

# Digital Forensics

## Lecture Week 3

# Web Browsers

# Web Tracking

# Browser History

Readings

# Our Focus today

- We think the suspect visited some interesting websites
- How can we check?
- Does a Web browser leave any traces?
- What files does a web browser open?
- What footprint does it leave in memory?
- Can we attribute the browser use to a person?

# Objectives

- To use Web Browsers for Forensics
- To use web tracking for Forensics
- To locate and examine cookie files
- To locate and examine web browser history
- To locate and examine temporary internet files

# Forensics on browser data

- We want to locate evidence of a suspect visit to a website
- We want to locate the files involved
- We want to search the files **post mortem**
- This means **no** browser
- Where are the **chrome** files?
- We need to match a running **process** with its files

# A web browser

- Quite a complex process
- several http conversations on different tcp ports
  - Nowadays encrypted as TLS
- html/xml rendering for the screen display
- Often multiple processes
- tabbed windows create more tcp ports and processes
- several files opened and locked

# Windows Resource Monitor - PG

Disk Activity			0 B/sec Disk I/O
Filtered by chrome.exe, chrome.exe			
Image	PID	File	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Cache\data_4	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Cache\data_2	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Cache\data_3	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Cache\data_1	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\History-journal	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Cookies-journal	
chrome.exe	3536	C:\\$LogFile (NTFS Volume Log)	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\9F97.tmp	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Favicons-journal	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\871C.tmp	
chrome.exe	3536	C:\System Volume Information\{153b4464-3412-11e5-8000-902b34d9c2fb}\{3808876b-c1	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Favicons	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\History	
chrome.exe	3536	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Cookies	

# SysInternals Process Explorer - PG

chrome.exe		1608 Google Chrome	Google Inc.
<			
Type	Name		
File	\Device\DeviceApi		
File	C:\Windows\SysWOW64\en-US\MMDevAPI.dll.mui		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Web Data		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Visited Links		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\lockfile		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\History		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Service Worker\Dat...		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Service Worker\Dat...		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Top Sites		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Application Cache\...		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Service Worker\Dat...		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Service Worker\Dat...		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\Favicons		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\File System\Origins\...		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\QuotaManager		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\DATABASES\Databas...		
File	C:\Users\Graham\AppData\Local\Google\Chrome\User Data\Default\GCM Store\LOG		

# Your Firefox Profile - 1

- **Bookmarks, Downloads and Browsing History:**

The *places.sqlite* file contains all your Firefox bookmarks and lists of all the files you've downloaded and websites you've visited.

- **Passwords:**

Your passwords are stored in the *key4.db* and *logins.json* files..

- **Site-specific preferences:**

- **Search engines:**

- **Personal dictionary:**

- **Autocomplete history:**

The *formhistory.sqlite* file remembers what you have searched for






# Your Firefox Profile - 2

- **Cookies:** Cookies are all stored in the *cookies.sqlite* file.
- **DOM storage:** DOM Storage is designed to provide a larger, more secure, and easier-to-use alternative to storing information in cookies. Information is stored in the *webappsstore.sqlite* file for websites
- **Extensions:**  
The *extensions* folder, if it exists, stores files for any [extensions](#) you have installed..
- **Stored session:** The *sessionstore.jsonlz4* file stores the currently open tabs and windows.


# Web forms in Firefox

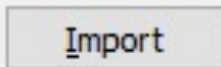
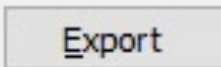
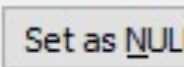
Database Structure **Browse Data** Edit Pragma's Execute SQL

Table: **webappsstore2**   **New Record**

	Attrit	originKey	scope	key	value
	<b>kees</b>		Filter	Filter	Filter
1		ua.moc.kees.www.:https:443	ua.moc.kees.www.:https:443	impression-tracking-logger	"[]"
2		ua.moc.kees.www.:https:443	ua.moc.kees.www.:https:443	job-tracking	"[]"
3		ua.moc.kees.www.:https:443	ua.moc.kees.www.:https:443	lastSearchWhere	"All Sydney NSW"
4		ua.moc.kees.www.:https:443	ua.moc.kees.www.:https:443	tealium_timing	{"domain":"www.seek

Edit Database Cell

Mode: Text 

`{"domain":"www.seek.com.au","pathname":"/digital-forensics-jobs/in-All-Sydney-NSW","query_string":"","timestamp":1531973447621,"ins":18,"connect":41,"response":`

# Identifying a Web Client

- We are given a packet capture file (pcap)
- We are told to look for forensic evidence
- The first step is to identify the Web Client and the OS
  - Preferably before we examine the pcap file
- How can we do this?
- One way is to use the browser to access a special device fingerprinting website
  - (After we acquire and save an image of the device)

# Identifying a Web Client #2

- The http request string is informative
- The detail may identify a suspect's PC
  - Even with inprivate browsing
- This process is called **device fingerprinting**
- It can be used to regenerate deleted cookies.
- We need to access the PC to confirm identity.
- <https://www.browserleaks.com/>
- <https://coveryourtracks.eff.org/learn>

Browser Characteristic	bits of identifying information
User Agent	10.14
HTTP_ACCEPT Headers	9.55
Browser Plugin Details	15.38
Time Zone	7.15
Screen Size and Color Depth	4.5
System Fonts	19.08
Are Cookies Enabled?	0.43
Limited supercookie test	0.96

# The BrowserLeaks website

- Web Browser Fingerprinting
- Displays personal identity data leaked when the suspect surfs the Internet
- <https://www.browserleaks.com/>

Shows Your IP :

IP Address

Host Name

IP Address Location :

Country

State/Region

City

Organization

ISP

AS Number

Timezone

Local Time

Latitude/Longitude

# Web Servers

- Lots of tracking software
  - Linked to Social Media and Search Engines
  - Analysis by [builtwith.com](https://builtwith.com)
- Lots of Technologies
  - Analysis by [W3Techs.com](https://w3techs.com)
- Very detailed logs of visitors
  - Often Apache2
  - `/var/log/apache2/access.Log`



Find out what websites are  
Built With

# Builtwith analysis of UTS.edu.au

Log In · Signup for Free


**builtwith** Tools ▾ Features ▾ Plans Customers Resources ▾

[Home](#) / [uts.edu.au Technology Profile](#)

## UTS.EDU.AU

Technology Profile Detailed Technology Profile Meta Data Profile Relationship Profile Redirect

### Analytics and Tracking [View Global Trends](#)

 **iGoDigital**

[iGoDigital Usage Statistics](#) · [Download List of All Websites using iGoDigital](#)

Analyzes individual shopper behavior and provides personalized product recommendations. Now owned by ExactTarget.

