FISCALIA
MINISTERIO PUBLICO DE CHILE

APEC
Asia-Pacific
Economic Cooperation

Challenges of Gathering Evidence from the Internet

Colin Ehren
ACTWG Workshop
June 11-13, 2013
Santiago, Chile

Colin Ehren
candse@gmail.com

Overview

- Search Engines
- Invisible Web
- Social Media
- Compromise Issues & Internet Footprints
- Gathering & Evidencing content.

Colin Ehren
candse@gmail.com

Search Engines

What does Google know?

2007 - Eric Schmidt (Google CEO) estimated Google had indexed roughly 0.004% of the Internet.

July 2008 – Google had identified 1 trillion (1,000,000,000,000) unique URL's.

Everything is on the Internet.

It is estimated that 80% of Open Source information exists in Books, Magazines, Literature and other Media.

Colin Ehren
candse@gmail.com


Google
Ask .com UK
altavista
bing
exalead
GASTA .co.uk
excite
blekko
mirago UK
A9
2GOGO for YOU by YOU
YAHOO! SEARCH
HOTBOT
ooBdoo UK
Scotsmart Scottish Directory
LYCOS
mamma The Mother of All Search Engines
GIGABLAST
Click2Britain
DuckDuckGO
Mojeek United Kingdom
JAYDE The B2B Search Engine
dogpile
Яндекс Yandex
searchers united kingdom
Puruze
searchgo
TEOMA
ASK - DIRectory find fast human reviewed quality sites
iZito
Put My Finger Searching the web, made simple
findtarget
Indexplex
clush
SurfWax enabling knowledge
Norwichportal
Baidu 百度 www.baidu.com
ALLTHEUK .CO.UK
Gimpsy Active Sites for Active People
Qango
Global Search Directory

