in the web server.
- A directory listing shows a list of files in a directory as opposed to presenting a web page.
- Directory listings can provide very useful information.

Directory Example

a query of

intitle:"Index of"

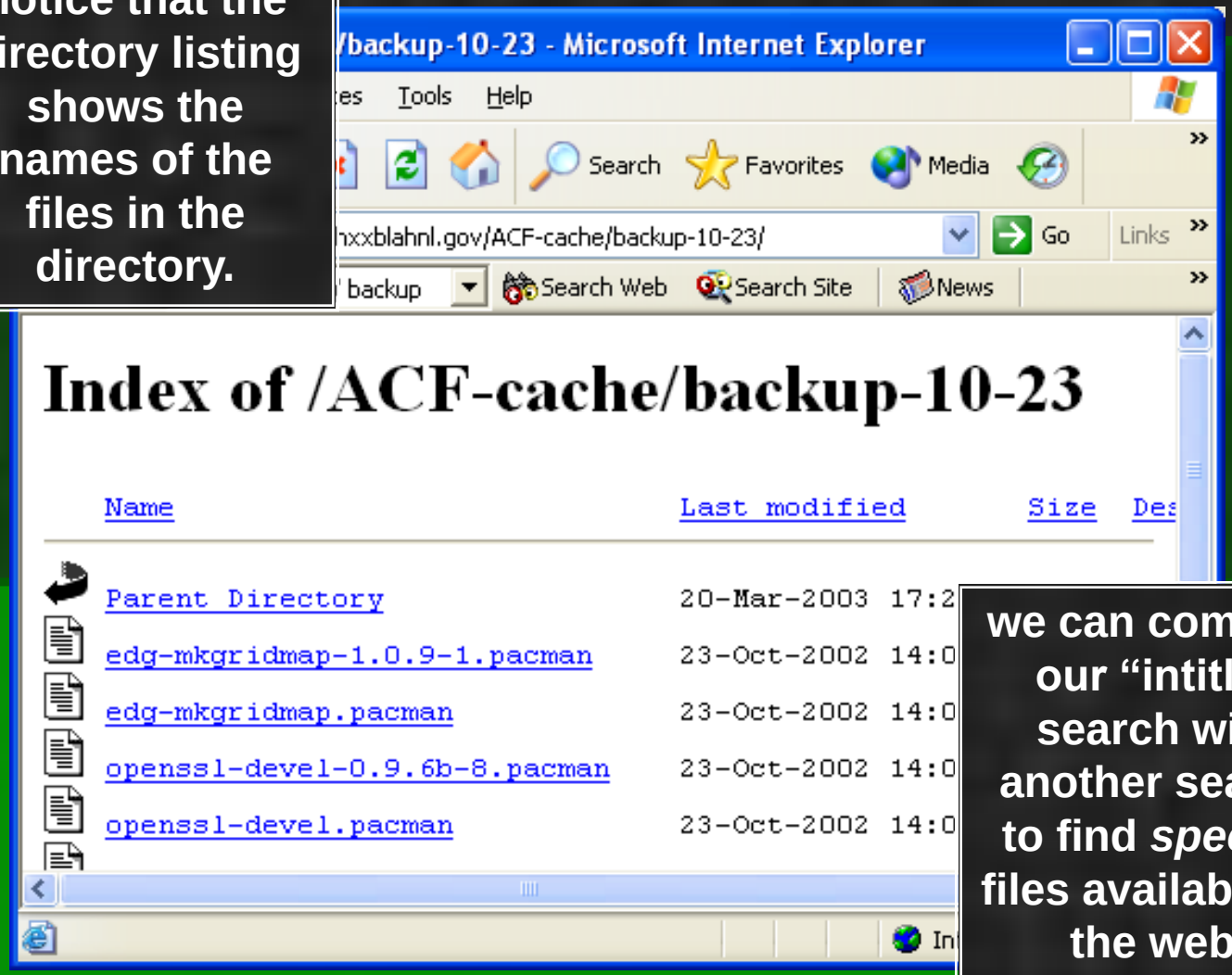
reveals sites like this one.



The "intitle" keyword is one of the most powerful in the google master's arsenal...

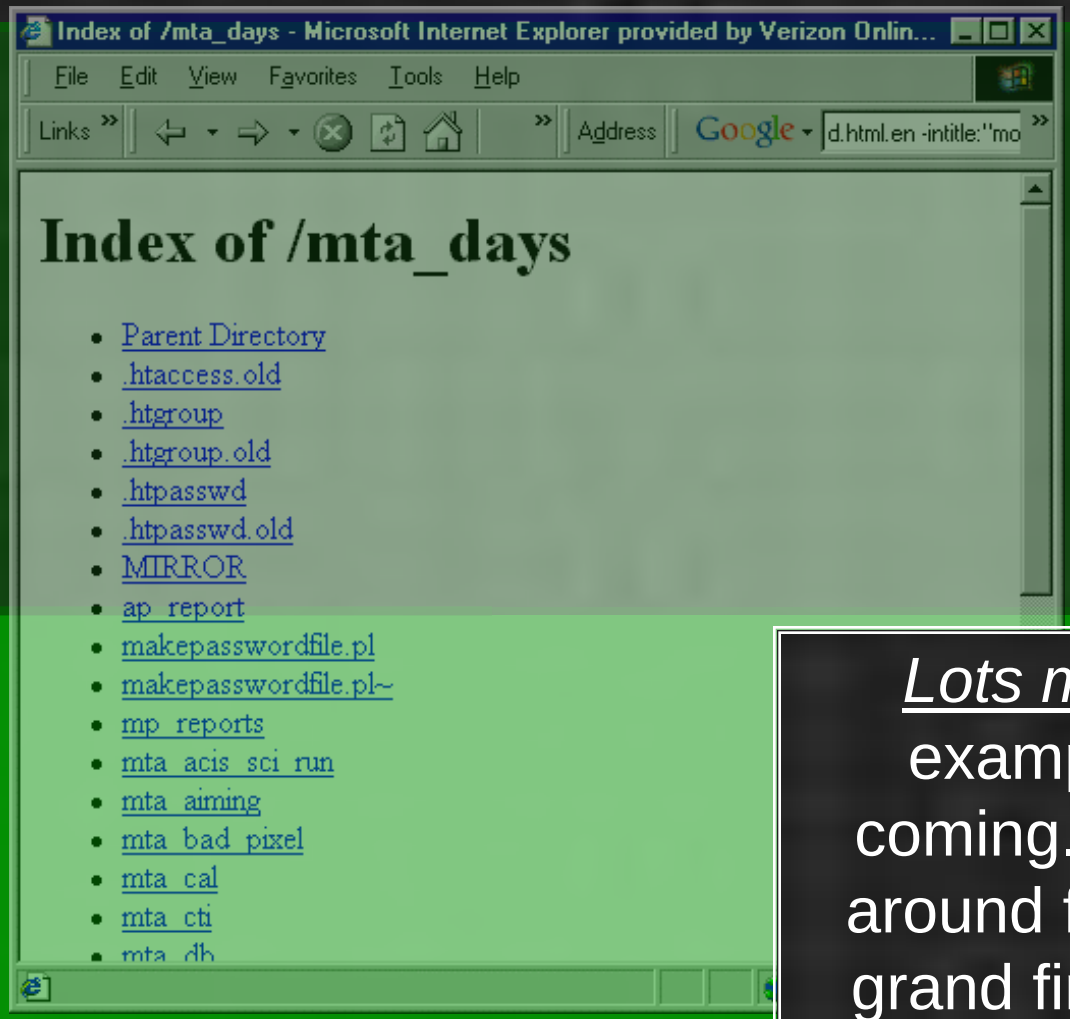
Directory Example

notice that the directory listing shows the names of the files in the directory.



we can combine our “intitle” search with another search to find *specific* files available on the web.

Intitle:”Index of” .htpasswd



Lots more examples coming. Stick around for the grand finale...

finding interesting stuff

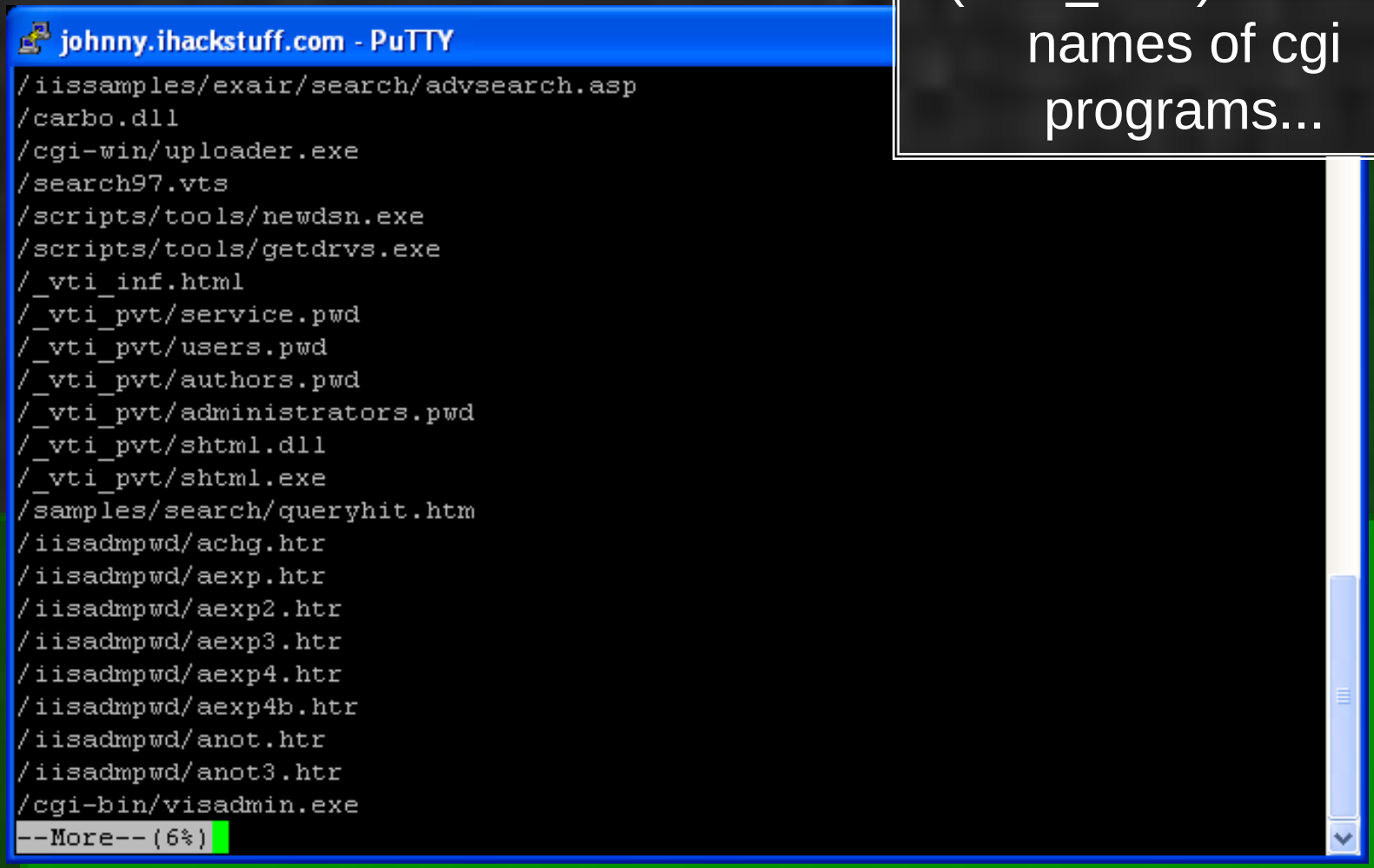
automation: googlescan

Googlescan

- With a known set of file-based web vulnerabilities, a vulnerability scanner based on search engines is certainly a reality.
- Let’s take a look at a painfully simple example using nothing more than UNIX shell commands...

Googlescan.sh

first, create a file (vuln_files) with the names of cgi programs...



Googlescan.sh

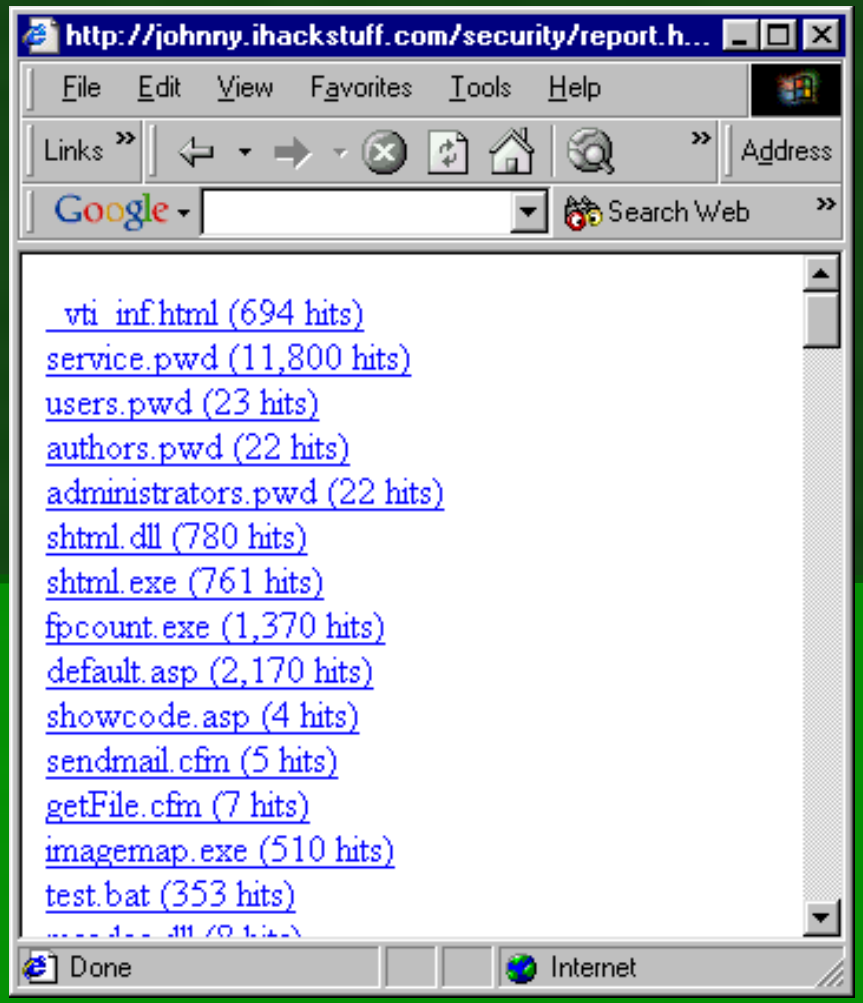
...then, use this shell script...

```
rm temp
awk -F"/"
  '{print $NF"|http://www.google.com/search?q=
    intitle%3A%22Index+of%22+"$NF}' vuln_files > queries

for query in `cat queries`
do
      echo -n $query"|"  >> temp
      echo $query | awk -F"|" '{print $2}'
      lynx -source `echo $query | awk -F"|" '{print $2}'` |
      grep "of about"                                     |
      awk -F "of about" '{print $2}'                      |
      awk -F"." '{print $1}'                              |
      tr -d "</b>[:cntrl:] "                           >> temp
      echo " "                                         >> temp
Done

cat temp |
   awk -F"|" '{print "<A HREF=\"" $2 "\">" $1 " (" $3 "hits)
   </A><BR><BR>"}' | grep -v "(1,770,000" > report.html
```

Googlescan.sh output



...to output an html list of potentially vulnerable or interesting web servers according to Google.

http://johnny.ihackstuff.com/googledorks.shtml

more interesting stuff

Rise of the Robots

Rise of the Robots

- “Rise of the Robots”, Phrack 57-10 by Michal Zalewski: autonomous malicious robots powered by public search engines
- Search engine crawlers pick up malicious links and follow them, actively exploiting targets

Rise of the Robots: Example

Michal presents the following example links on his indexed web page:

http://somehost/cgi-bin/script.pl?p1=../../../../attack
http://somehost/cgi-bin/script.pl?p1=;attack
http://somehost/cgi-bin/script.pl?p1=|attack
http://somehost/cgi-bin/script.pl?p1=`attack`
http://somehost/cgi-bin/script.pl?p1=$(attack)
http://somehost:54321/attack?`id`
http://somehost/AAAAAAAAAAAAAAAAAAAAA...