Conversion Optimization

Analytics and Tracking  
View Global Trends  
iGoDigital

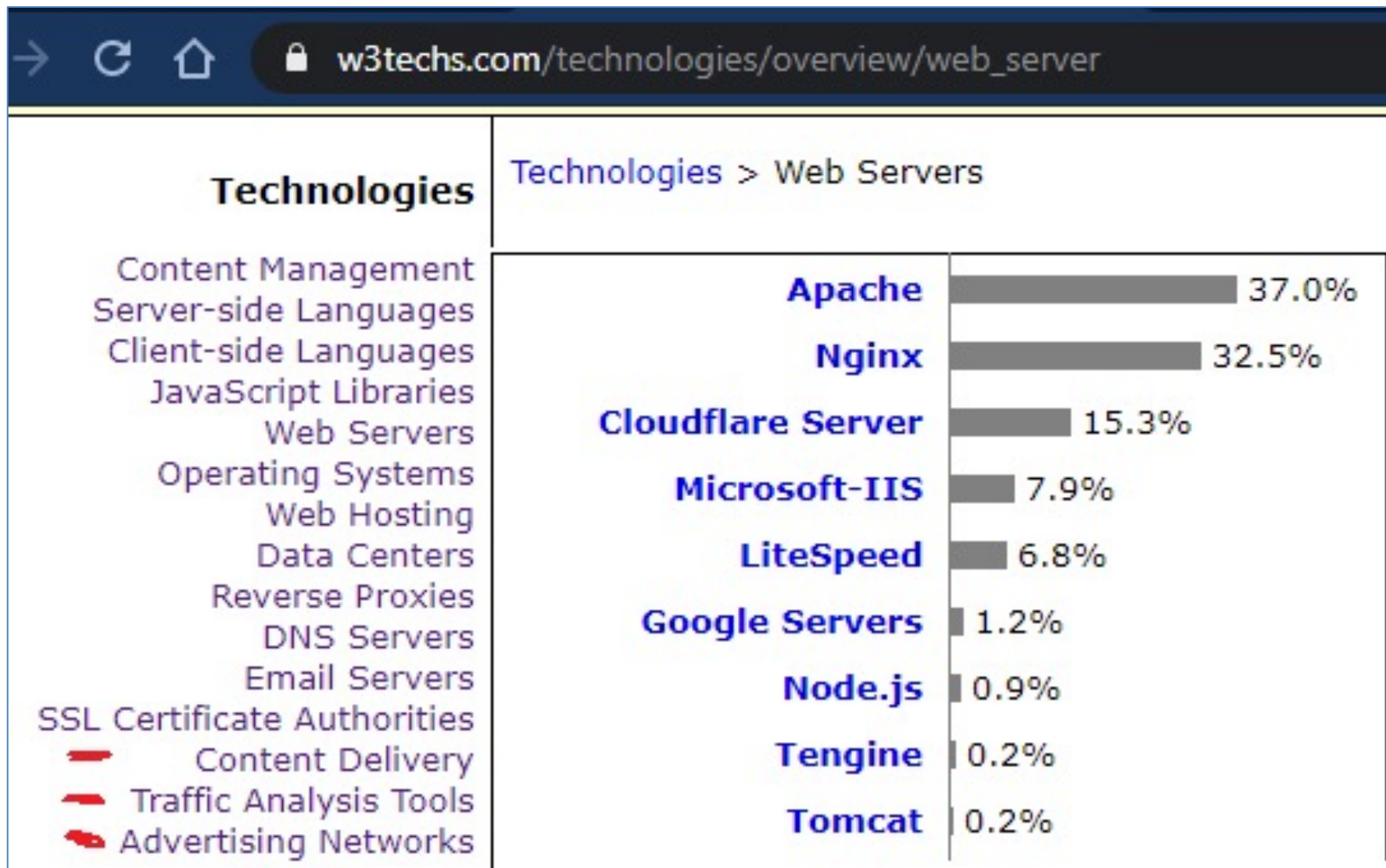
Audience Measurement  
CrazyEgg

Site Optimization  
Google Optimize 360  
New Relic

Application Performance  
Google Analytics  
DoubleClick Floodlight

Conversion Optimization  
Google AdWords Conversion  
Facebook Signal  
Facebook Pixel  
LinkedIn Insights  
Baidu Analytics







# Objectives

- To use Web Browsers for Forensics
- To use web tracking for Forensics
- To locate and examine cookie files
- To locate and examine web browser history
- To locate and examine temporary internet files

# User Tracking

- A web server needs to track a web client
- by ip address
- by the http referer tag
- by a cookie saved on the target
  - http cookie, web cookie, browser cookie
    - three names for the same thing
    - Cookies have gone out of fashion as insecure.
- by embedded code on the web page
  - Tracking on the website
  - Using third parties to do remote tracking



# Web Analytics

- Web Page Tracking is also used by Advertisers
- A lot of time, money and effort goes into tracking
- See this week's **Journey Mapping** reading.
- See this week's **Web Analytics** reading.
- The **user profile** is an important marketing tool
  - See **social networking** websites