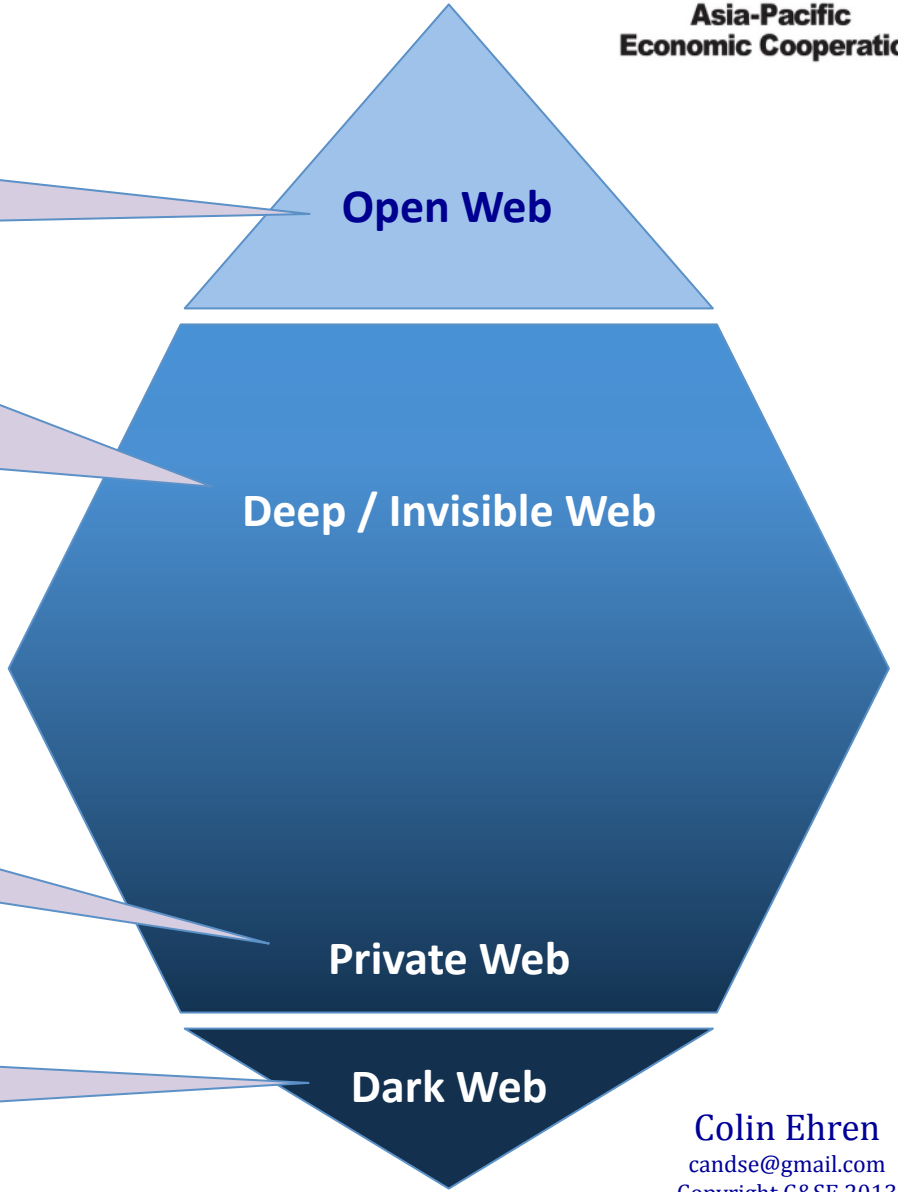
Invisible Web

Static websites (few content changes). Generic and Niche Information. *Easily indexed by Search Engines*

Specialist / niche sites using Robots.txt. Websites with billions of pages stored in databases which are built dynamically for each search. *Search Engines cannot easily index this content*

Company Intranets, Private Websites and Networks. *Search Engines cannot index this content*

Criminal Area of the Internet. *Accessed using specialised software. Search Engines cannot access this content*



Colin Ehren
candse@gmail.com

FISCALIA
MINISTERIO PUBLICO DE CHILE

Invisible Web

APEC
Asia-Pacific Economic Cooperation

July 2001, Michael Bergman produced a white paper for www.brightplanet.com

- 400 to 550 times larger than the World Wide Web.
- 7,500 terabytes of information compared to 19 terabytes on WWW.
- 550 billion documents compared to 30 billion on the WWW.
- 200,000+ deep Web sites.
- 60 of the largest sites collectively contained over 40x the information on the WWW.

2004 Study - identified 330,000+ Deep Web sites.

Has grown almost exponentially since.

Colin Ehren
candse@gmail.com

Social Media

- Austria, Belgium , Croatia, Cyprus, Czech Republic, Denmark, Finland, FYR of Macedonia, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Malta, Norway, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey & UK - **Facebook**, **Youtube**
- Bulgaria – **Facebook**, **VBox7**, **Youtube**
- Estonia – **Facebook**,**Youtube**, **VKontakte**
- France – **Facebook**, **Skyrock**, **Youtube**
- Hungary – **Facebook**, **Iwiw**, **Youtube**
- Latvia – **Youtube**, **Draugiem**, **Facebook**
- Lithuania – **Facebook**, **Youtube**, **One**
- Netherlands – **Facebook**, **Hyves**, **Youtube**
- Poland – **Facebook**, **Youtube**, **Chomikuj**
- Portugal – **Facebook**, **Youtube**, **Twitter**
- Romania – **Facebook**, **Youtube**, **Hi5**

Colin Ehren
candse@gmail.com

Social Media

- China.
  - Qzone, Tencent Weibo, Sina Weibo, RenRen
- Russia.
  - Vkontakte, Youtube, Odnoklassniki, Facebook, Livejournal
- India.
  - Facebook, Youtube, Orkut, Ibibo
- Georgia
  - Facebook, Youtube, Odnoklassniki, VKontakte
- Brazil
  - Facebook, Youtube, Twitter, Orkut
- Iran
  - Cloob, Facenama, Facebook?