Rise of the Robots: Results

- Within Michal's study, the robots followed all the links as written, including connecting to *non-http ports!*
- The robots followed the "attack links," performing the attack completely unawares.
- Moral: Search engines can attack for you, and store the results, all without an attacker sending a single packet directly to the target.

Prevention

Locking it down

Google's advice

- This isn't Google's fault.
- Google is very happy to remove references. See http://www.google.com/remove.html.
- Follow the webmaster advice found at http://www.google.com/webmasters/faq.html.

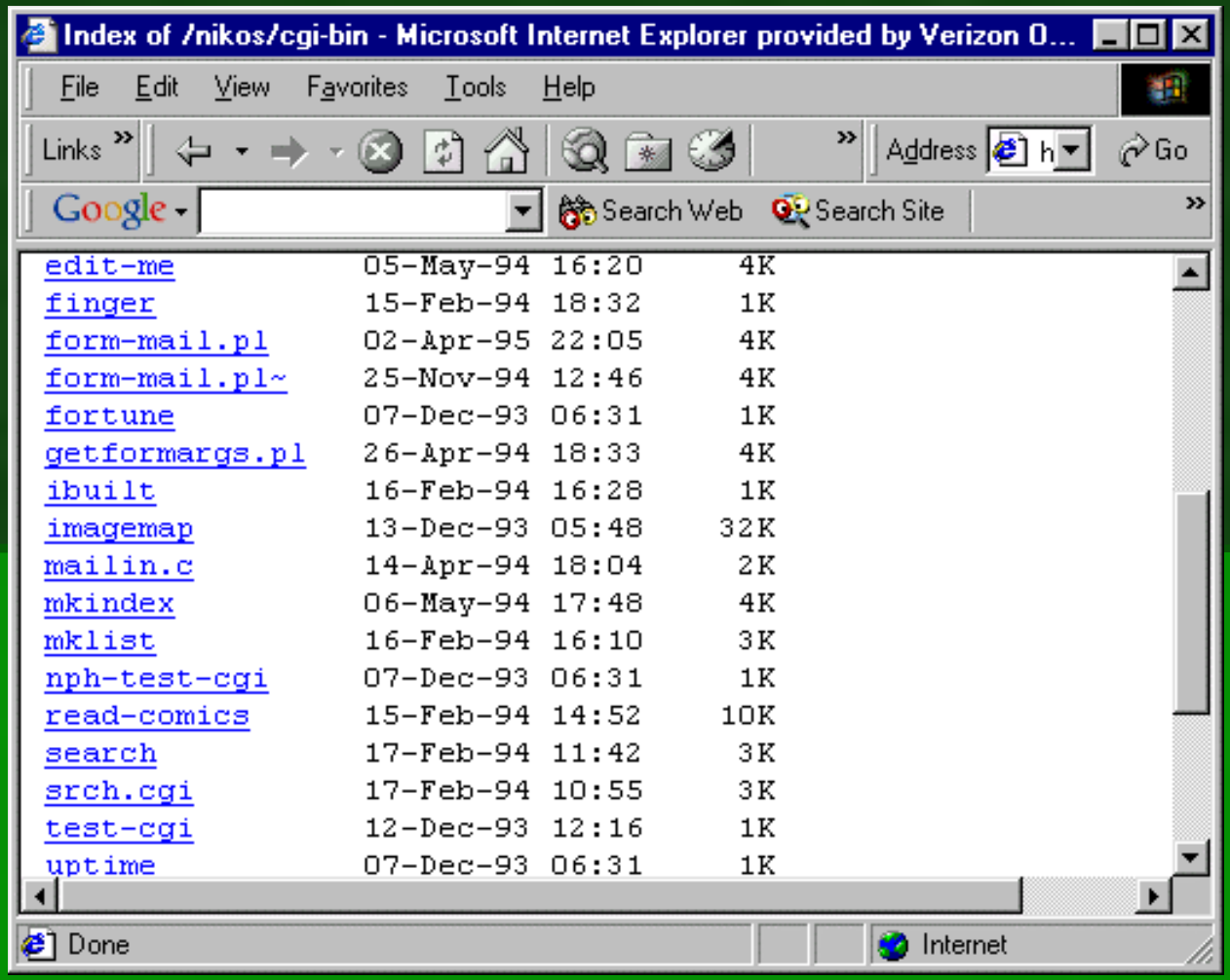
My advice

- Don’t be a dork. Keep it off the web!
- Scan yourself.
- Be proactive.
- Watch googledorks (http://johnny.ihackstuff.com/googledorks.shtml)

Finally….

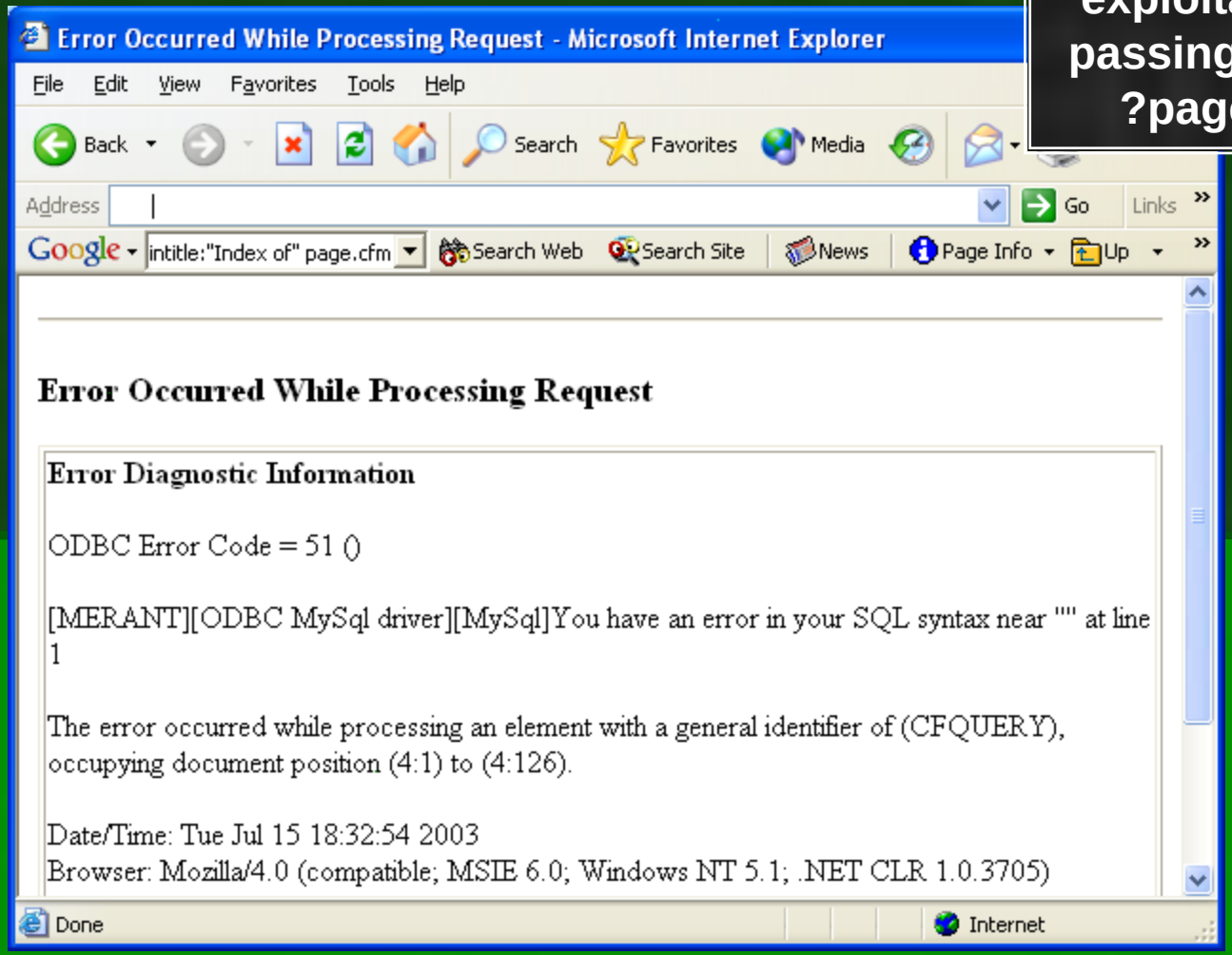
The Grand Finale!

intitle:index.of test-cgi

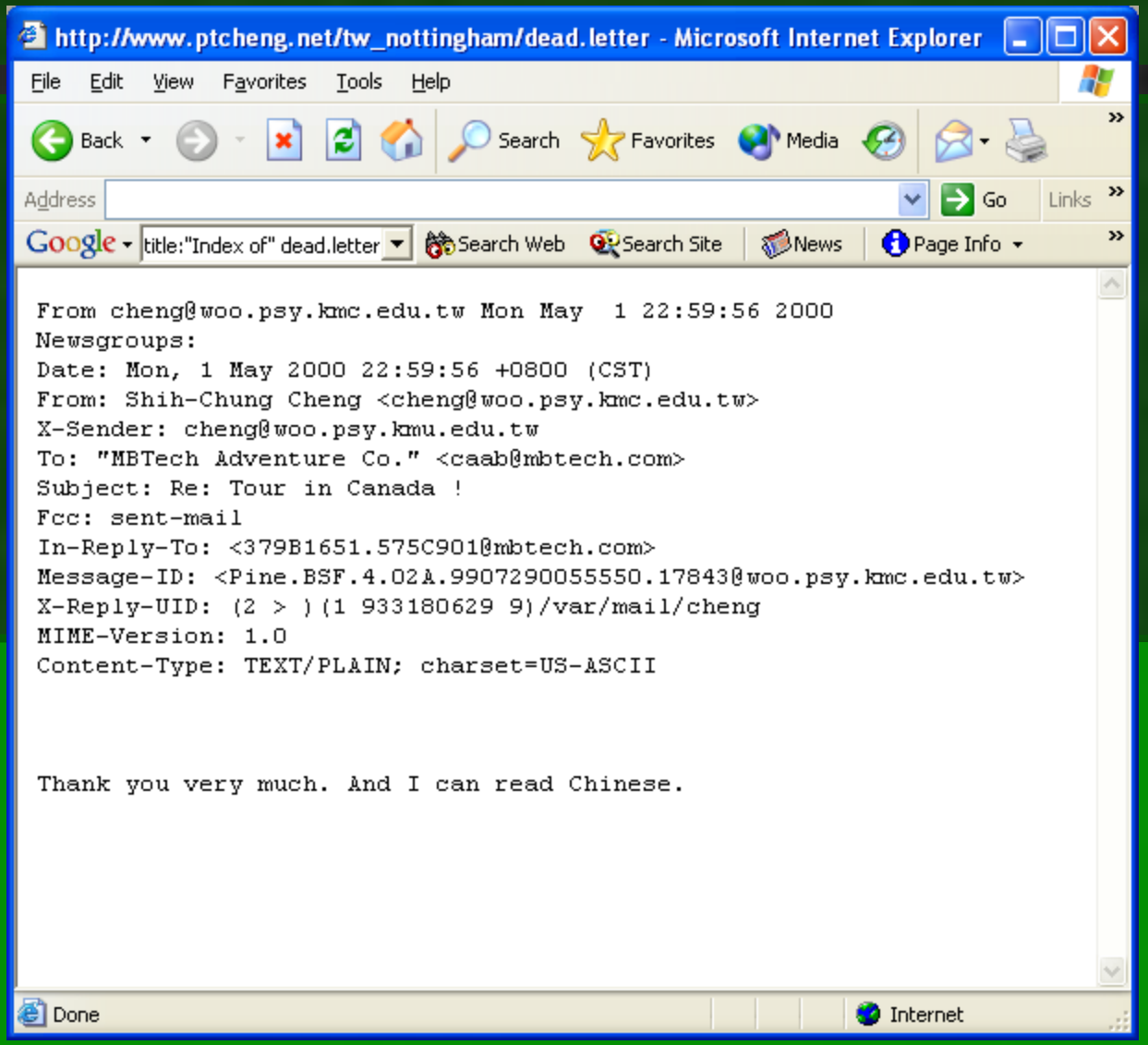


intitle:index.of page.cfm

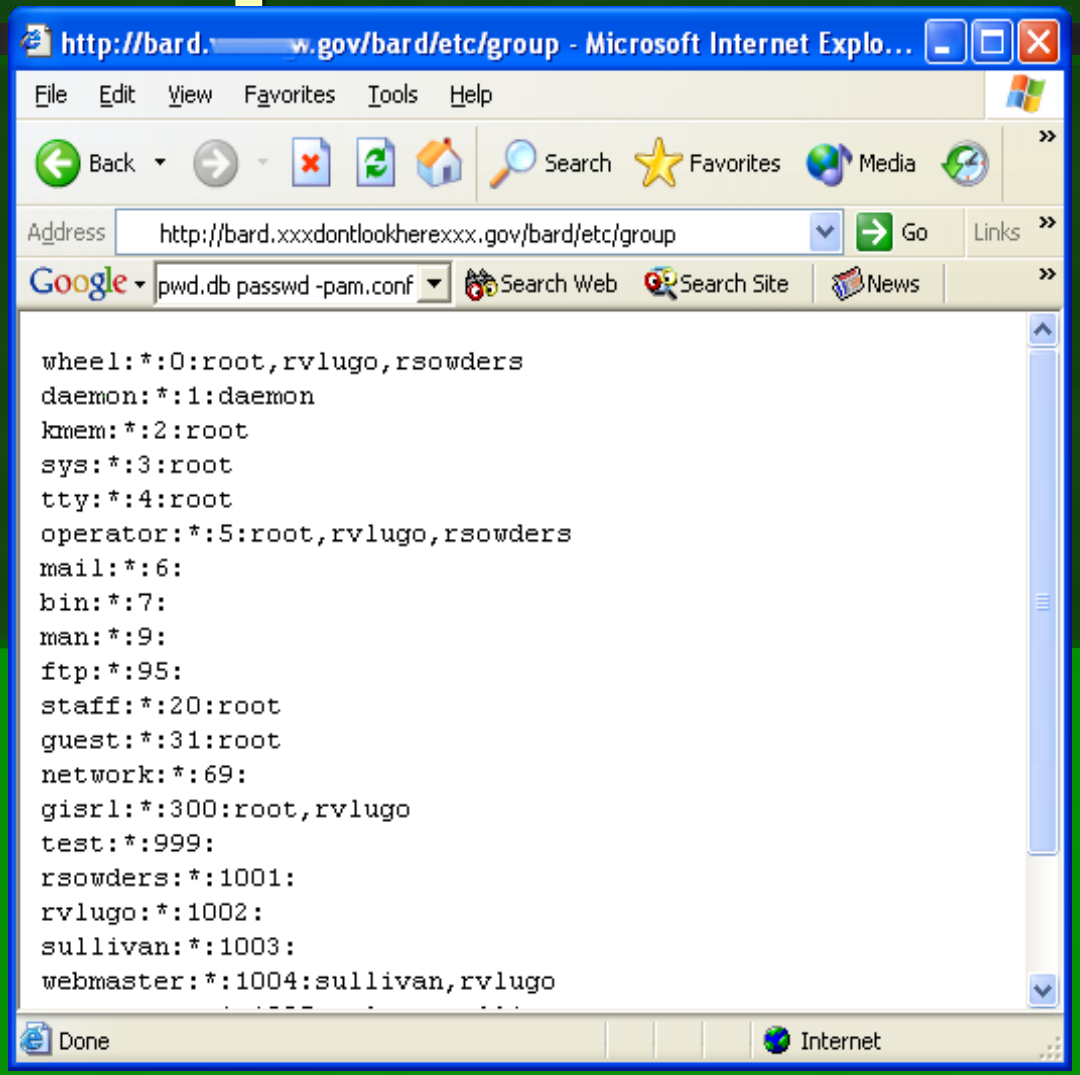
exploitable by passing invalid ?page_id=