# User tracking for profit

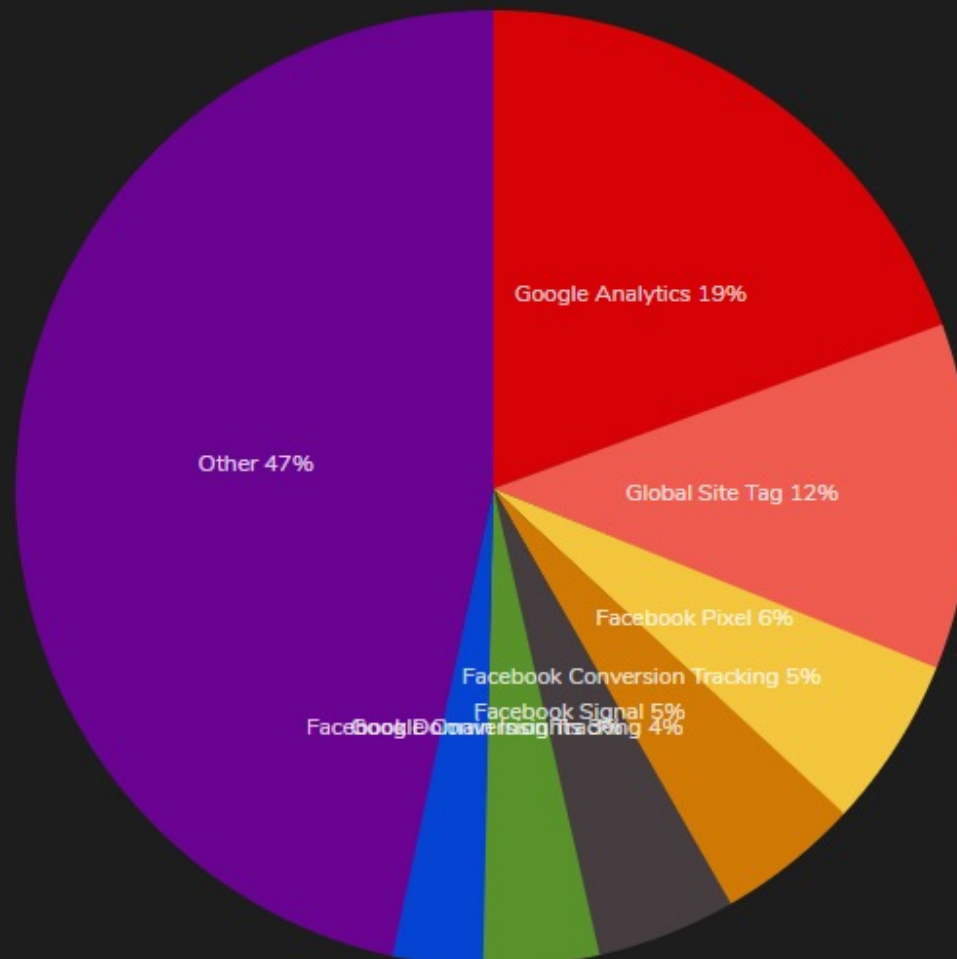
- If you analyse now a user got to your website and purchased an item you will be able to entice many more too follow.
- By combining your visits to many websites many times, the analytics company will be able to understand and influence your future behaviour.
- We can use these results for forensics

# Journey Mapping

- Follow the target as they navigate the internet and end up purchasing an item.
- Visits to Social Media and Search Engines are **converted** into a visit to the website
- Build a timeline based on the time stamp of every stage.
- The person is called the **actor**.
- The **scenario** identifies the expectations of the actor
- **Opportunities** arise to entice further purchases

# Analytics Usage Distribution in the Top 1 Million Sites

Distribution for websites using Analytics technologies



# Google Analytics

- Two methods using two different JavaScript libraries
  - **ga.js** drops a set of `_utm` cookies.
    - Good forensics but now out of fashion due to privacy.
  - **analytics.js** drops two cookies
    - ga            lifetime 2 years
    - gid           lifetime 1 day
- poor forensics but popular as little private detail is visible

# UTM Cookie Formats

- Google bought Urchin Software in 2005
- Google Analytics (GA) uses UTM
- Urchin Tracking Module (UTM) codes
- **UTMA** tracks dates and visits
- **UTMB/C** indicate session expired
- **UTMZ** is for tracking the user
  - Referrer, keyword, ad campaign, etc
- More detail in readings