Colin Ehren
candse@gmail.com


Create Journals
Update Journals
Journals
Find Users
Random
Read
Search
Create New
Communities
blurty
flickr
myGamma
myWorld
mobile Community
8CarDomain
years old!
facebook
myspace.com
a place for friends
listography, your life in lists
create. share. read.
sconex
Your High School Online
bebo
photobucket
Video and Image Sharing
care2
Make a Difference!
Bolt
eons
downelink
BETA
hi5
kaneva
beta
LIVEJOURNAL
Create an Account
Post to Journal
xanga.com
THE WEBLOG COMMUNITY
YouTube
Faceparty
THE BIGGEST PARTY ON EARTH
FOTOLOG
VampireFreaks.com
Mashable!
social networking 2.0
All
MySpace
YouTube
Bebo
Facebook
Xanga
last·fm
the social music revolution
Gaia
ONLINE
Music forte
friendster
Windows Live Spaces
MOG
VALIDATING MUSIC GEEKS SINCE 2006
BETA
MULTIPLY
SECOND LIFE
VOX
profileheaven
...FUN IN THE AFTERLIFE!
ODEO
myyearbook
YOU'VE GOT FRIENDS!
BETA
www.mymidishare.com
sound
MIDIs and MP3s
imvu
beyond instant messaging
orkut
beta
WAYN
WHERE ARE YOU NOW?
TagWorld
LibraryThing
BETA
StumbleUpon
PICZO
OUTERWORLDS
Where Virtual Reality
And Real People Meet!

Traditional Social Media Tools

- Internet Relay Chat
- Usenet Talk Groups
- Google Groups / Yahoo Groups
- MSN/Skype, Yahoo, AOL Messengers and Chat Rooms
- E-mail
- Dedicated Discussion Forums
- Dating – Muslim Match, Uniform Match, Adult Friend
- Reunion Sites.

Colin Ehren
candse@gmail.com

Compromise Issues

Multiple ways to access the Internet;

- Corporate networked PC's
- Corporate stand-alone PC's
- Re-claimed stand-alone PC's
- Covert / Unattributed stand-alone PC's
- Covert / Unattributed networked PC's
- Working from Home (Stand-alone or Networked)
- Mobile Devices
- Internet Café's

Colin Ehren
candse@gmail.com

Compromise Issues

General Recommendation

All detailed or sensitive Internet research or Open Source Investigations should be undertaken on a covert or unattributed and registered PC, using a covert or unattributed Internet connection.

Why?

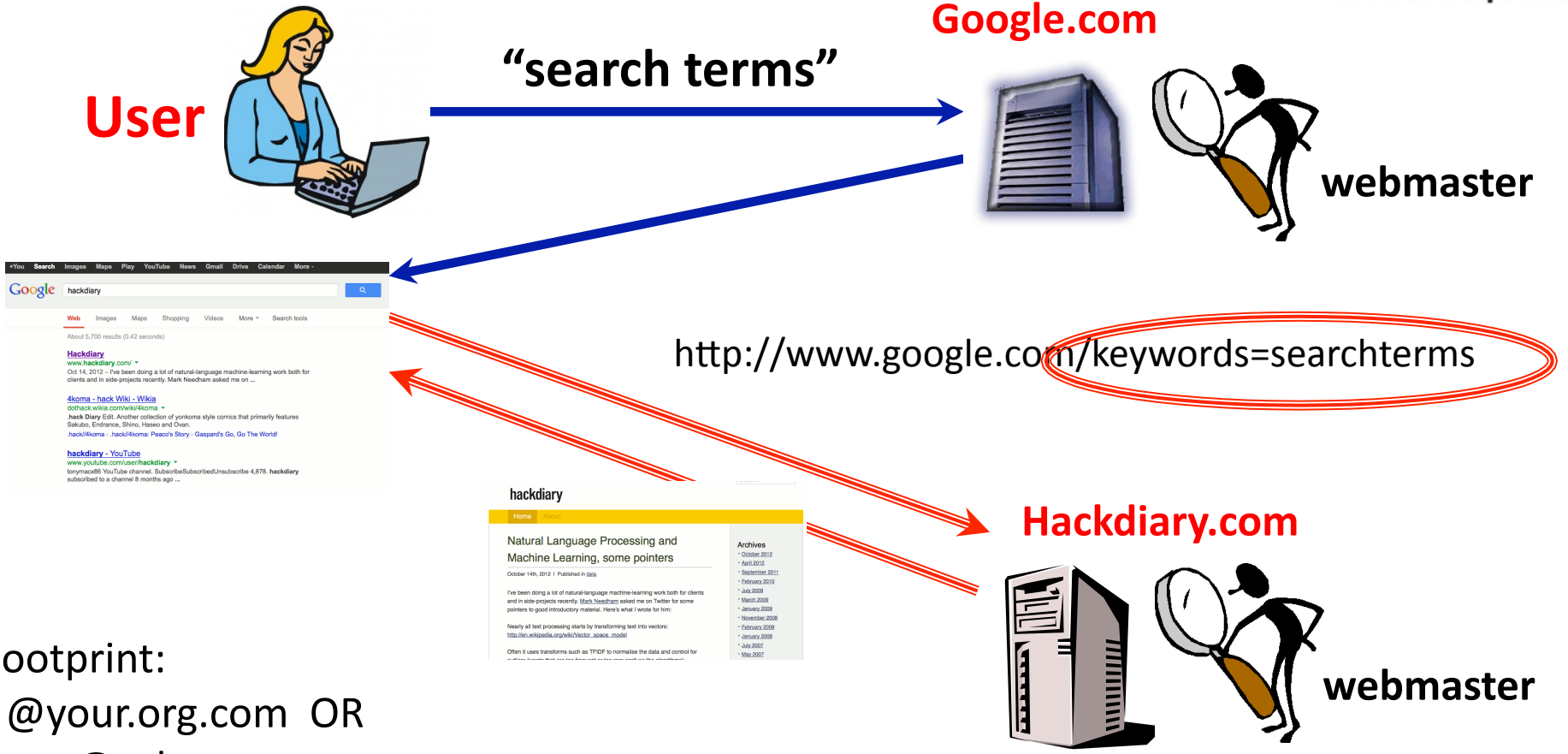
Your Internet Footprint could compromise yourself, your colleagues or an Investigation or Intelligence Operation that your Organisation or a Partner may be engaged in.