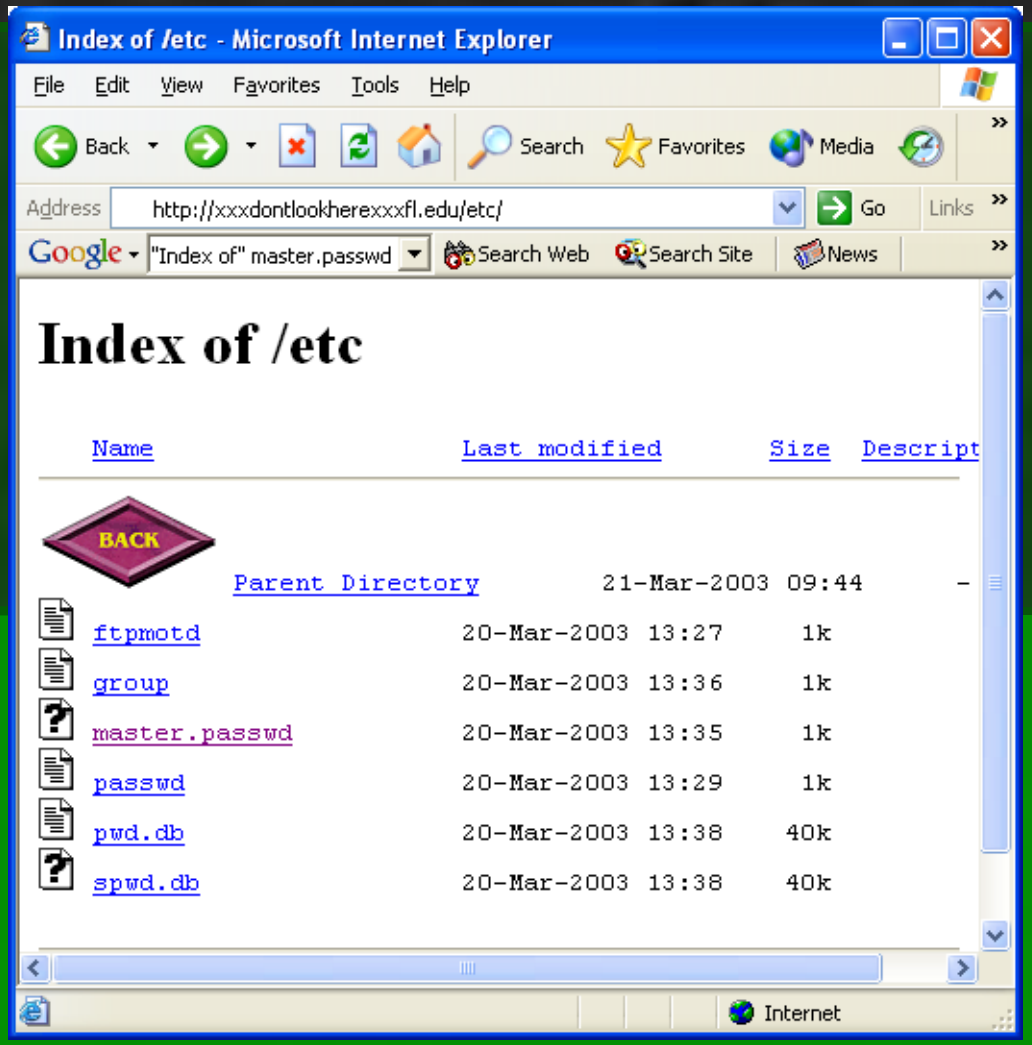
intitle:index.of dead.letter



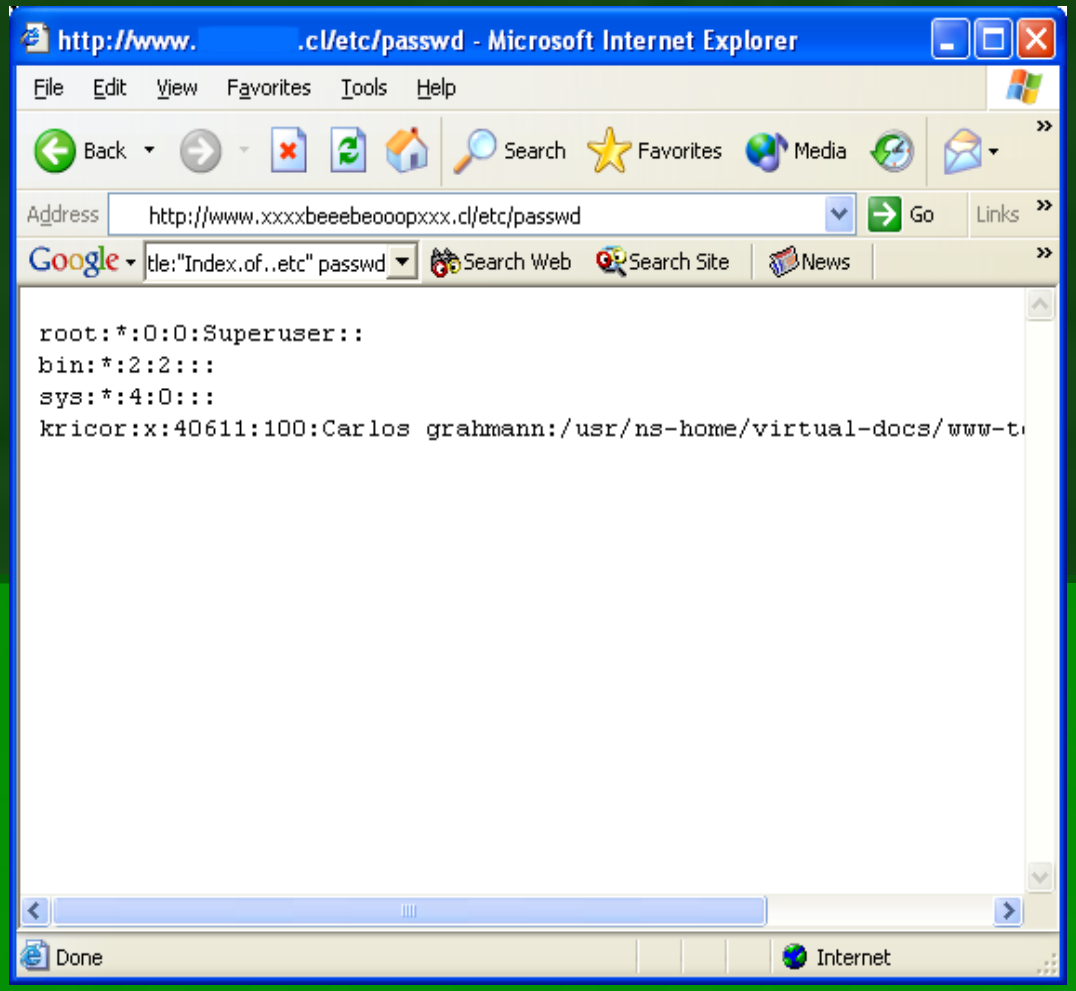
intitle:index.of pwd.db passwd –pam.conf