UTMA - The Visitor Identifier

UTMB - 30 Minute session identifier

UTMC - On Exit session identifier

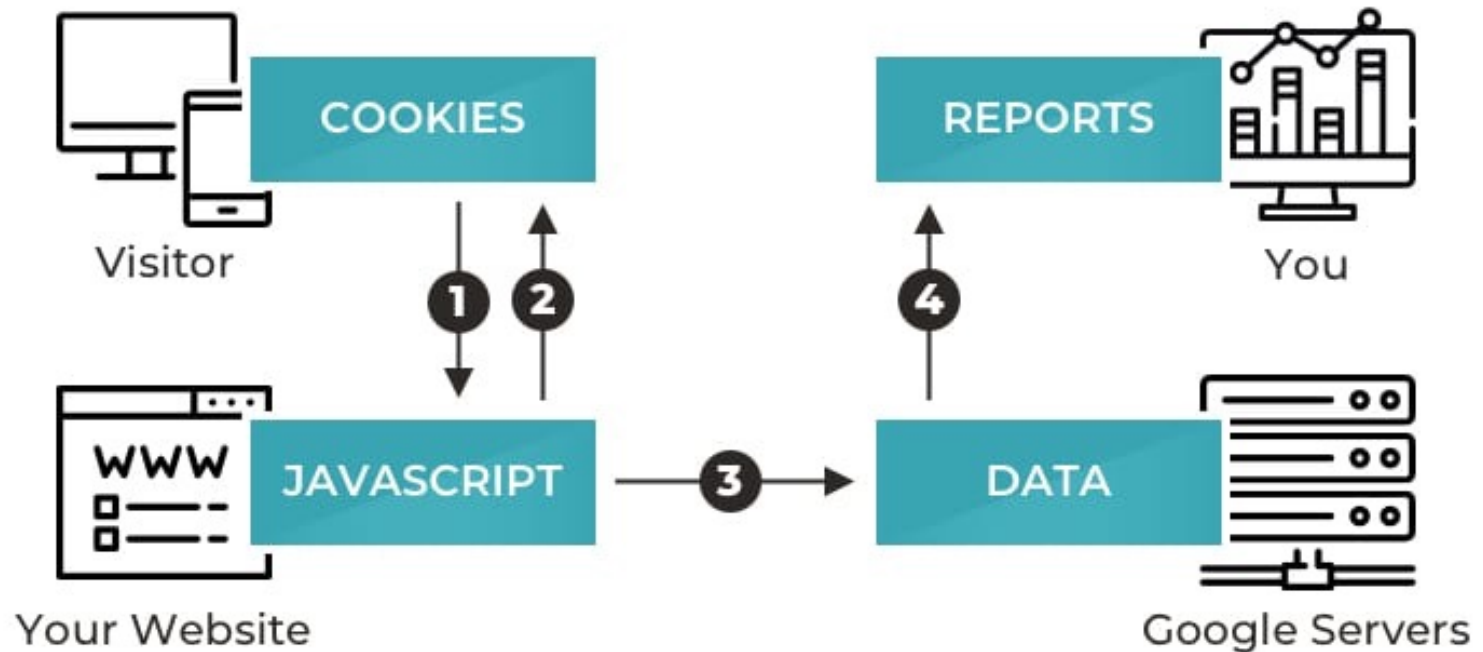
UTMV - Custom Variable Cookie

UTMZ - Visitor segmentation



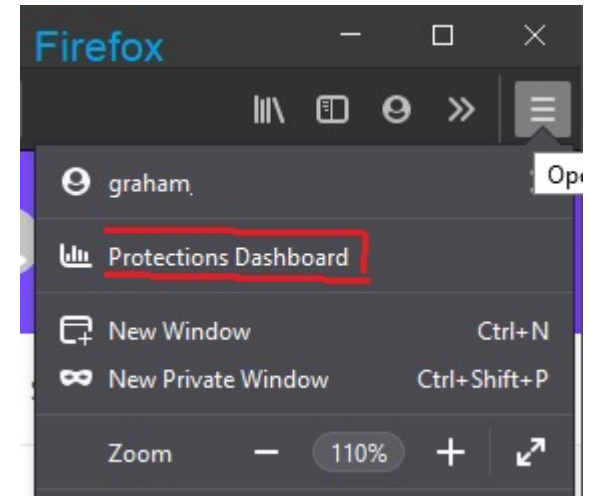
# GA data flow

## How Google Analytics works

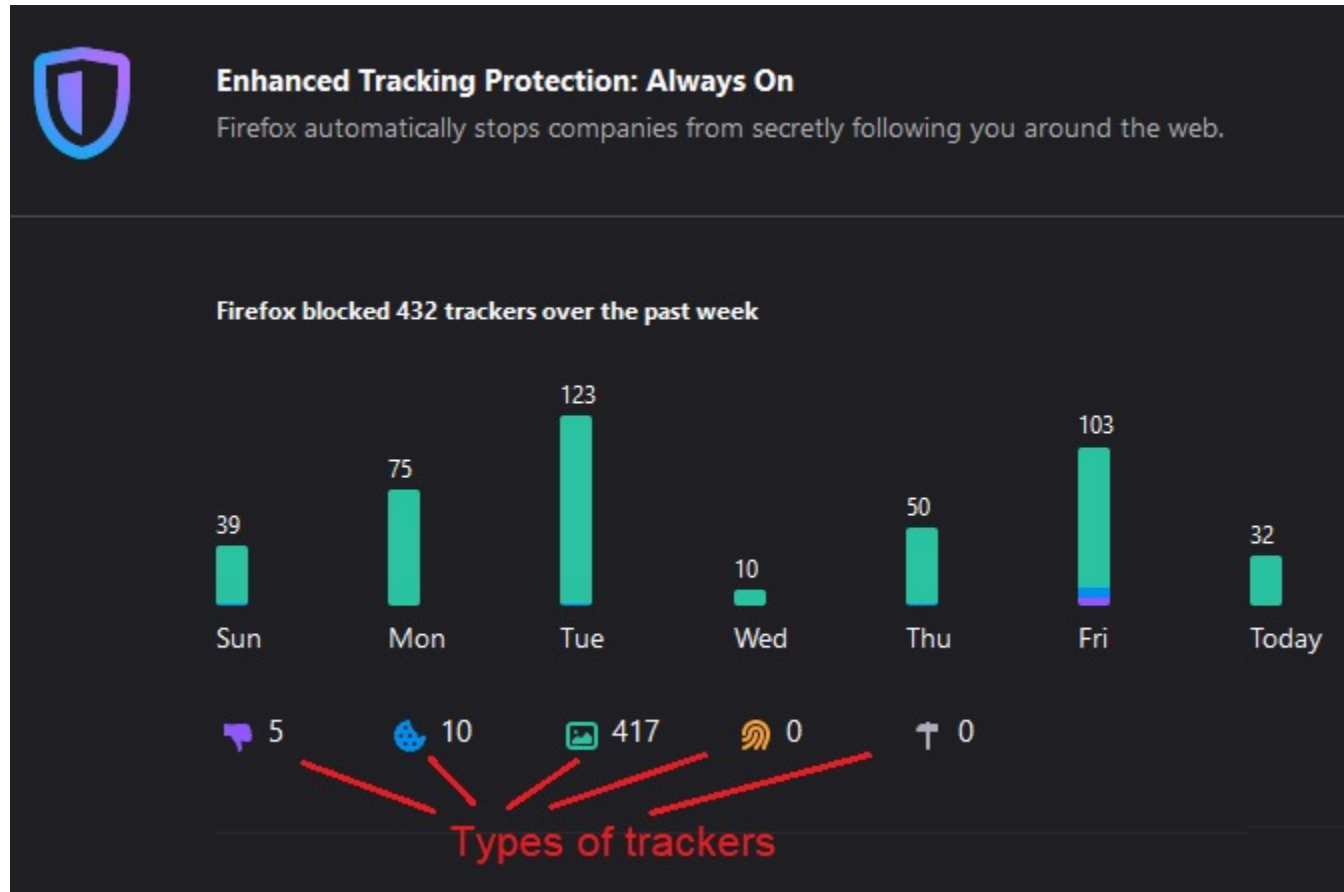


# Firefox Protection Dashboard

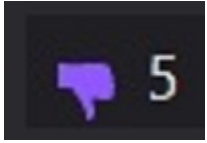
- Firefox can block trackers with its Protections Dashboard
- Provides tracking protection for a desktop or mobile device
- Provides analytics on the blocked trackers
- Uses an addon from Disconnect
  - <https://disconnect.me/>
- There are more than 2500 tracking websites



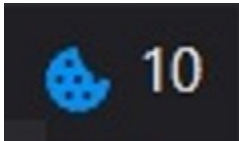
# Tracking Analytics



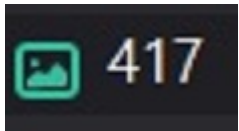
# Tracker types



- Social Media
  - Social networks place these trackers on other websites to follow you.



- Cross Site Tracking cookies
  - Analytics companies place these trackers to follow you



- Tracking Content
  - The website loads external content targeted at you



- Fingerprints
  - Collect browser settings to identify you



- Cryptominers – use your device to mine money

# Objectives

- To use Web Browsers
- To use web tracking
- To locate and examine cookie files
- To locate and examine web browser history
- To locate and examine temporary internet files

•



# A forensic traces example

- 1988: Prosecutors upgraded the charges against a suspect to murder on the basis of evidence of premeditation found on his office computer.
- Tech Support had done a routine investigation during an upgrade and checked the suspect's cookies
- They found websites with cookies that were referred from Google.
- The search terms included kill+spouse, accidental+death, smother, poison, homicides and murder

# Reason for Cookies

- Web Pages are transferred over the Internet using HTTP (Hypertext Transfer Protocol)
- HTTP is **Stateless** so we need some method of saving viewer choices
- Cookies save state on the client as a file on disk
  - *Cookies are small and fast (lightweight)*
- Cookies are also used to save state for session key negotiation (Wireless and VPNs)

# More reasons for Cookies

- Personalisation
  - The server remembers what you liked on your last visit
- Data Capture
  - The server remembers what you asked for
- Sales tracking using a shopping basket
  - cannot be undone (do not click the back button!)
- Authentication
  - no need for a password for a repeated login



# Cookie forensics

- Deleting cookies will disable many websites
- Modern web sites only drop very basic cookies
- Viewing Cookies is a useful forensic tool
  - websites visited
  - actions taken/pages visited