Colin Ehren
candse@gmail.com

Compromise Issues

FISCALIA
MINISTERIO PUBLICO DE CHILE

APEC
Asia-Pacific Economic Cooperation



Footprint:
- @your.org.com OR
- you@yahoo.com

google.com webmaster knows your "search terms"

hackdiary.com webmaster knows what "search terms" you used to find them.

Colin Ehren
candse@gmail.com

Compromise Issues

Browser ID : Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0; AWARELockDown2001; .NET CLR 1.1.4322)
Browser:Micro$oft Internet Exploder Version 6.0
Operating System:Windows 2000
Referrer: http://www.google.co.uk/search?q=am+i+secure&hl=en&lr=&start=10&sa=N
Your IP Address : 195.173.172.10
IP Address Behind Firewall
Your hostname: mailgate.exitstrategy.co.uk
Your AS : 2529 ARIN ASN block
Your Country: GB
Your Abuse Contact : abuse@demon.net
Additional Software :

Colin Ehren
candse@gmail.com

Compromise Issues

Home / Forums Index / Hardware and OS Related Technologies / Website Technology Issues

Forum Library : Charter : Moderators: lammert

Website Technology Issues

These terms have been highlighted:
awarelockdown2001 [remove highlighting]

AWARELockDown2001
What user agent is this?