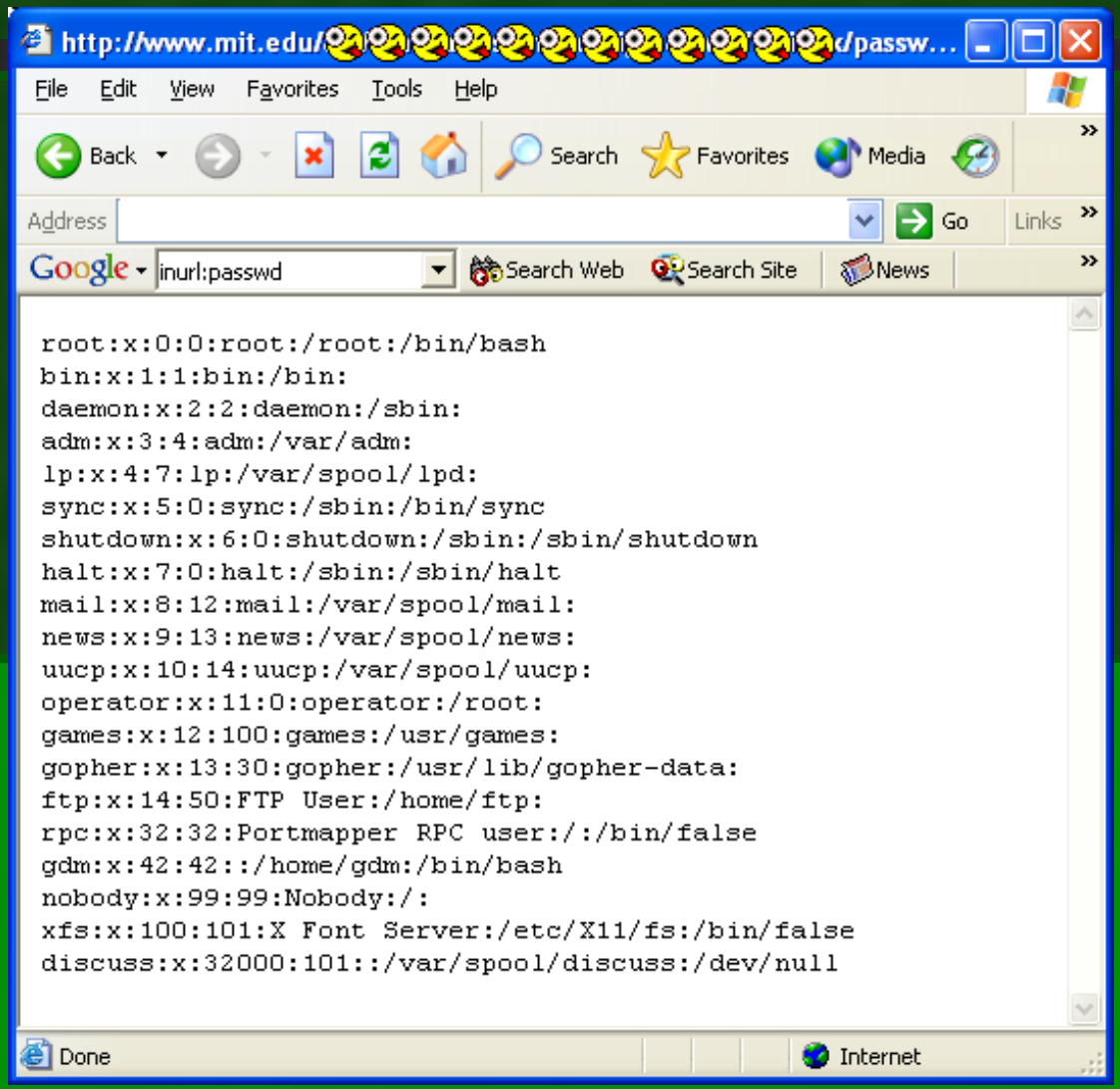
intitle:index.of master.passwd



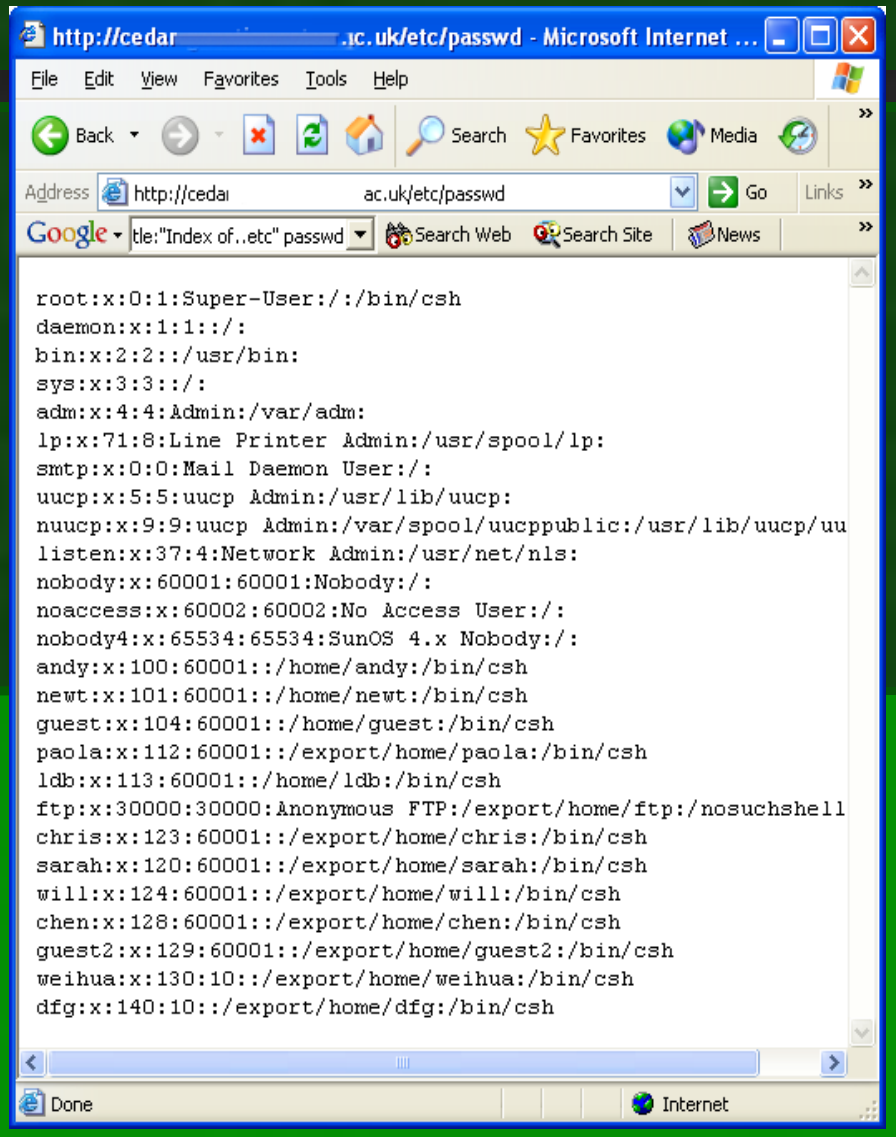
intitle:index.of..etc passwd



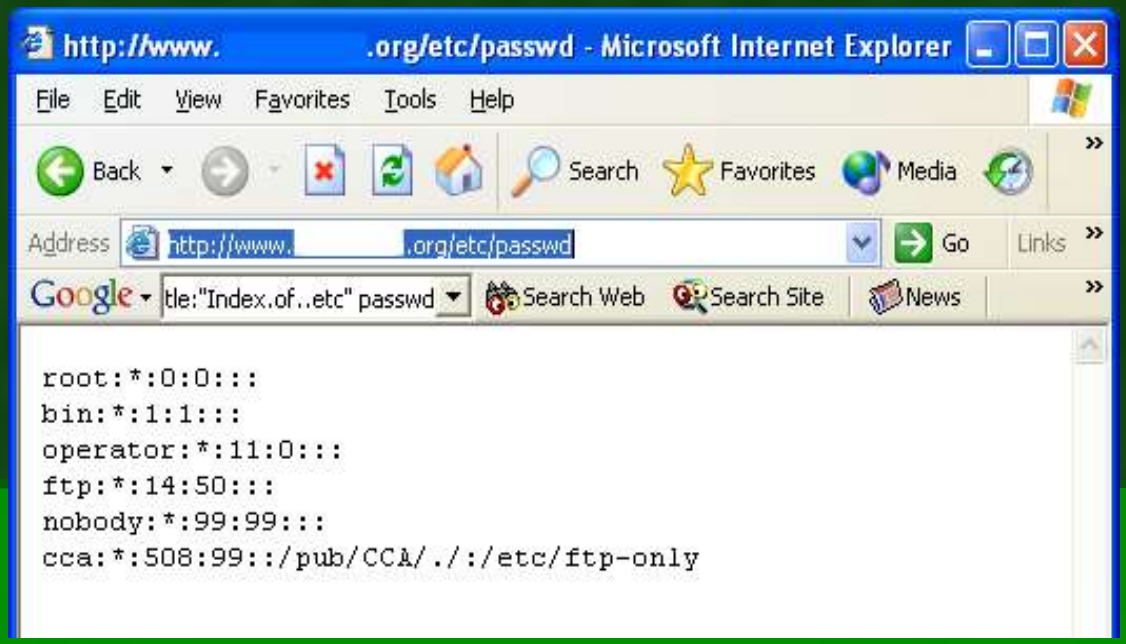
intitle:index.of passwd



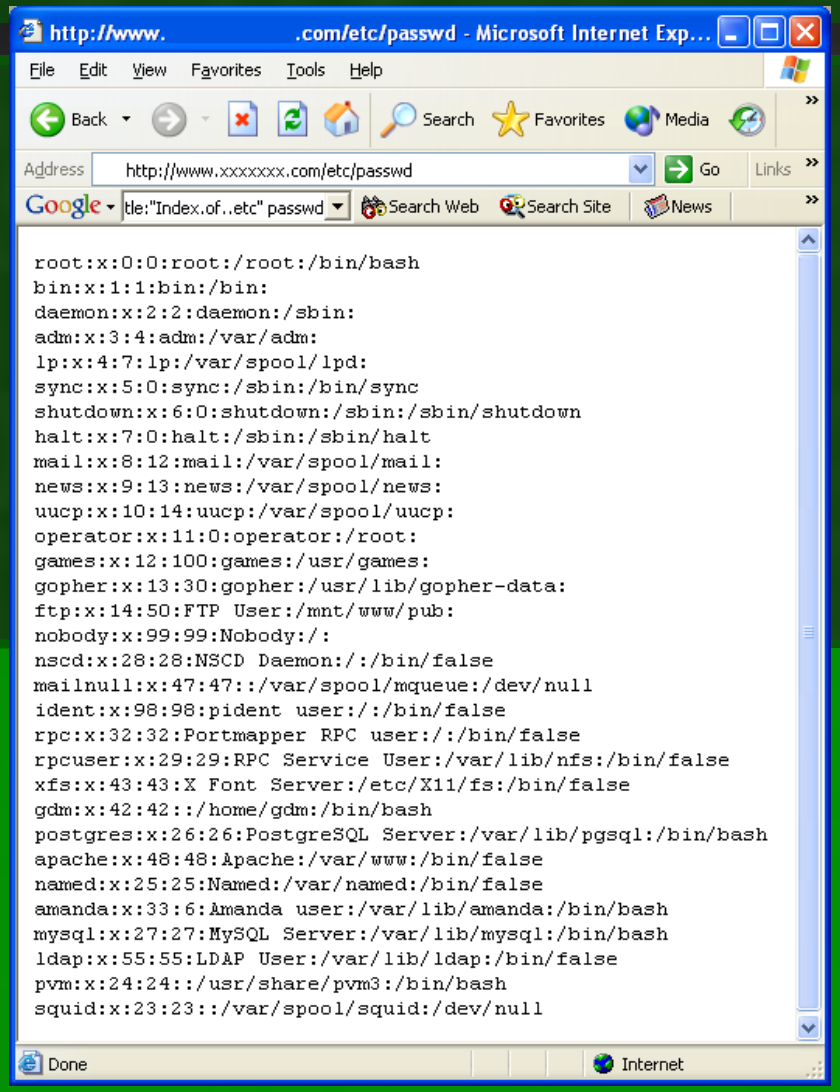
intitle:"Index.of..etc" passwd