  - date of first visit
  - date of last visit
  - number of visits



# Setting cookies

- The web client asks for a web page using http
  - GET /index.html HTTP/1.1
- The web server **sets a cookie** when it replies
  - HTTP/1.1 200 OK
  - Set-Cookie: name=value
- The cookie is returned each time the page is accessed
- The server keeps a log of cookies to track viewers
- viewer=ip address+referrer+cookie
- See Readings, http cookies

# Set-Cookie

- Server has code to set the cookie

```
<?php
$expire=time()+60*60*24*90;
setcookie("user", "CEH Student",
$expire);
?>
```

- Browser asks for the server page

```
GET /login_p.php HTTP/1.1
Accept: text/html, application/xhtml+xml
Accept-Language: en-AU,en-GB;q=0.8,en-GB;q=0.7,en;q=0.6
User-Agent: Mozilla/5.0 (compatible;
Accept-Encoding: gzip, deflate
Host: 10.10.10.38:8080
Connection: Keep-Alive
Cookie: user=CEH+Student
```

- Server sets the cookie

```
HTTP/1.1 200 OK
Date: Sun, 21 Apr 2013 20:48:10 GMT
Server: Apache/2.2.14 (Ubuntu)
X-Powered-By: PHP/5.3.2-1ubuntu4.18
Set-Cookie: user=CEH+Student; expire:
Vary: Accept-Encoding
Content-Encoding: gzip
Content-Length: 317
Keep-Alive: timeout=15, max=100
Connection: Keep-Alive
Content-Type: text/html
```

- Cookie file appears on client PC

– the dates are in cookie format

userCEH+Student10.10.10.38/1536190769587230311818374128194030293839\*

# Cookie types

- Session cookie
  - no expiry date, deleted by the browser when the session ends
- Persistent cookies (tracking cookies)
  - expiry date in the future
- Secure cookie
  - sent encrypted using https
- Third party cookies
  - set from a different URI domain

# GA \_gid cookie example

- GA1.3.1180290148.1531954273
- GA = Google Analytics
- 1.3 = Version
- 1180290148 = random number session ID
- 1531954273 = Unix timestamp (visit)

# Unix Date format



The screenshot shows the EpochConverter website. The browser's address bar displays "Secure | https://www.epochconverter.com". The website's logo, a clock icon, is followed by the text "EpochConverter". Below this, the heading "Epoch & Unix Timestamp Conversion Tools" is displayed in a dark red font. A section titled "The current Unix epoch time is" shows the value "1531955905" in a light blue box. Further down, a heading "Convert epoch to human readable date and vice versa" is shown. Below this heading, there is a text input field containing "1531954273", a button labeled "Timestamp to Human date", and a link "[batch convert timestamps to human]". The results of the conversion are shown below: "GMT: Wednesday, 18 July 2018 22:51:13" and "Your time zone: Thursday, 19 July 2018 08:51:13 GMT+10:00". A dashed line is at the bottom of the results section.