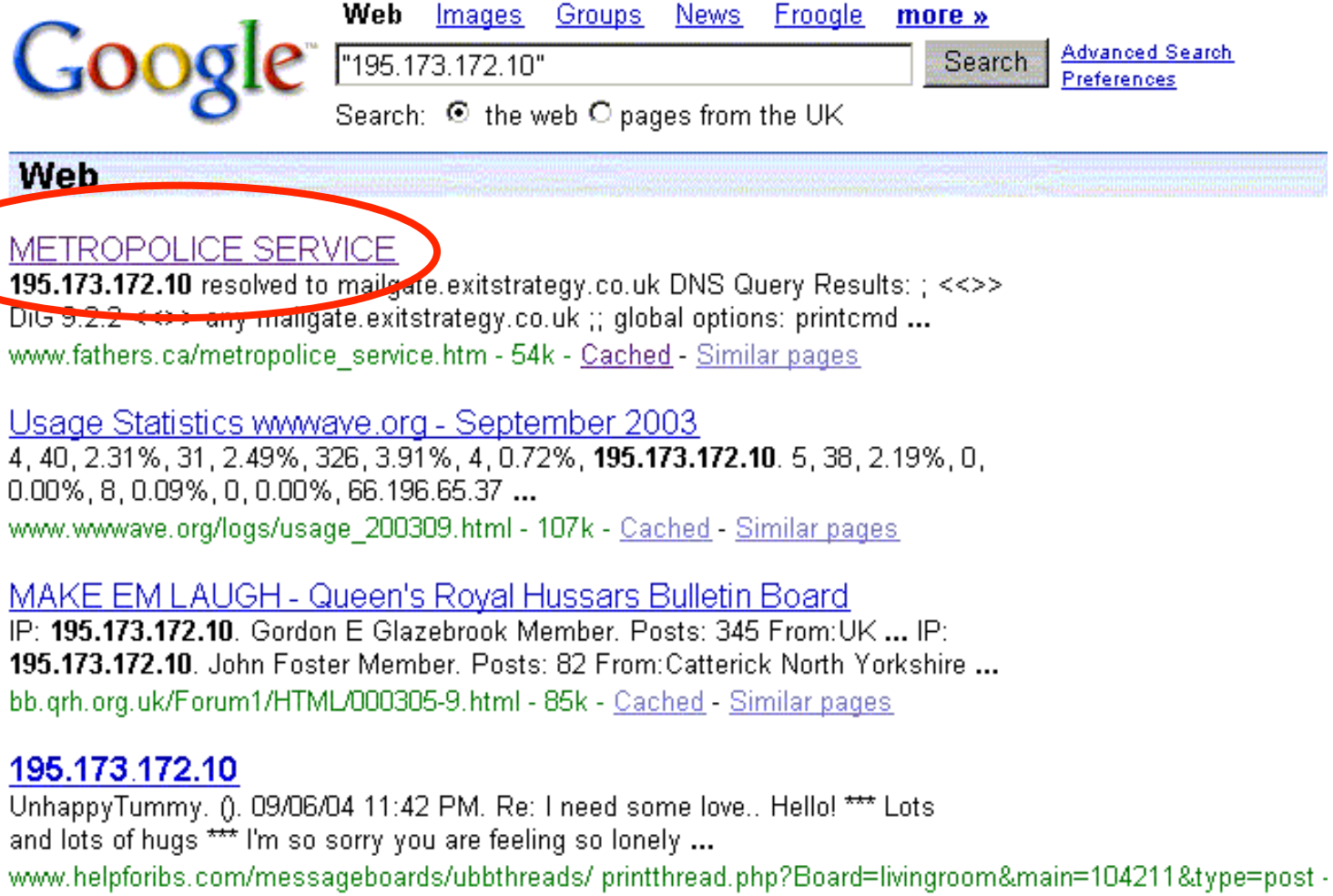
dunne

#:672992

What user agent would "**awarelockdown2001**" be, anyone know? The host IP traces back to ls4689-s0.metpolice.router.uk.quza.net, which is, ah, "interesting".

Colin Ehren
candse@gmail.com

Compromise Issues



Colin Ehren
candse@gmail.com

Compromise Issues

Colin Ehren
candse@gmail.com

Compromise Issues

Network Whois record

Queried whois.ripe.net with "-B 195.173.172.10"...