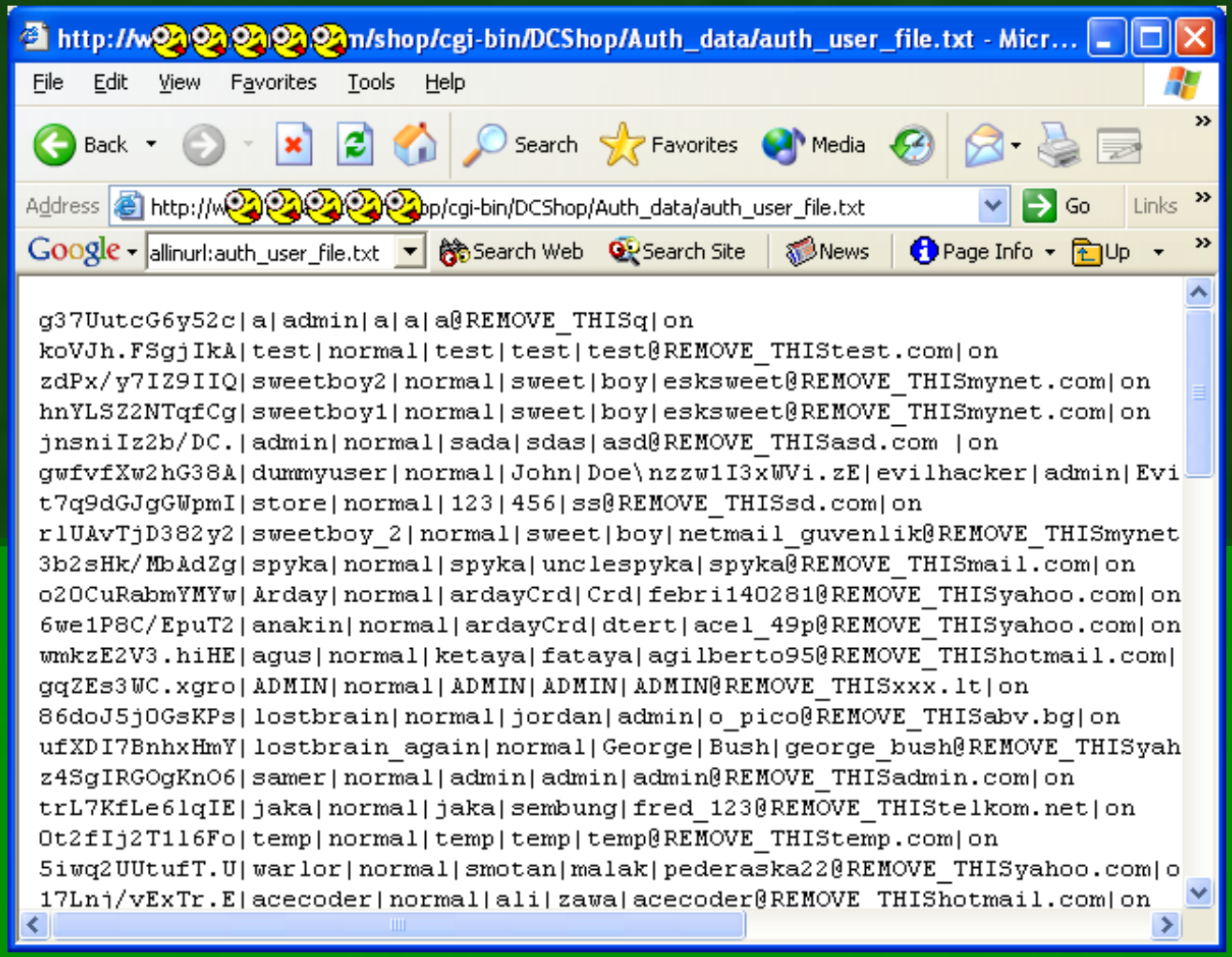
intitle:"Index.of..etc" passwd



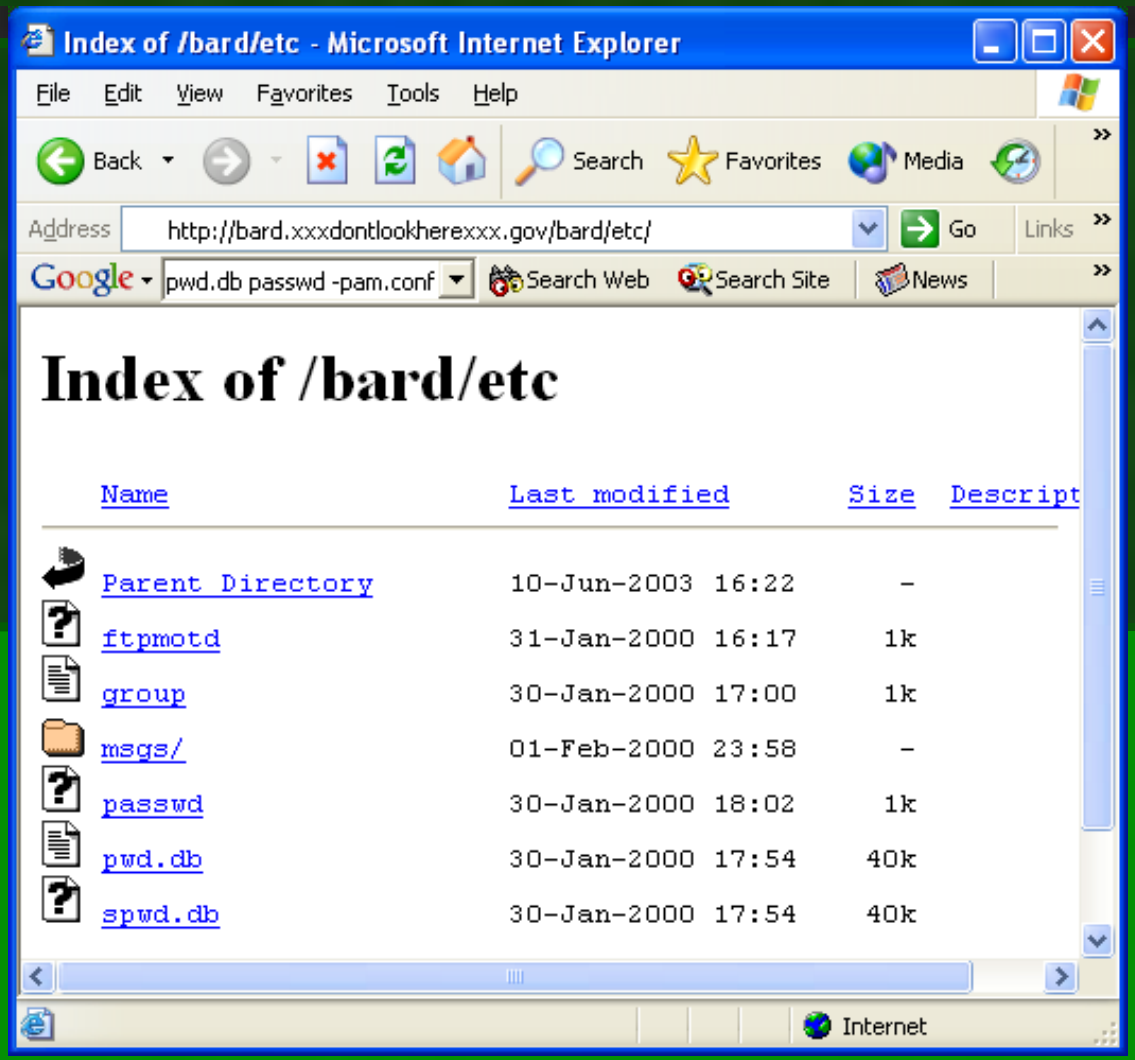
intitle:"Index.of..etc" passwd



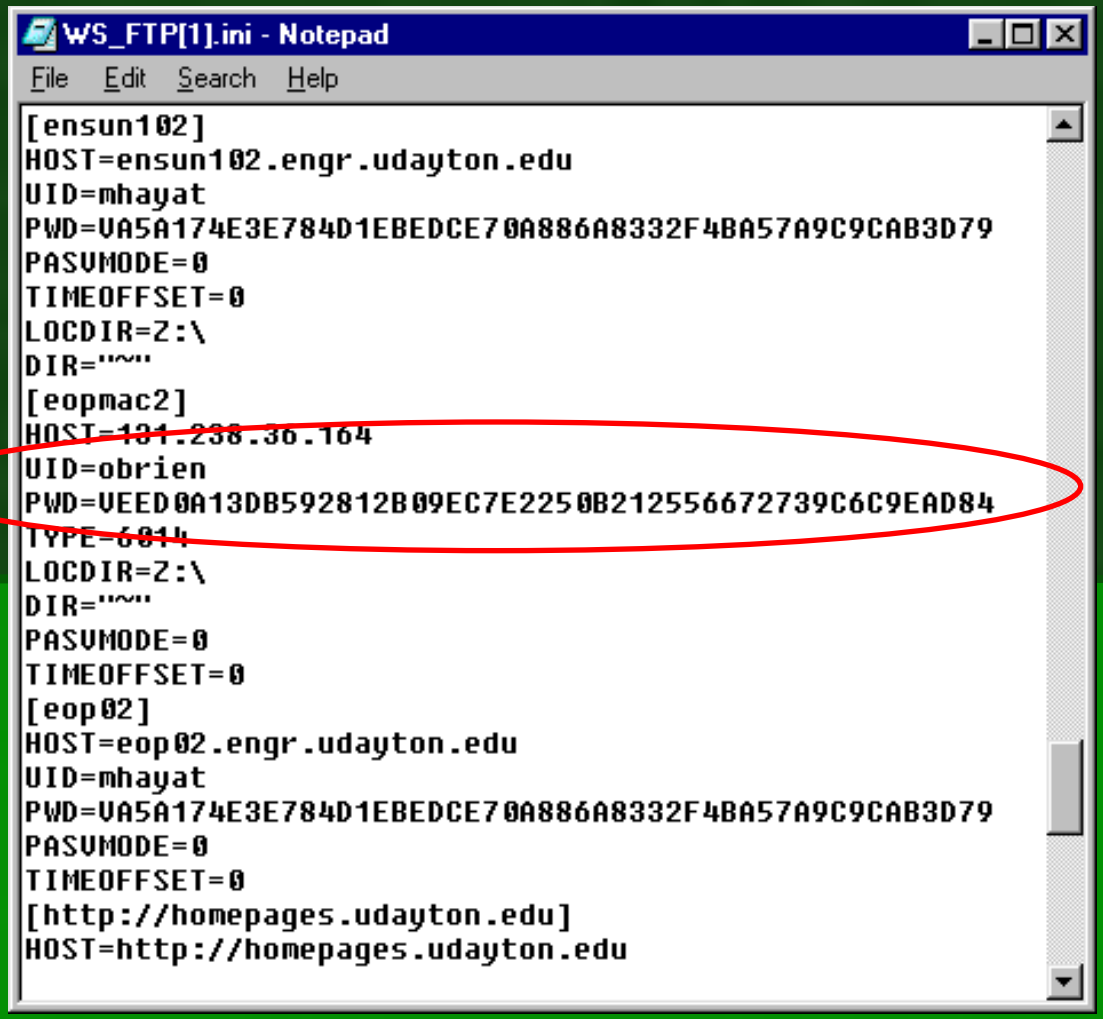
intitle:index.of auth_user_file.txt



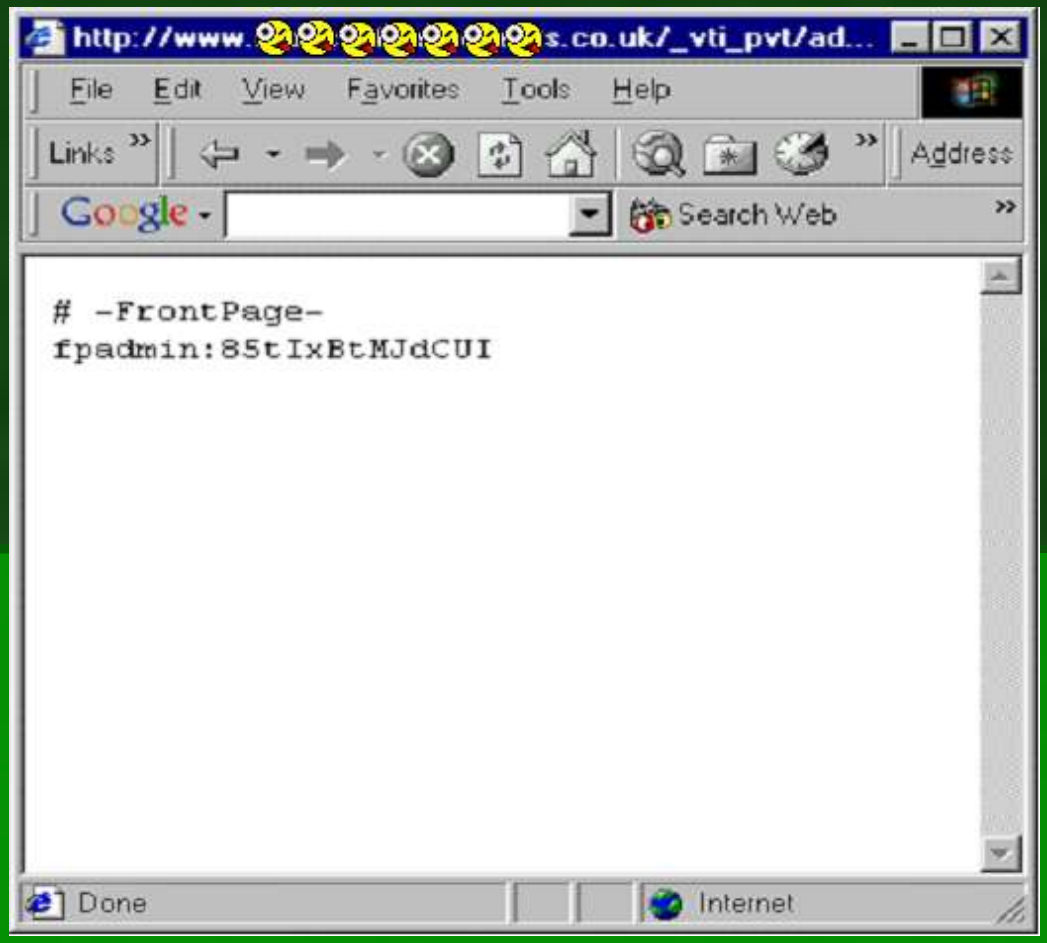
intitle:index.of pwd.db passwd –pam.conf