→ ↻ 🏠 Secure | https://www.epochconverter.com

 EpochConverter

## Epoch & Unix Timestamp Conversion Tools

The current Unix epoch time is **1531955905**

### Convert epoch to human readable date and vice versa

Timestamp to Human date [\[batch convert timestamps to human\]](#)

**GMT:** Wednesday, 18 July 2018 22:51:13  
**Your time zone:** Thursday, 19 July 2018 08:51:13 GMT+10:00

-----

# Some cookie data examples

- `_ga` cookie
  - GA1.3.0164af9713b500100f85468f61190004e00a500d00bd0
  - SHA Hash
- GUID
  - Globally Unique ID (see Week 6 - Registry)
  - 3b83fca5-3223-342c-459f-64e0fcf78633
  - 5 parts
- % Encoded
  - To treat control characters as plain text
  - %5B%5B%27SEM-GGL-SRC-FY16Q3-5463%27%2C%271531954291555%27%5D%5D
  - [['SEM-GGL-SRC-FY16Q3-5463','1531954291555']]

# Goodhart's Law

- People try and corrupt the system.
- A salesman will distort the sales figures so the analytics will give her more attribution and thus more bonuses.
- Goodhart's Law
- when a measure becomes a target it ceases to be a measure.
- A famous example is the Cobra bounty



# The Cobra Effect

- In British India there were many deaths due to Cobra snake bites
- The government set a bounty on cobra heads
- Farmers started to breed many cobras
- The cobra bounty was cancelled
- The farmers released their cobras
- Many more people died from cobra bites

# Third Party Cookies

- Third Parties provide content
  - advertisements
  - [like me](#) on Facebook
- This is done to make money
  - The [Pay per click](#) (PPC) business model
- User details are sent to the third party
  - used for marketing
- This can be stopped by [InPrivate/Incognito](#) Filtering

# Blocking Cookies

- Browsers have add-ons that block certain cookies
- Firefox **Ghostery** for example
- Blocks Advertisements
- Stops data tracking to preserve privacy
- Provides a viewer to show add/tracking activity
- One million users

# Google Third party cookies

- Do not relate to the page visited
- Designed to encourage you to buy an unrelated product
- Include
  - adwords.com
  - doubleclick.net
  - googleadservice.com
  - gstatic.com
  - youtube.com

# Cookie storage

- Each browser maintains its store of cookies separately
- Cookies are saved in a compressed format for speed
- We can use the browser [cookie manager](#) to view its cookies
- We can use third party cookie viewers when the browser is not running
- Each user has a separate cookie store

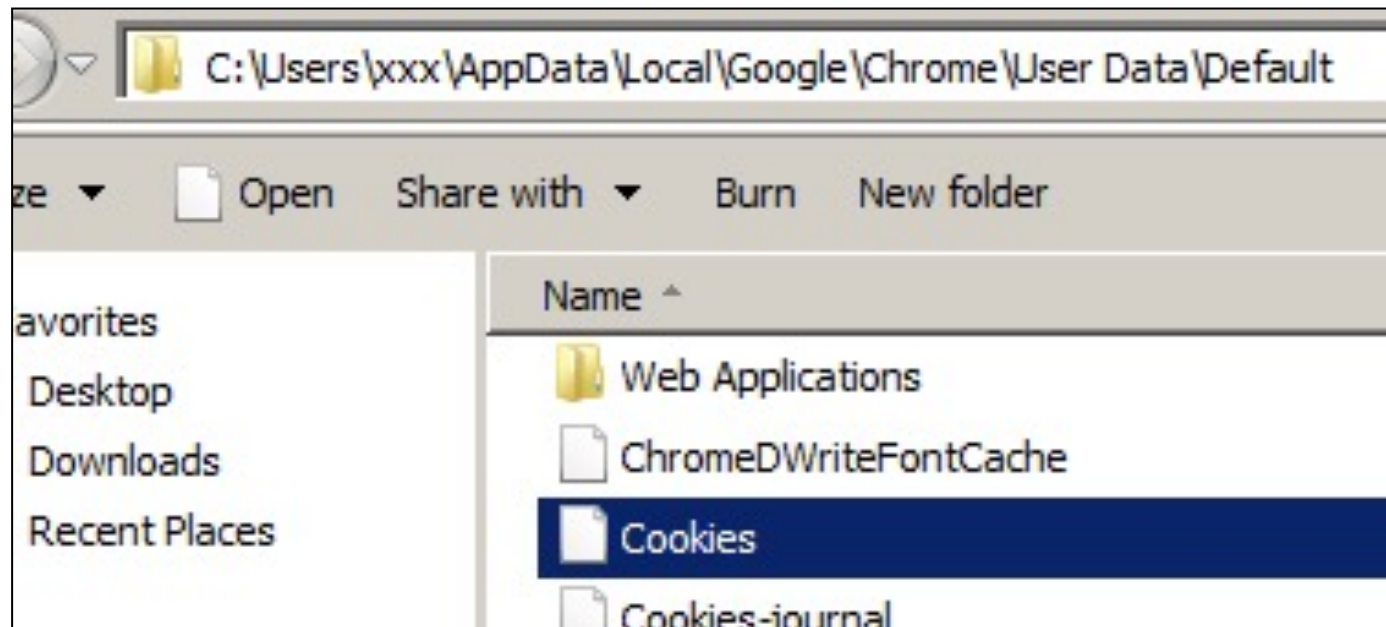


# Chrome Cookies

Site	Locally stored data
accounts.google.com	1 cookie
google.com.au	<div>3 cookies, Channel ID</div> <div>Channel ID NID PREF SNID</div> <div><div>Domain: google.com.au</div><div>Certificate Type: ecdsa_sign</div><div>Created: Friday, July 31, 2015</div><div>Remove</div></div>
www.google.com.au	Database storage, Service Workers