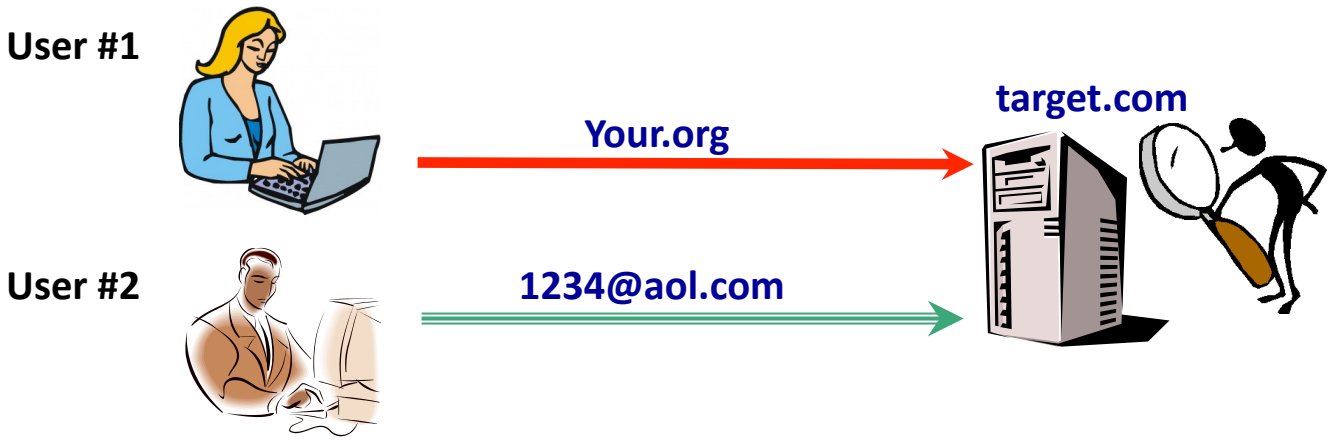
```
% This is the RIPE Whois query server #2.
% The objects are in RPSL format.
%
% Note: the default output of the RIPE Whois server
% is changed. Your tools may need to be adjusted. See
% http://www.ripe.net/db/news/abuse-proposal-20050331.html
% for more details.
%
% Rights restricted by copyright.
% See http://www.ripe.net/db/copyright.html

% Information related to '195.173.172.0 - 195.173.172.15'

inetnum:        195.173.172.0 - 195.173.172.15
netname:        METROPOLICE
descr:          Metropolitan Police Service
descr:          London SW1H
country:        GB
admin-c:        [redacted]
tech-c:         AC2375-RIPE
status:         ASSIGNED PA
mnt-by:         AS2529-MNT
mnt-lower:      AS2529-MNT
mnt-routes:     AS2529-MNT
notify:         hostmaster@demon.net
changed:        hostmaster@demon.net 20030121
source:         RIPE

person:         Alan Cooper
address:        Metropolitan Police Service
address:        London SW1H
phone:          +44-20 8649 3658
notify:         hostmaster@demon.net
mnt-by:         AS2529-MNT
nic-hdl:        AC2375-RIPE
changed:        hostmaster@demon.net 20030121
source:         RIPE
```


Colin Ehren
candse@gmail.com

Compromise Issues

The “parallel surfing” Problem...

- User #1: leaves “your.org” footprints whilst visiting “target.com”
- User #2: leaves “Covert” footprints whilst visiting “target.com”