intitle:index.of ws_ftp.ini

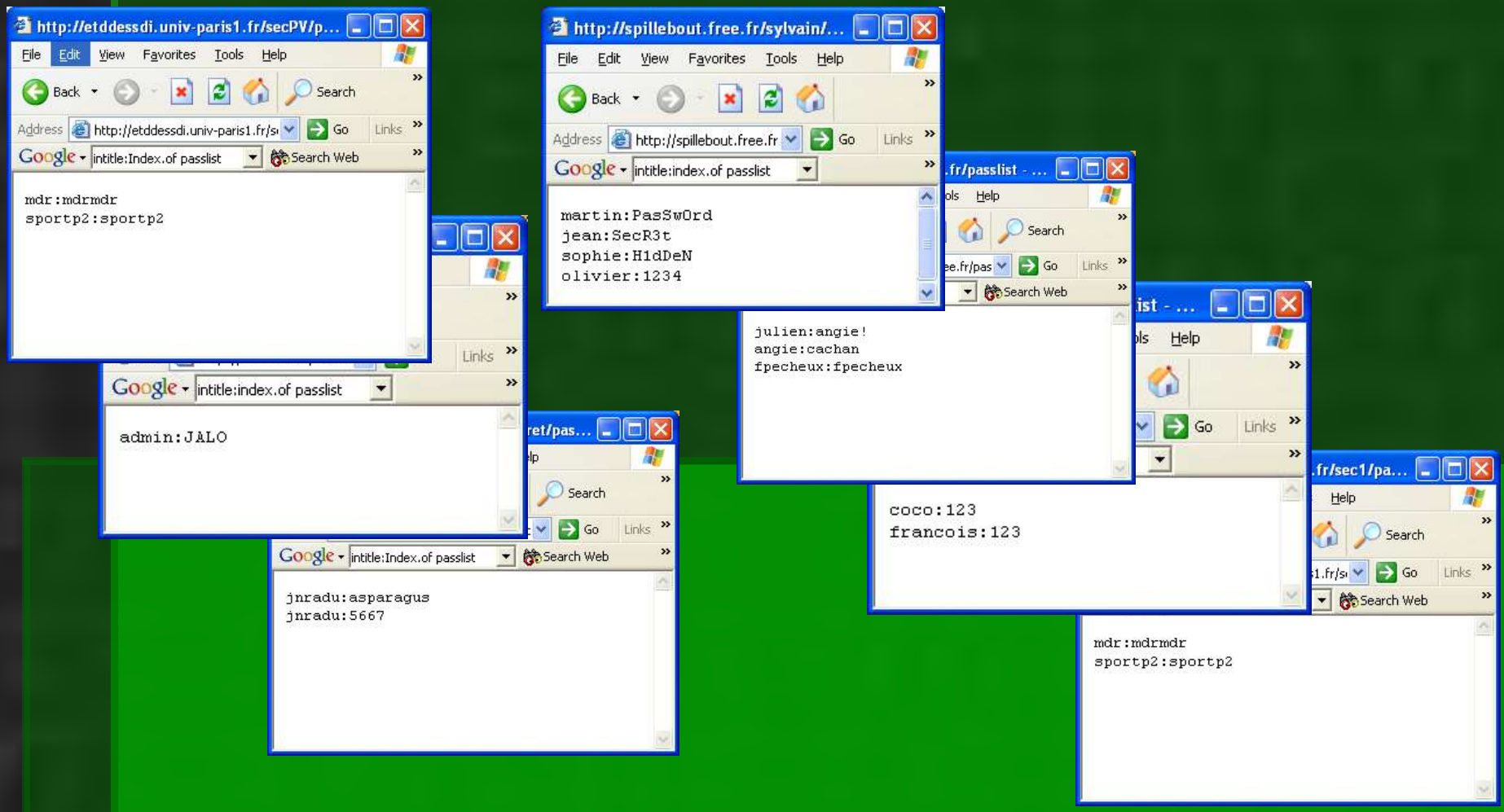


intitle:index.of administrators.pwd

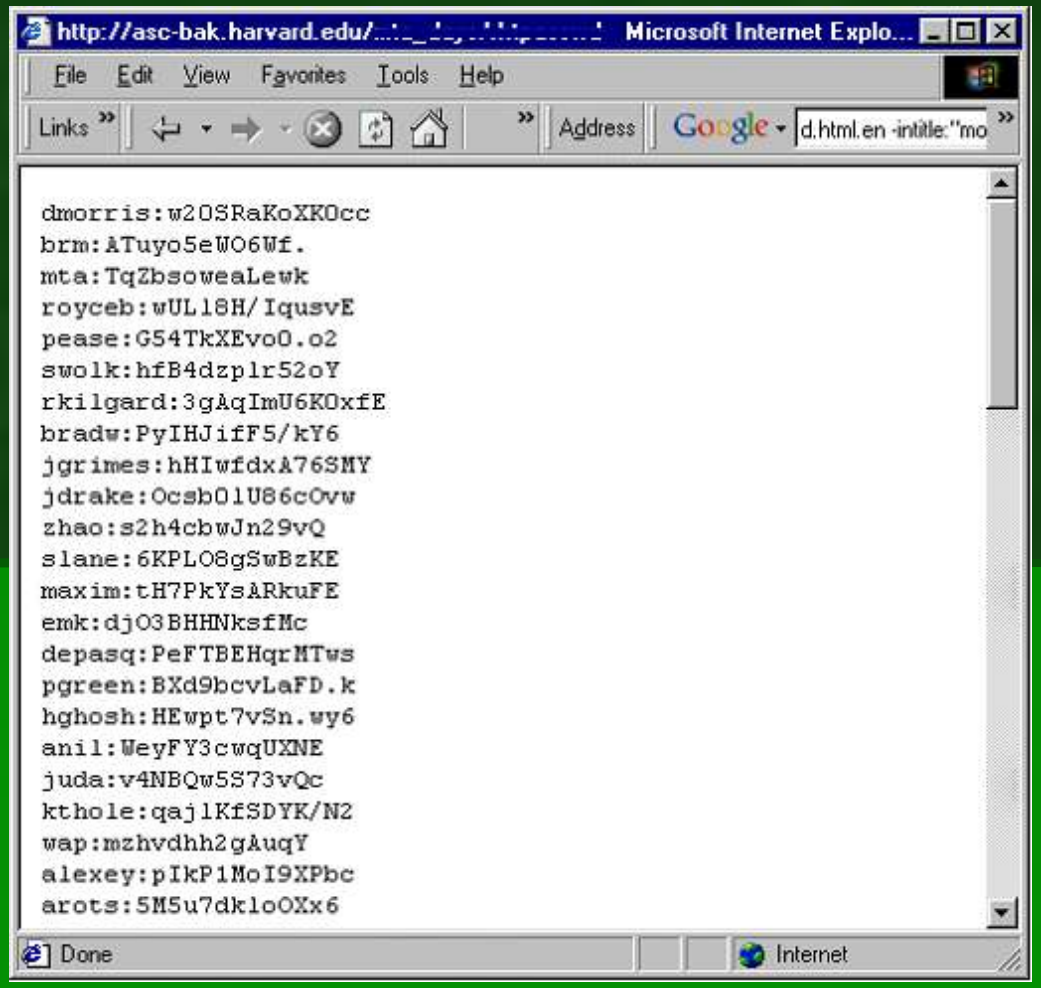


intitle:index.of people.lst

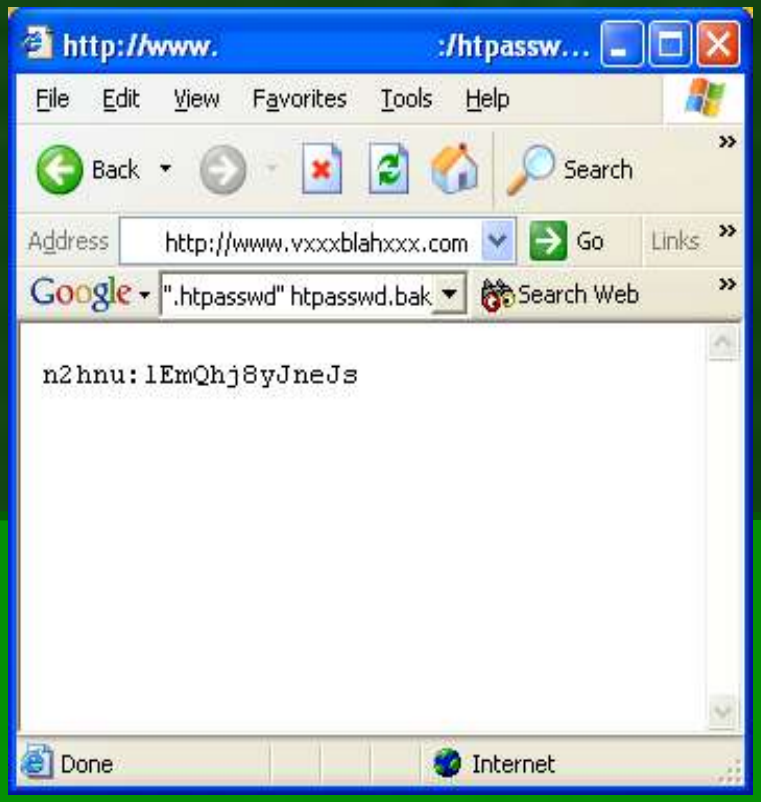
intitle:index.of passlist



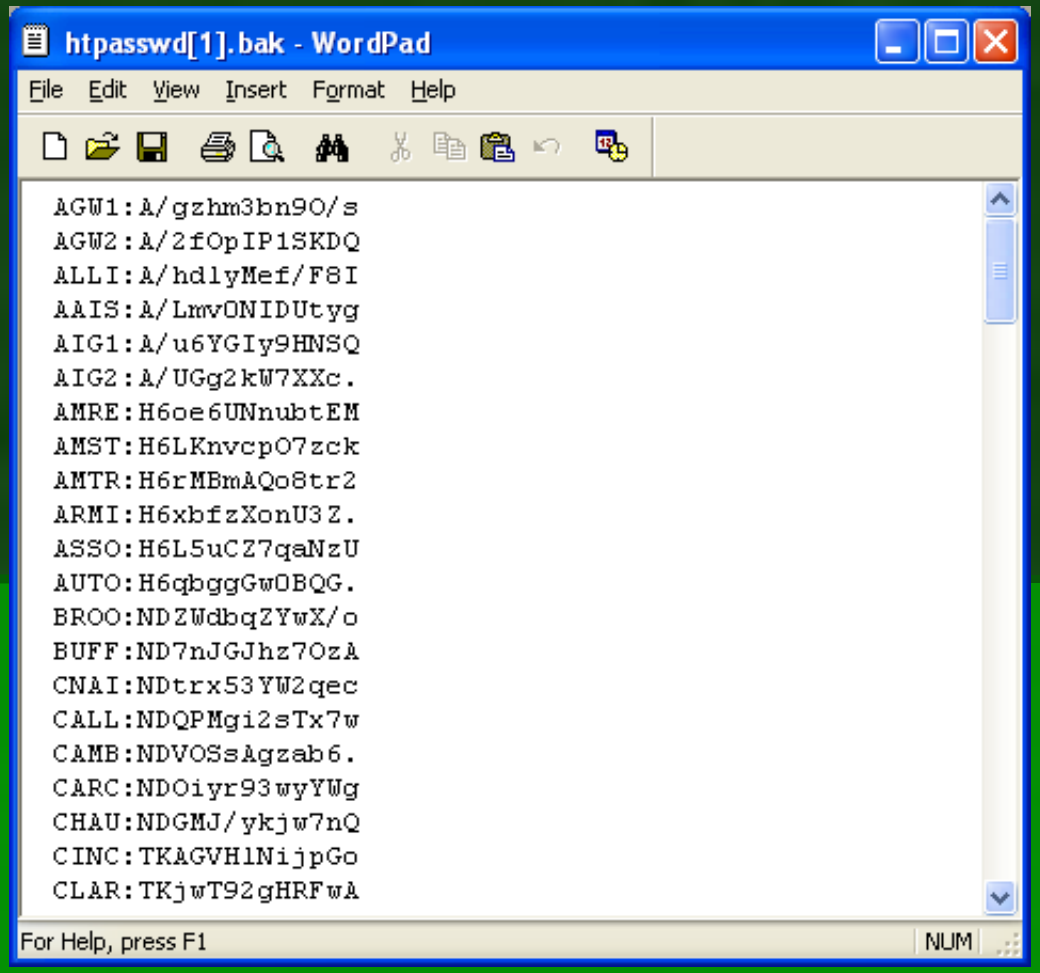
intitle:index.of .htpasswd



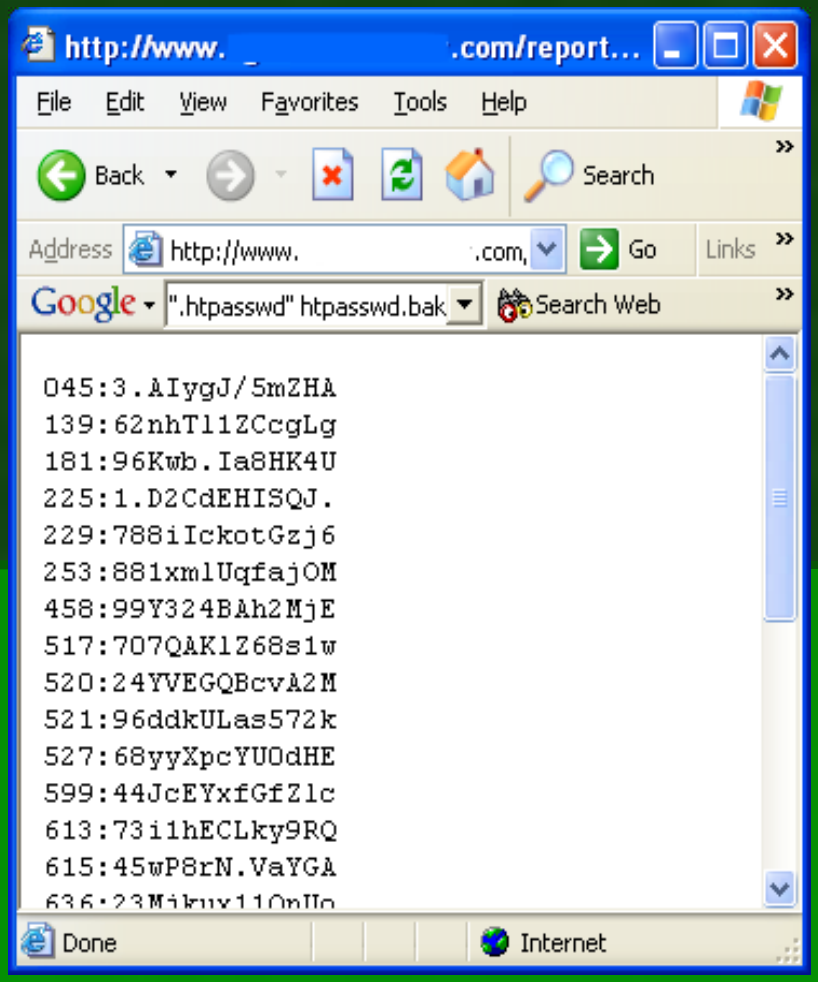
intitle:index.of “.htpasswd” htpasswd.bak



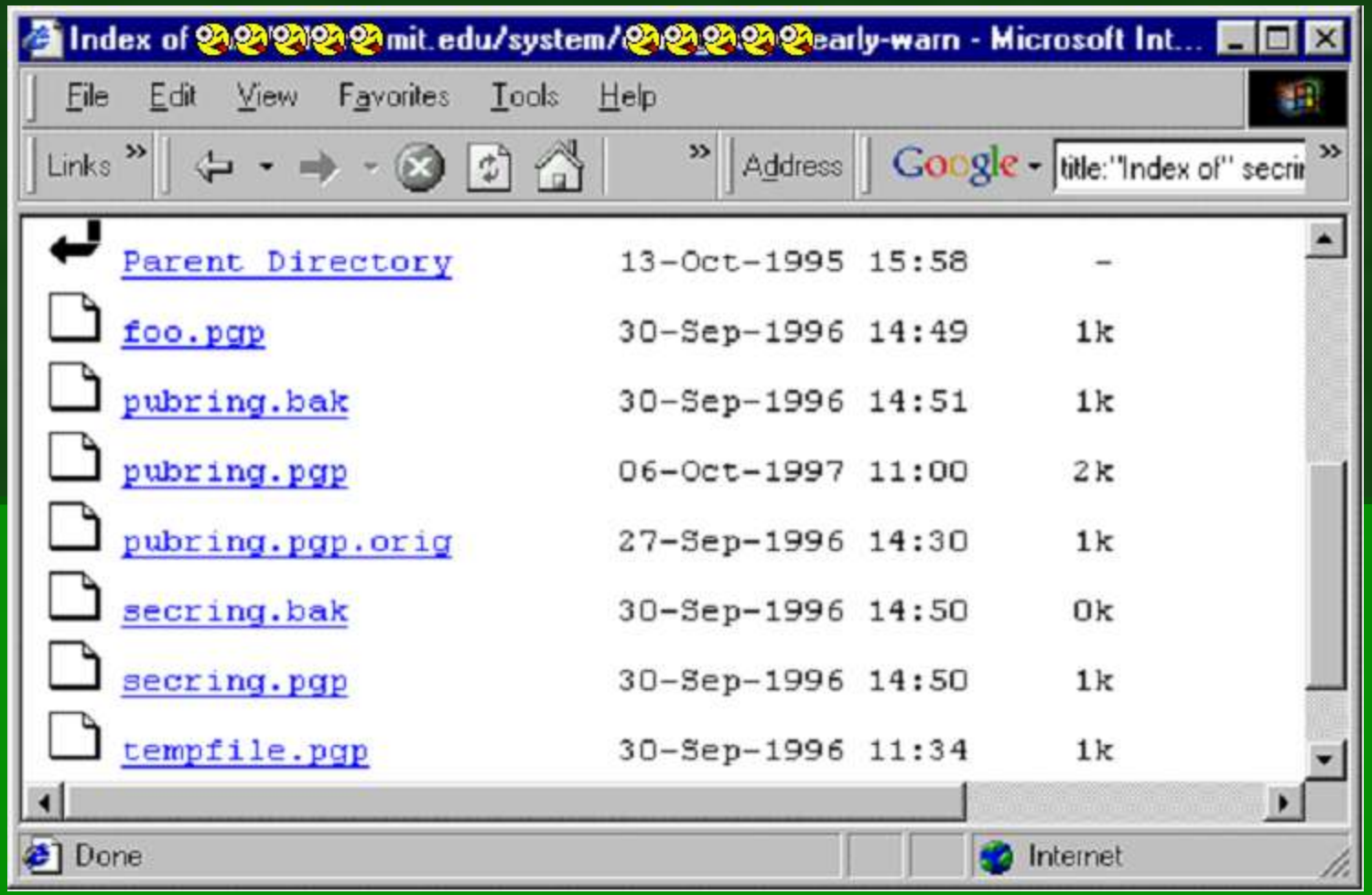
intitle:index.of “.htpasswd” htpasswd.bak



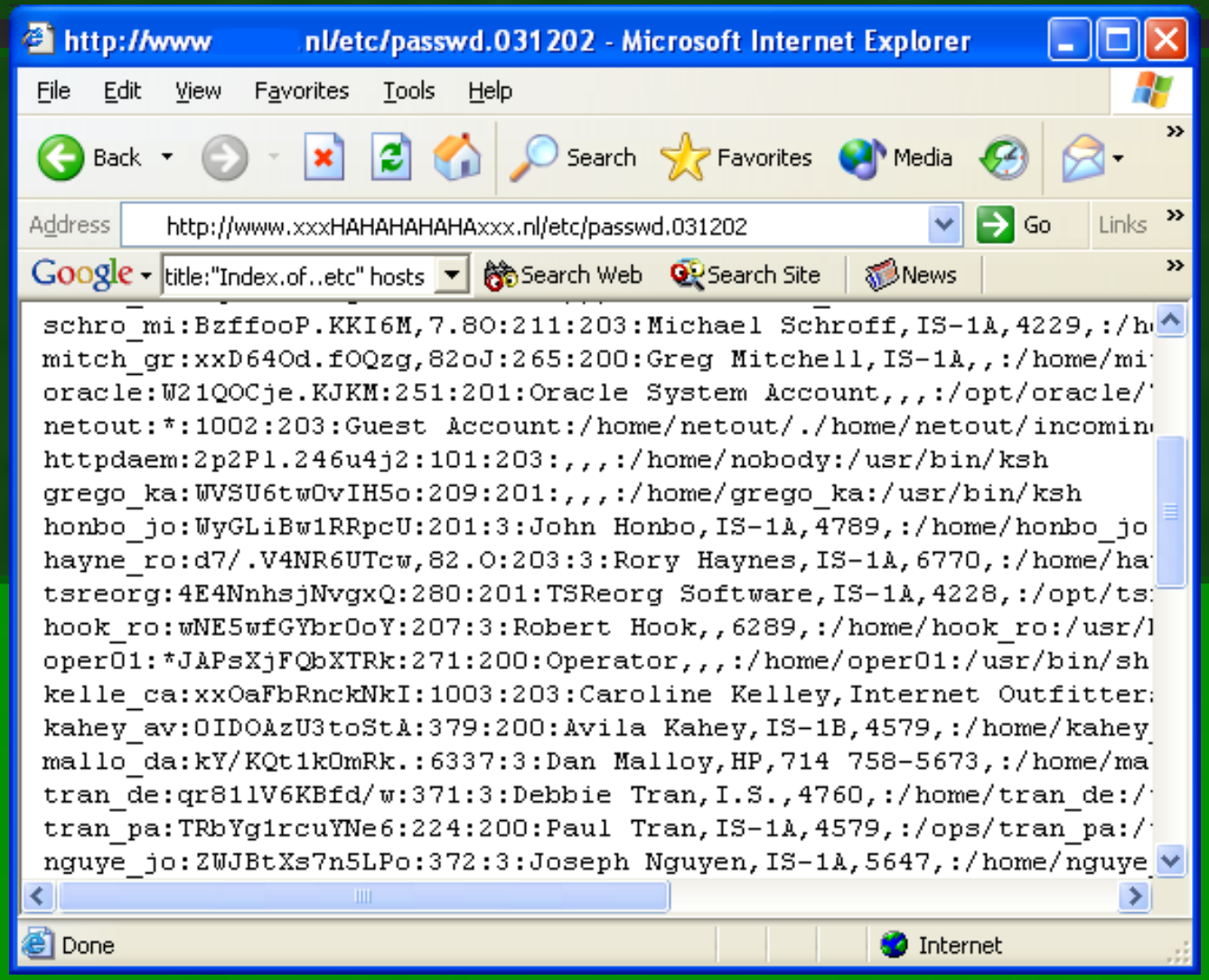
intitle:index.of “.htpasswd” htpasswd.bak