# Chrome Cookie File

- Keeps a list of cookies from all sites visited
- Can see the site name in plain text
- The user can delete these from chrome settings



# Viewing the Chrome cookie file - with find

- Copy the Cookies file to your Windows work folder
- Filter by your search term using find

```
C:\Users\graha>find "seek" C:\transfers\Cookies

----- C:\TRANSFERS\COOKIES
..æ°¼||.seek.com.au_gat_tealium_0/
.æý£ø4u.seek.com.au_gac_UA-63897908-1/
.æŸ0úÍ®.seek.com.au_ga/
.æŸ0nu.seek.com.aus_cc/
.æ°{Ñ.seek.com.aumain/
.æ°z.seek.com.au_gat_tealiumga/
.|.æý£ø/-seek.com.au_gid/
```



# Viewing the Chrome cookie file – with **grep**

- Copy the **Cookies file** to your **Linux** work folder
- Pull out the **ascii** using **strings**
- Filter by your search term using **grep**

```
group11~$ strings /mnt/c/transfers/Cookies | grep seek
.seek.com.au_gat_tealium_0/
4u.seek.com.au_gac-UA-63897908-1/
.seek.com.au_ga/
nu.seek.com.aus_cc/
.seek.com.aumain/
U.seek.com.aus_ev59/
z.seek.com.au_gat_tealiumga/
.seek.com.au_gid/
```

# Windows Subsystem for Linux (WSL)

- Runs Linux on Windows 10
- Runs Linux commands in a bash shell
- Can run bash, Python, MySQL, Apache, sshd

```
C:\Forensics_Graham>bash
root@PowerPC:/mnt/c/Forensics_Graham# cat /etc/issue
Ubuntu 14.04.5 LTS \n \l
```

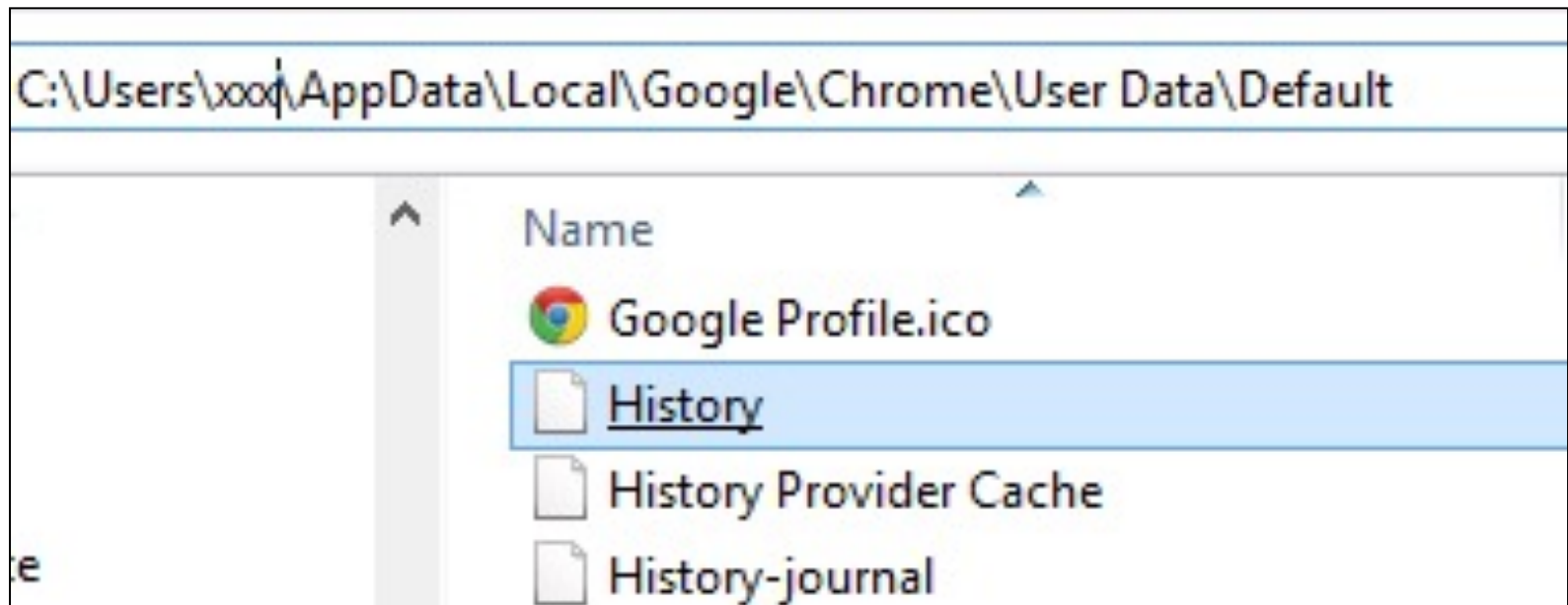
- Can install a Distro (Ubuntu for example)
- See Windows 10 Subsystem for Linux in Readings

# Objectives

- To use Web Browsers for Forensics
- To use web tracking for Forensics
- To locate and examine cookie files
- To locate and examine web browser history
- To locate and examine temporary internet files

# Web History Files

- The browser stores the history of visited pages
- This is usually large (MB) so cannot read directly.
- The browser has methods of deleting web history



# Viewing the Chrome web history file

- Copy the **history file** to your work folder
- Pull out the ascii using **strings**
- Filter by your search term using **grep**

```
C:\Forensics>strings history | grep hostworks
https://www.google.com.au/search?q=abn&oq=abn&aqs=chrome.
filetype:pdf+hostworks
filetype:pdf hostworks
filetype:pdf hostworks
site:seek.com.au hostworks
```