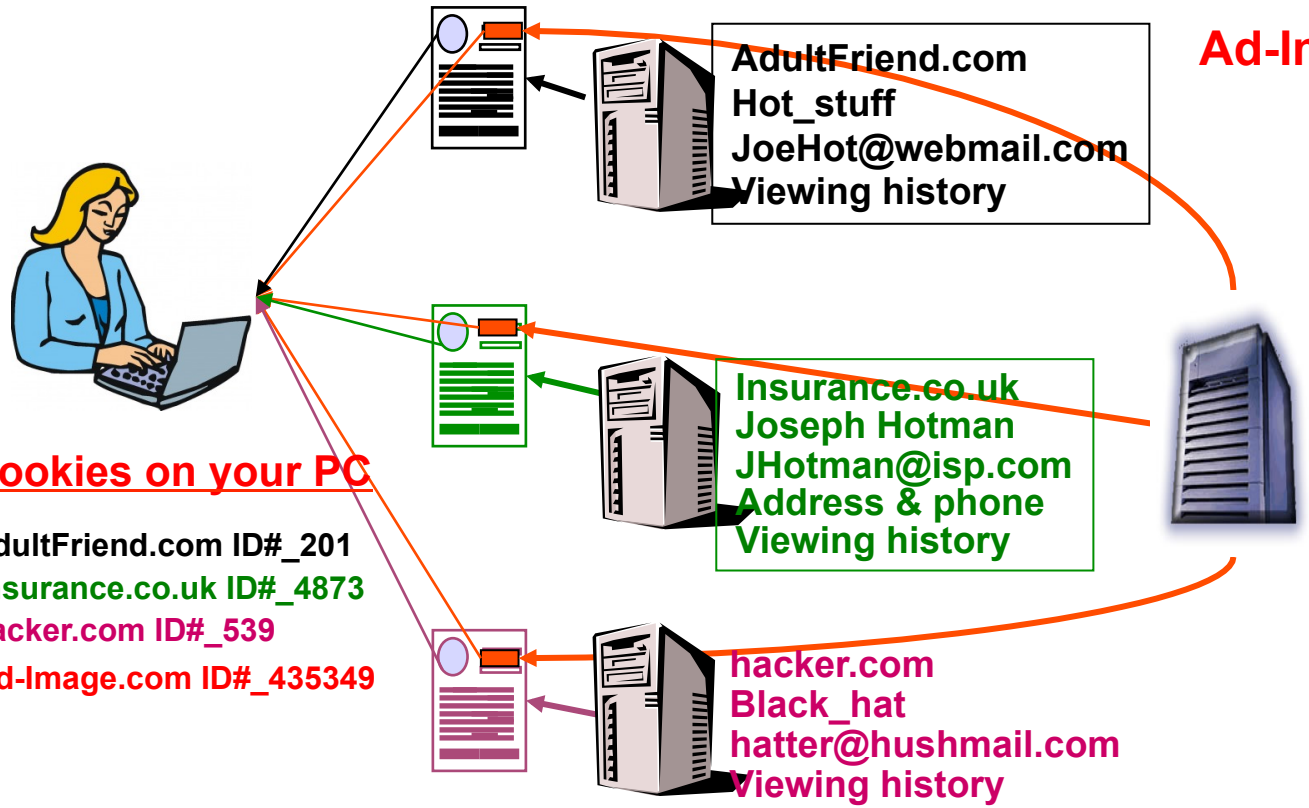


The “Covert” User may now be recognized as an “your.org” visitor.

Colin Ehren
candse@gmail.com

Compromise Issues

Web pages can include images or adverts from third parties.



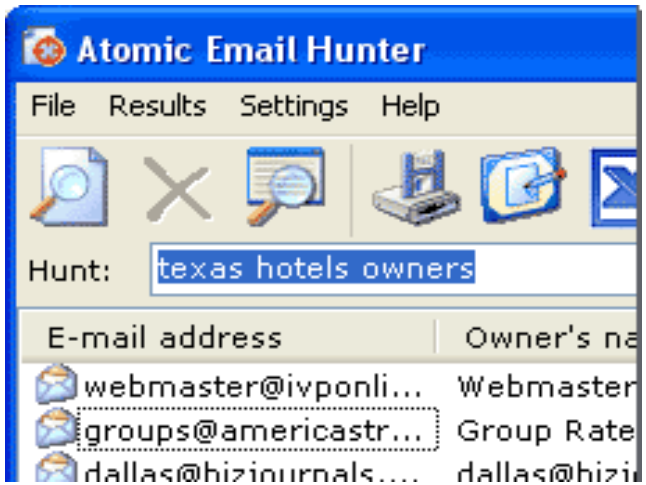
Companies, such as “Ad-Image.com” are able to compile a significant profile on you and your surfing habits, which they are able to trade or sell to their partners or customers.