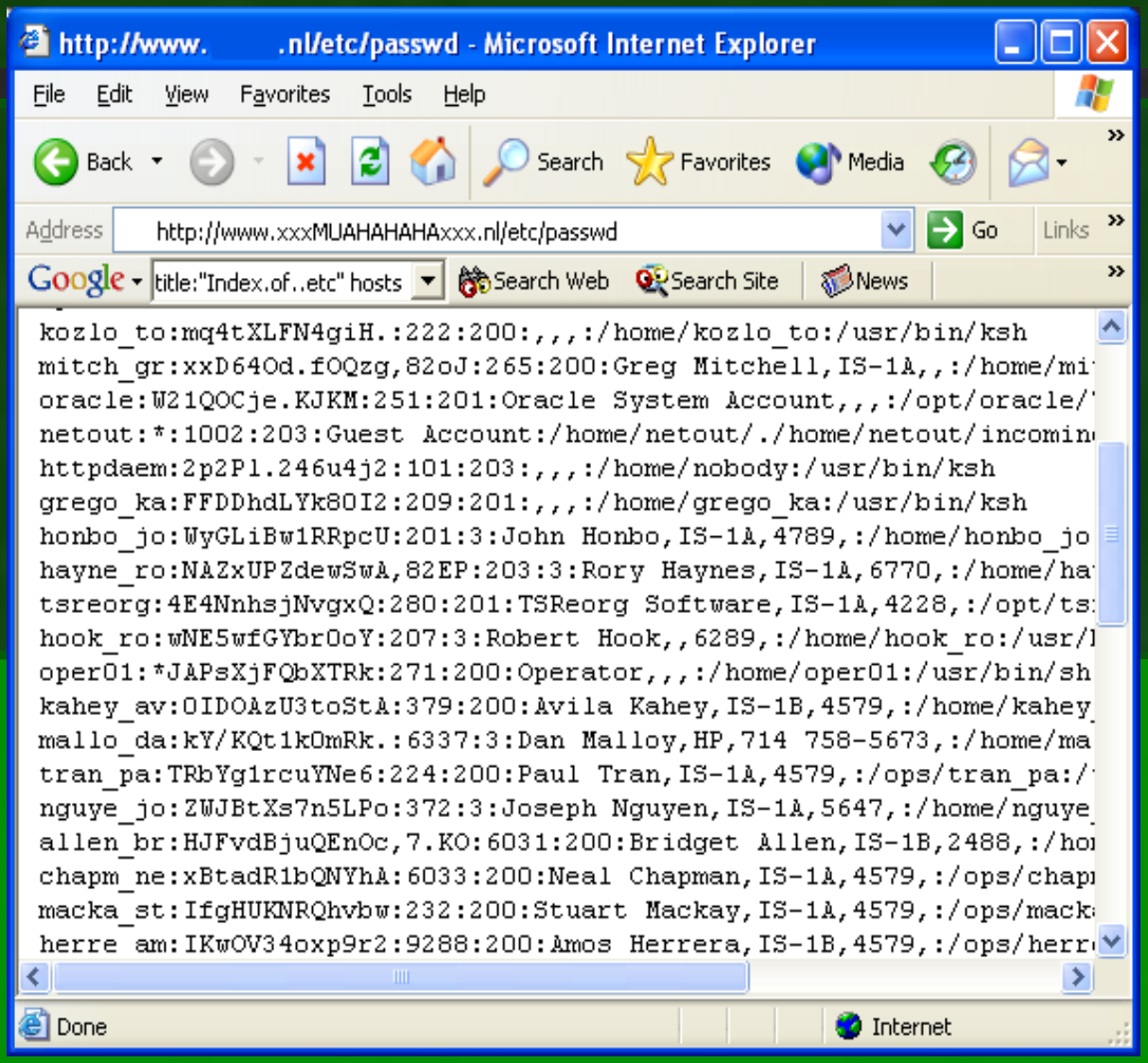
intitle:index.of secring.pgp



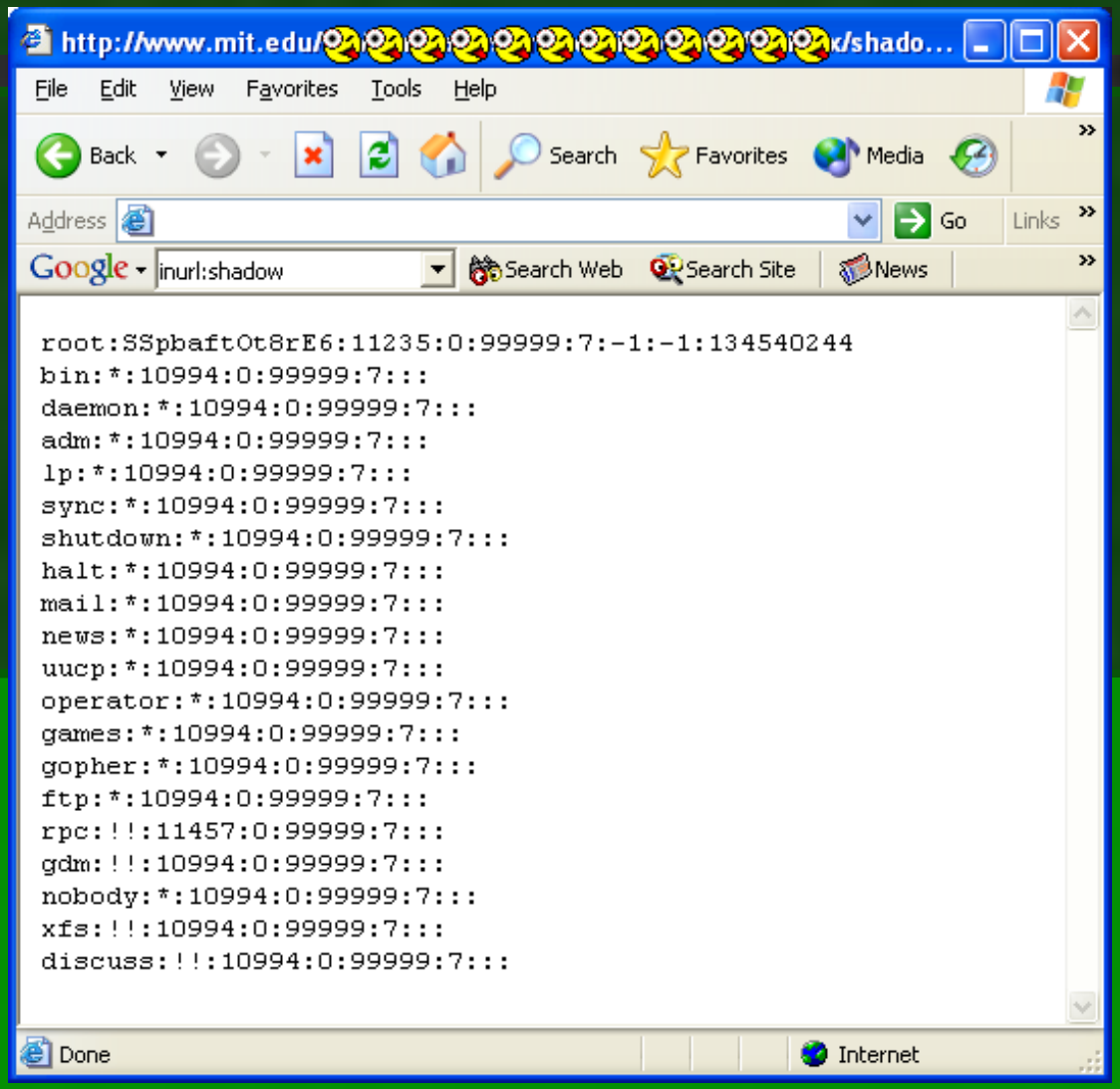
intitle:index.of..etc hosts



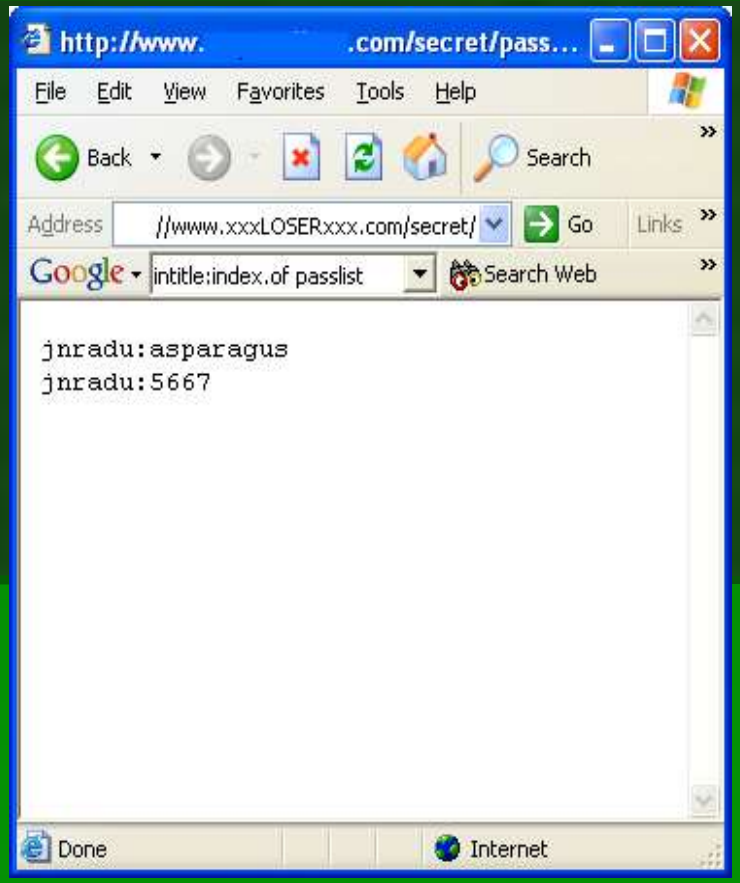
intitle:index.of..etc hosts



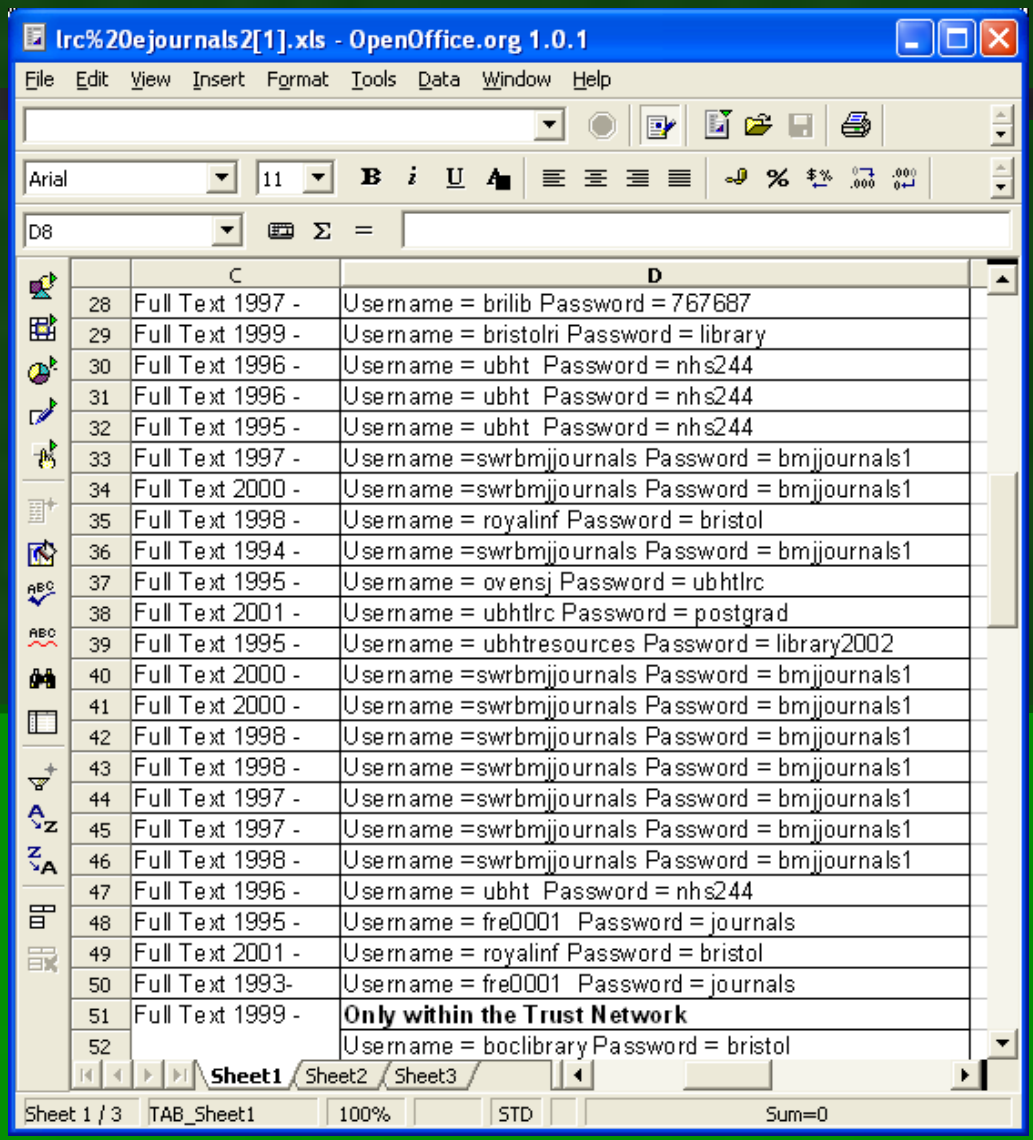
intitle:Index.of etc shadow



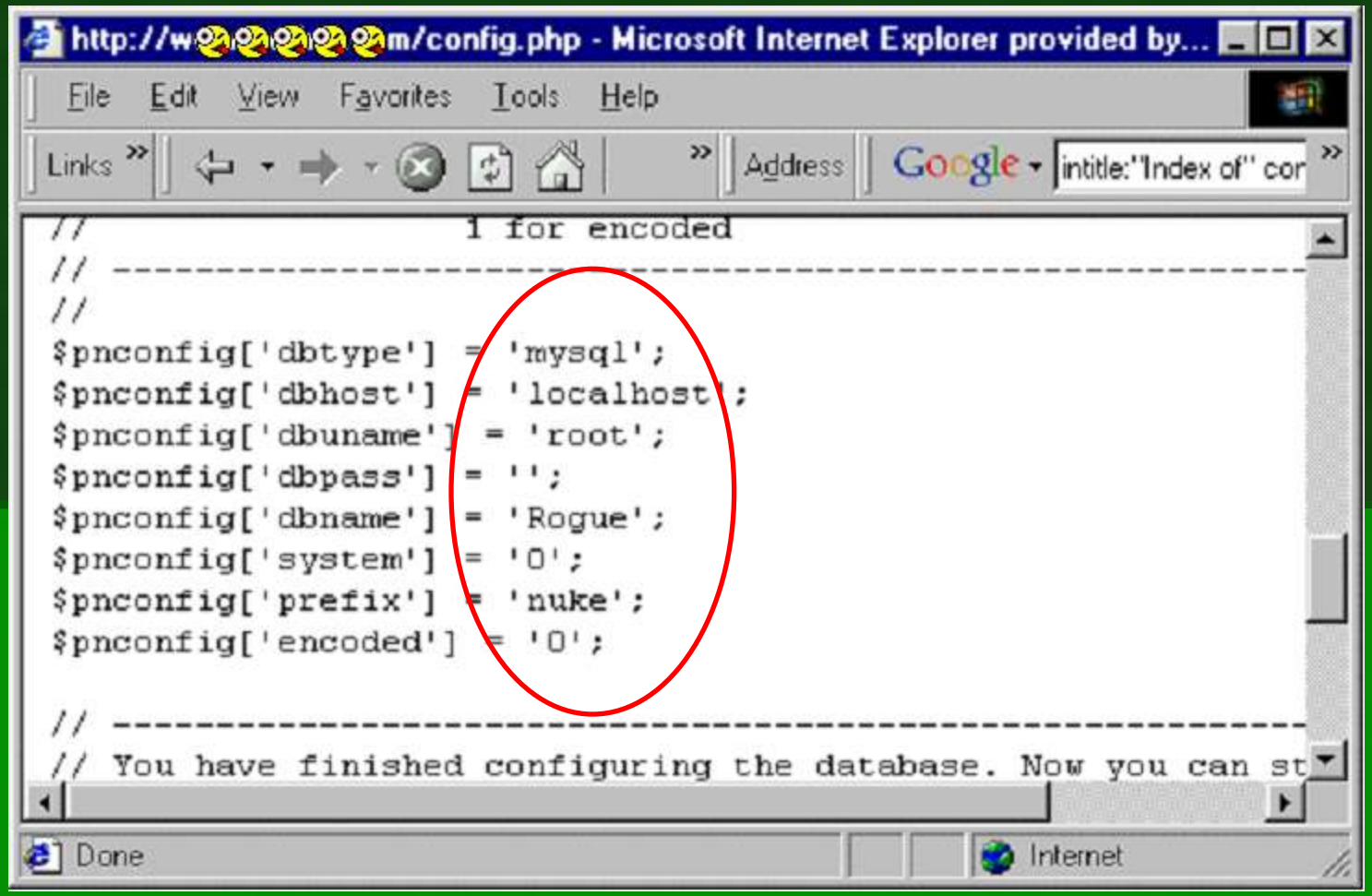
intitle:index.of passlist



filetype:xls username password email

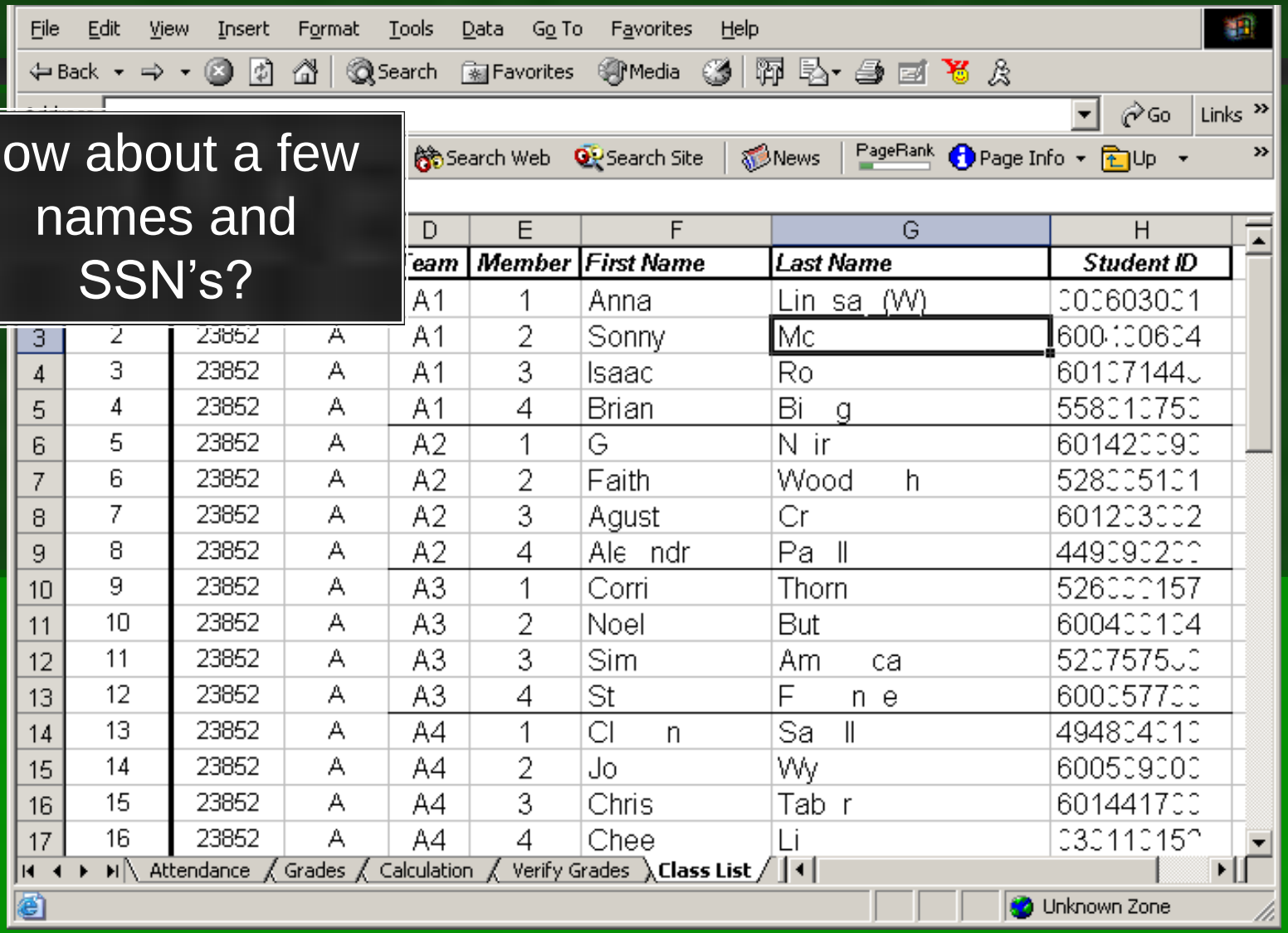


intitle:index.of config.php



social security numbers

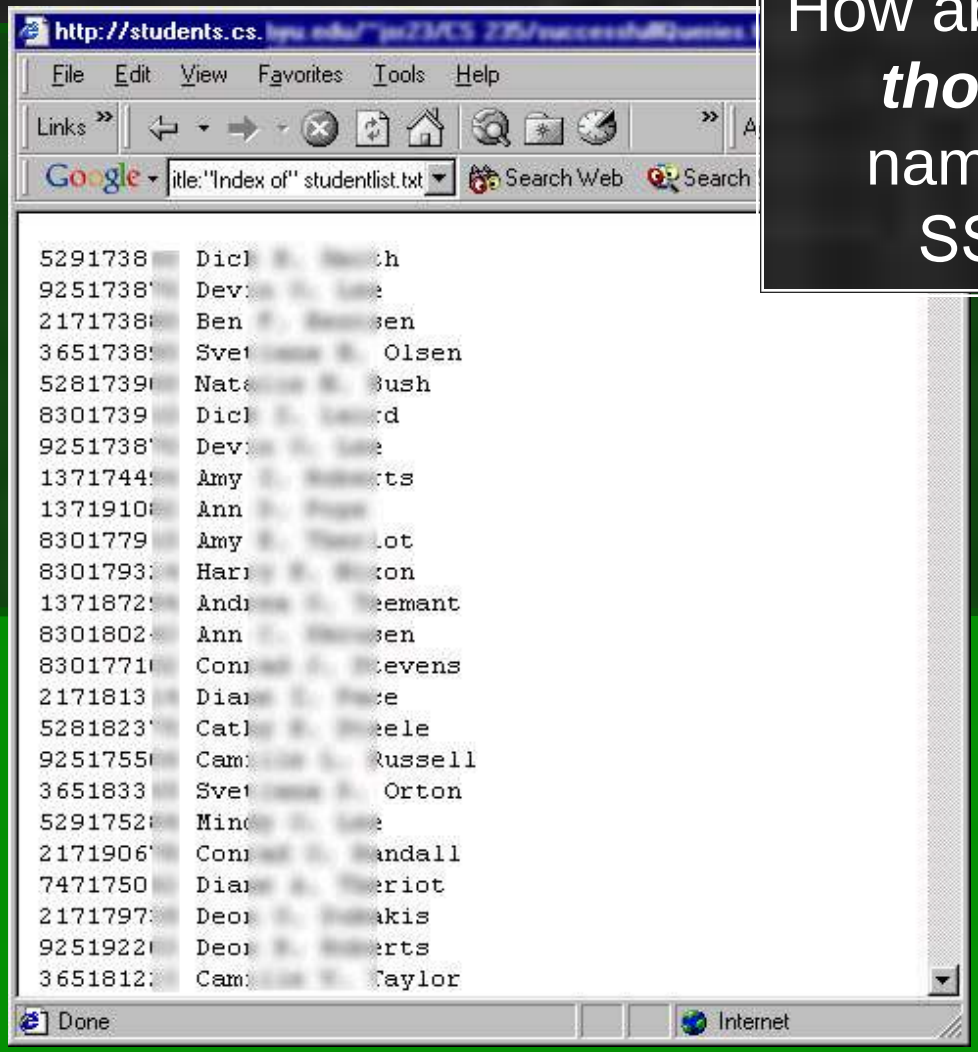
how about a few names and SSN's?



				D	E	F	G	H
				eam	Member	First Name	Last Name	Student ID
				A1	1	Anna	Lin sa (W)	00603001
3	2	23852	A	A1	2	Sonny	Mc	600[illegible]0604
4	3	23852	A	A1	3	Isaac	Ro	601[illegible]7144[illegible]
5	4	23852	A	A1	4	Brian	Bi g	558[illegible]1[illegible]75[illegible]
6	5	23852	A	A2	1	G	N ir	60142[illegible]9[illegible]
7	6	23852	A	A2	2	Faith	Wood h	528[illegible]51[illegible]1
8	7	23852	A	A2	3	Agust	Cr	6012[illegible]3[illegible]2
9	8	23852	A	A2	4	Ale ndr	Pa ll	449[illegible]9[illegible]2[illegible]
10	9	23852	A	A3	1	Corri	Thorn	526[illegible]157
11	10	23852	A	A3	2	Noel	But	6004[illegible]1[illegible]4
12	11	23852	A	A3	3	Sim	Am ca	52[illegible]7575[illegible]
13	12	23852	A	A3	4	St	F n e	600[illegible]577[illegible]
14	13	23852	A	A4	1	Cl n	Sa ll	4948[illegible]4[illegible]1[illegible]
15	14	23852	A	A4	2	Jo	Wy	6005[illegible]9[illegible]0[illegible]
16	15	23852	A	A4	3	Chris	Tab r	601441[illegible]7[illegible]
17	16	23852	A	A4	4	Chee	Li	[illegible]3[illegible]11[illegible]15[illegible]

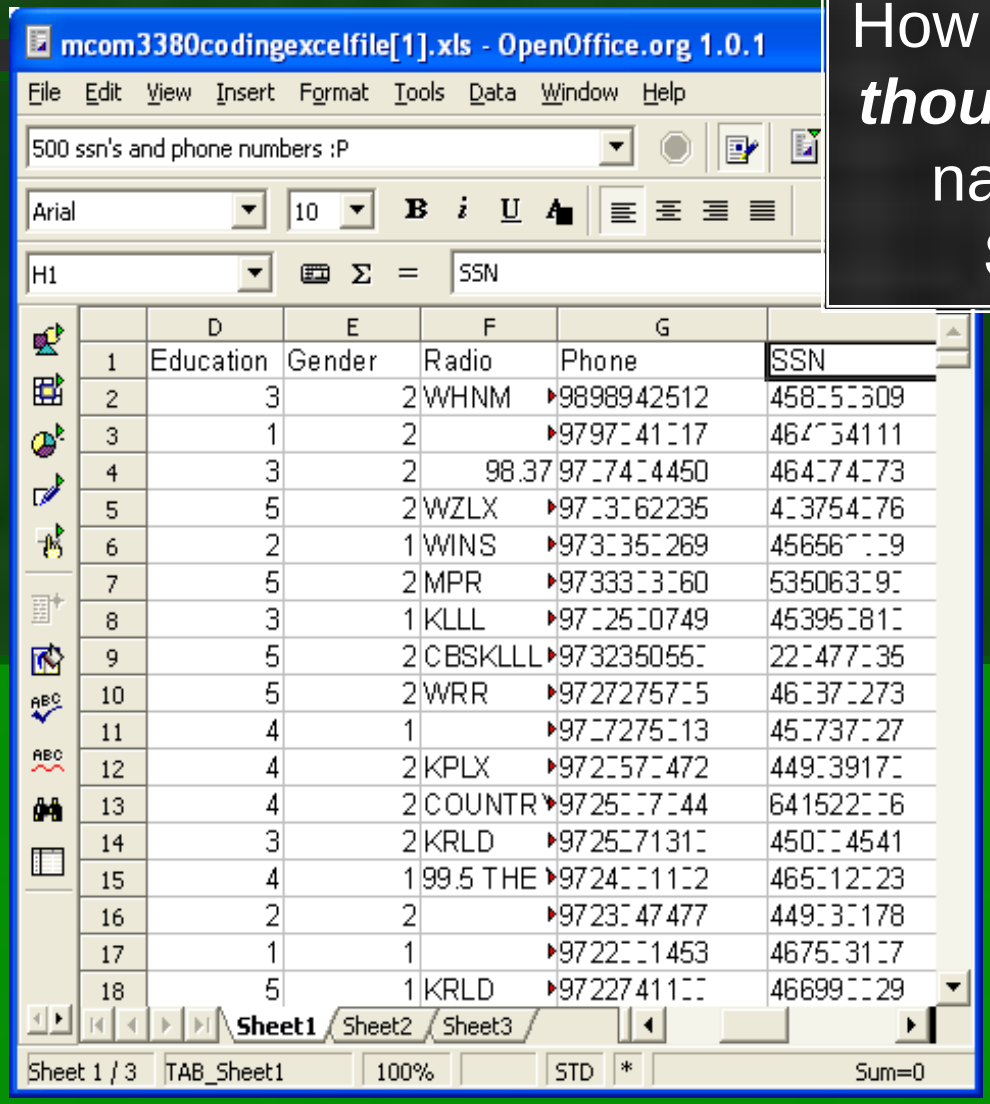
social security numbers II

How about a few ***thousand*** names and SSN's?



social security numbers III

How about a few ***thousand more*** names and SSN's?



	D	E	F	G	
1	Education	Gender	Radio	Phone	SSN
2	3	2	WHNM	9898942512	458□□309