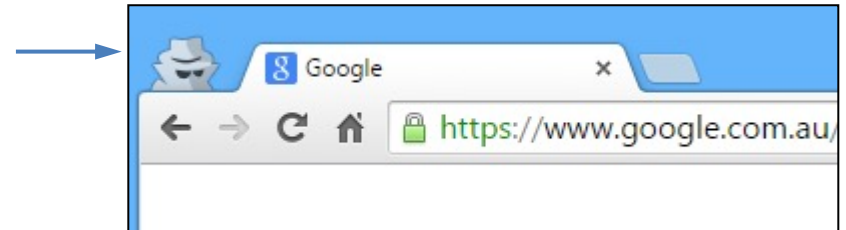
# Chrome top sites

- Records the most visited sites
- Can indicate a suspect's interest

```
C:\Forensics>strings.exe "Top Sites" | grep seek  
http://seek.com.au/t&  
http://seek.com.au/  
http://www.seek.com.au/ http://seek.com.au/?
```

# Hiding Web History

- Unfortunately for forensics, there are easy ways for users to minimise the evidence
- In chrome this is called **incognito** browsing



Going incognito doesn't hide your browsing from your employer, your internet service provider, or the websites you visit.

# Recovering hidden Web History

- A good forensic investigation can recover hidden evidence
- One likely place is volatile memory
  - process history
  - system history
- Another is disk
  - temporary files (including cached files)
  - swap files
- Yet another is the local dns server cache
- We will visit these later



# Objectives

- To use Web Browsers for Forensics
- To use web tracking for Forensics
- To locate and examine cookie files
- To locate and examine web browser history
- To locate and examine temporary internet files

# Temporary Internet Files

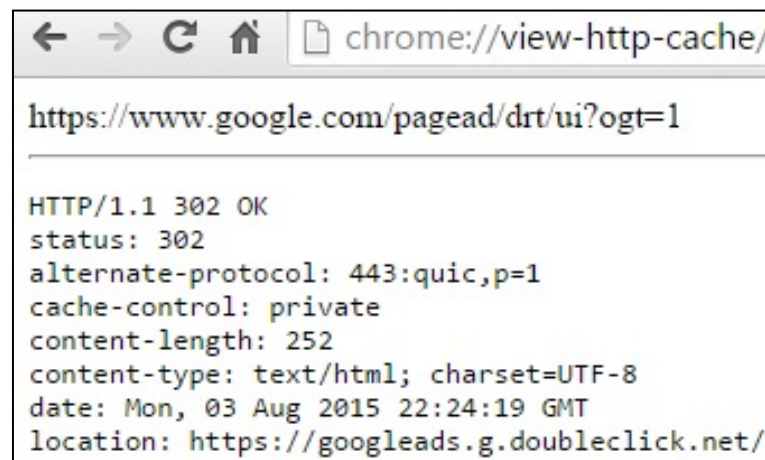
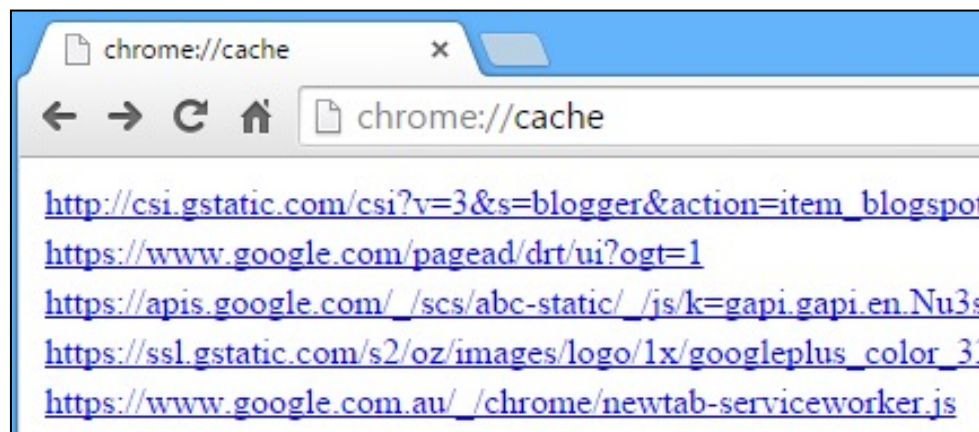
- http allows Web Browsers to **cache** recently visited pages
- When a viewer revisits a webpage, http checks the date on the cached page and decides whether to show the cached copy or refresh the page from the server.
- Caching cuts down on web traffic and speeds the rendering of the webpage
- Cached pages are a mine of forensic information

# Layout Engines

- The temporary file location is chosen by the web page layout engine
- The Layout Engine for IE is called [Trident](#)
- The Layout Engine for Firefox is [Gecko](#)
- The Layout Engine for Google Chrome is [Blink](#)
  - Blink is a fork of the WebCore part of WebKit
- The Layout Engine for Edge was [EdgeHtml](#)
- Current Edge uses Chromium with the [BLink](#) engine
- [en.wikipedia.org/wiki/Temporary\\_Internet\\_Files](https://en.wikipedia.org/wiki/Temporary_Internet_Files)
- [en.wikipedia.org/wiki/Trident\\_\(layout\\_engine\)](https://en.wikipedia.org/wiki/Trident_(layout_engine))

# Chrome Cache files

- These use data files, so no luck with strings
- Can use the chrome decoder



- Function removed since Chrome 66
- Can also use a Chrome Browser Forensics program

# Next week is Wireshark week

- You need to download and install Wireshark on your laptop for the Week 4 Lab.
- Check it works as expected.
- If you need a refresher, please do the Wireshark warmup Lab in Readings **before** the Week 4 Lab.

# Fin