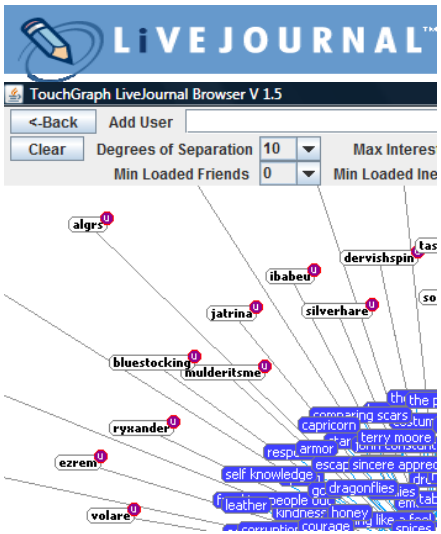
Colin Ehren
candse@gmail.com

Gathering Data

Common Issue – Large amounts of data

- Social Media Tools
- Data Extraction Tools.
- Visualisation Tools





Colin Ehren
candse@gmail.com

FISCALIA
MINISTERIO PUBLICO DE CHILE

Gathering Data

APEC
Asia-Pacific
Economic Cooperation

CutePDF™ ideas for PDF

Video DownloadHelper™
MEDIA DOWNLOAD & CONVERT FIREFOX EXTENSION

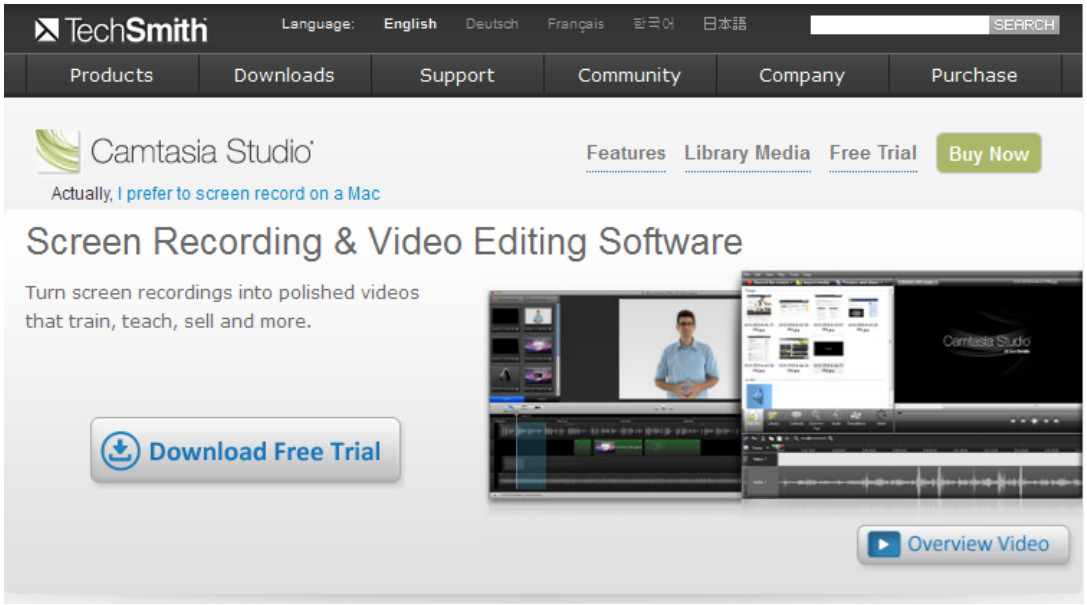
Abduction!
3.0.18

Colin Ehren
candse@gmail.com

Evidencing Data

Prove it or Lose it

All your efforts will be wasted unless you can prove what you found.



Colin Ehren
candse@gmail.com

Evidencing Data

Preserving the Evidential Chain

When saving and moving Data the Investigator must ensure the Evidential Chain is preserved.

- This is done with the aid of an MD5 Hash Extractor
  - A note is made of every number generated in relation to the files saved.
  - The Files are then copied to a CD/DVD Disk, and the Disk 'finalised' so that the files cannot be added to or deleted.
  - Then each file on the CD/DVD Disk is checked with the MD5 Hash Extractor again, and the resulting numbers noted.
  - The resulting numbers generated should be identical.

The CD/DVD disk should now be sealed in an Evidential Envelope or Bag, or placed in an envelope and sealed with an adhesive Exhibit / Evidence Label.

Colin Ehren
candse@gmail.com

FISCALIA
MINISTERIO PUBLICO DE CHILE

APEC
Asia-Pacific
Economic Cooperation

Challenges of Gathering Evidence from the Internet

Colin Ehren
+44-7941-338449
candse@gmail.com