3	1	2		9797□41□17	46□□34111
4	3	2	98.37	97□74□4450	464□74□73
5	5	2	WZLX	973□62235	4□3754□76
6	2	1	WINS	973□35□269	45656□□9
7	5	2	MPR	97333□3□60	535063□9□
8	3	1	KLLL	97□25□0749	45395□81□
9	5	2	CBSKLLL	9732350551	22□477□35
10	5	2	WRR	9727275715	46□37□273
11	4	1		97□7275□13	45□737□27
12	4	2	KPLX	972□57□472	449□3917□
13	4	2	COUNTRY	9725□□7□44	6415222□6
14	3	2	KRLD	9725□71311	450□□4541
15	4	1	99.5 THE	9724□□1112	465□12□23
16	2	2		9723□47477	449□3□178
17	1	1		9722□□1453	4675□31□7
18	5	1	KRLD	97227411□□	46699□□29

Final words...

other google press..

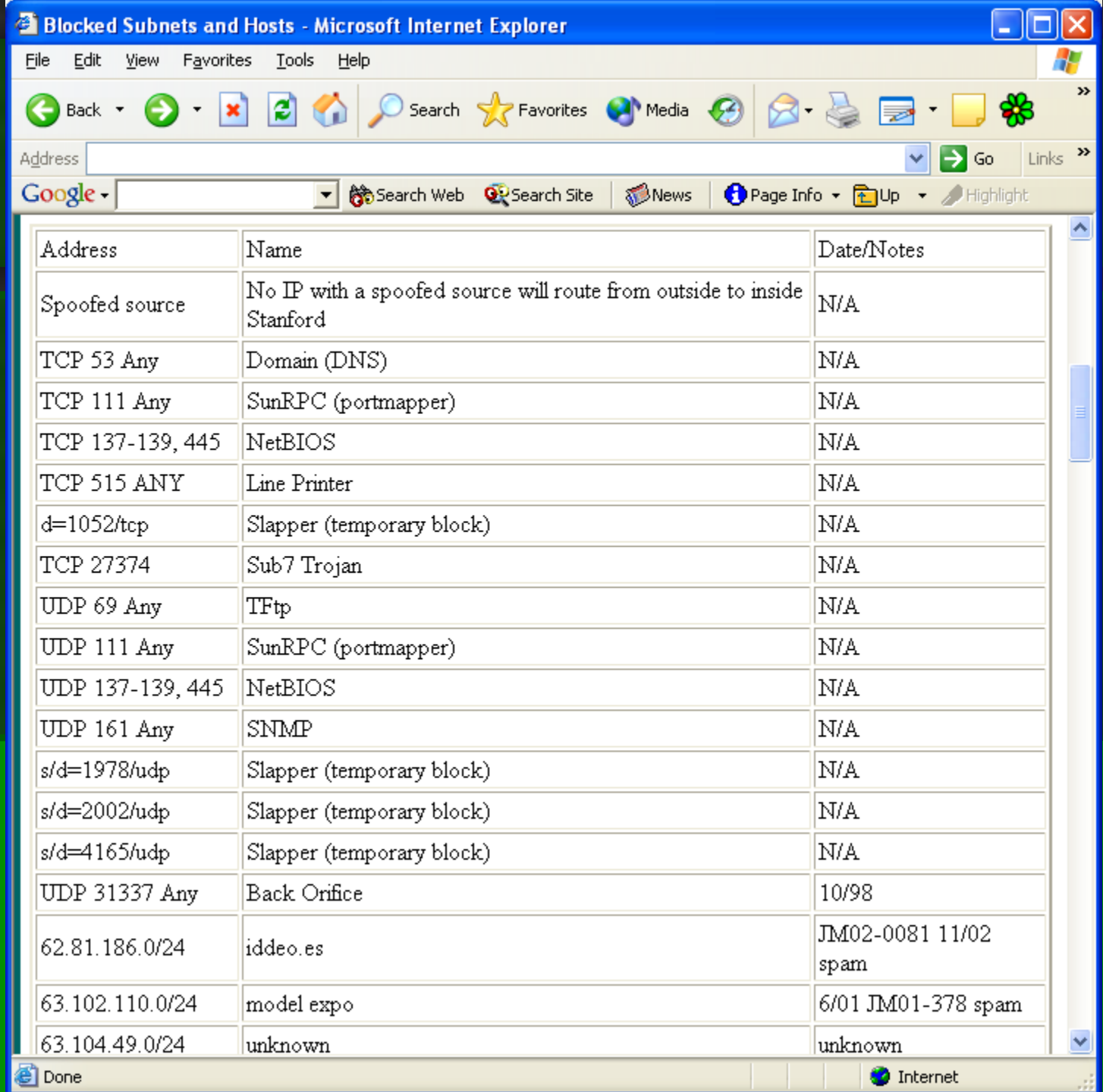
- “Mowse: Google Knowledge: Exposing Sensitive data with Google”
  - http://www.digivill.net/~mowse/code/mowse-googleknowledge.pdf
- “Autism: Using google to hack”
  - www.smart-dev.com/texts/google.txt
- “Google hacking”:
  - https://www.securedome.de/?a=actually%20report (German)
- “Google: Net Hacker Tool du Jour”
  - http://www.wired.com/news/infostructure/0,1377,57897,00.html

EOF

- <plug> Watch googleDorks. </plug>
- Questions?

Contact Me / Get stuff:
http://johnny.ihackstuff.com
johnny@ihackstuff.com

Special Thanks to j3n, m@c, tr3 and p3@nut! =)



Address	Name	Date/Notes
Spoofed source	No IP with a spoofed source will route from outside to inside Stanford	N/A
TCP 53 Any	Domain (DNS)	N/A
TCP 111 Any	SunRPC (portmapper)	N/A
TCP 137-139, 445	NetBIOS	N/A
TCP 515 ANY	Line Printer	N/A
d=1052/tcp	Slapper (temporary block)	N/A
TCP 27374	Sub7 Trojan	N/A
UDP 69 Any	TFtp	N/A
UDP 111 Any	SunRPC (portmapper)	N/A
UDP 137-139, 445	NetBIOS	N/A
UDP 161 Any	SNMP	N/A
s/d=1978/udp	Slapper (temporary block)	N/A
s/d=2002/udp	Slapper (temporary block)	N/A
s/d=4165/udp	Slapper (temporary block)	N/A
UDP 31337 Any	Back Orifice	10/98
62.81.186.0/24	iddeo.es	JM02-0081 11/02 spam
63.102.110.0/24	model expo	6/01 JM01-378 spam
63.104.49.0/24	unknown	unknown