MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Computer Security Principles

Lecture 09: Privacy and Cookies

www.mtu.ie

Overview

1. Referencing “Googling Security” by Greg Conti
2. Google’s Capabilities
3. Information Flow and Leakage
4. Footprints, Fingerprints & Connections
5. Searching the Web
6. Communications
7. Advertising & Embedded Content
8. Countermeasures

Google Overview

Global Organisation used by billions, 92.71% of all queries (10/2020)

Collect your information "to provide you with a better experience"

3.5 billion queries per day

1000 computers in 0.2 seconds per query, travels an average of 2000 km round trip to/from a data centre.

More than 1.8 billion Gmail users (43% of email service market share in 2020)

Accounts for 27% of all email opens

75% of all Gmail users access email on mobile device

Google Overview

15 Gbytes of free storage shared across Google Drive, Gmail and Google Photos

61% of 18-29 year-olds use Gmail.

YouTube has over 2 billion unique monthly users (10/2020)

YouTube has almost 1 hour of video for every person on earth!

How has Google done it?

- Top-tier intellectual talent
- World-class information-processing resources
- Years of interaction data

Google Services

Alerts

Calendar

Docs & Spreadsheets

Earth

Gmail

Wallet

Maps

Workspace

Web Search

YouTube

Images

Mobile

News

Blogger

etc. etc.

Your personal information for access to free services

Google Usage Risks

Sometimes people are of the opinion that each small piece of information that they disclose online is not valuable

What about "the sum of all of your activities"?

Most people are unaware of the fact that their personal information which they have disclosed over numerous services is being "aggregated, data mined and tied together across many sites and social groups."

Most people are unaware that their real-world identity can be discovered easily from a relatively small amount of online activity

Information Disclosure Scenarios

Examples:

- A research group discovers a new technology. They use Google's patent database to search for any related patents. Their searches reveal the core ideas behind their technology.
- A law enforcement agency uses an online search engine to check the background of potential leads and anonymous sources for most of their active cases.
- A company with a mobile workforce chooses to place its financial books in Google spreadsheet so sales reps can access the information from around the globe.

Don't underestimate the importance of your disclosures over time

Data Retention & Permanence

Our online activity and data is logged and retained, in some cases for an indeterminately long time

We need to be aware of what information is being stored about us and how it is being managed

Although Google do provide privacy policies, they are often vague and broad in nature

Data which is stored on your computer is not legally treated equal to data which is stored on 3rd Party servers

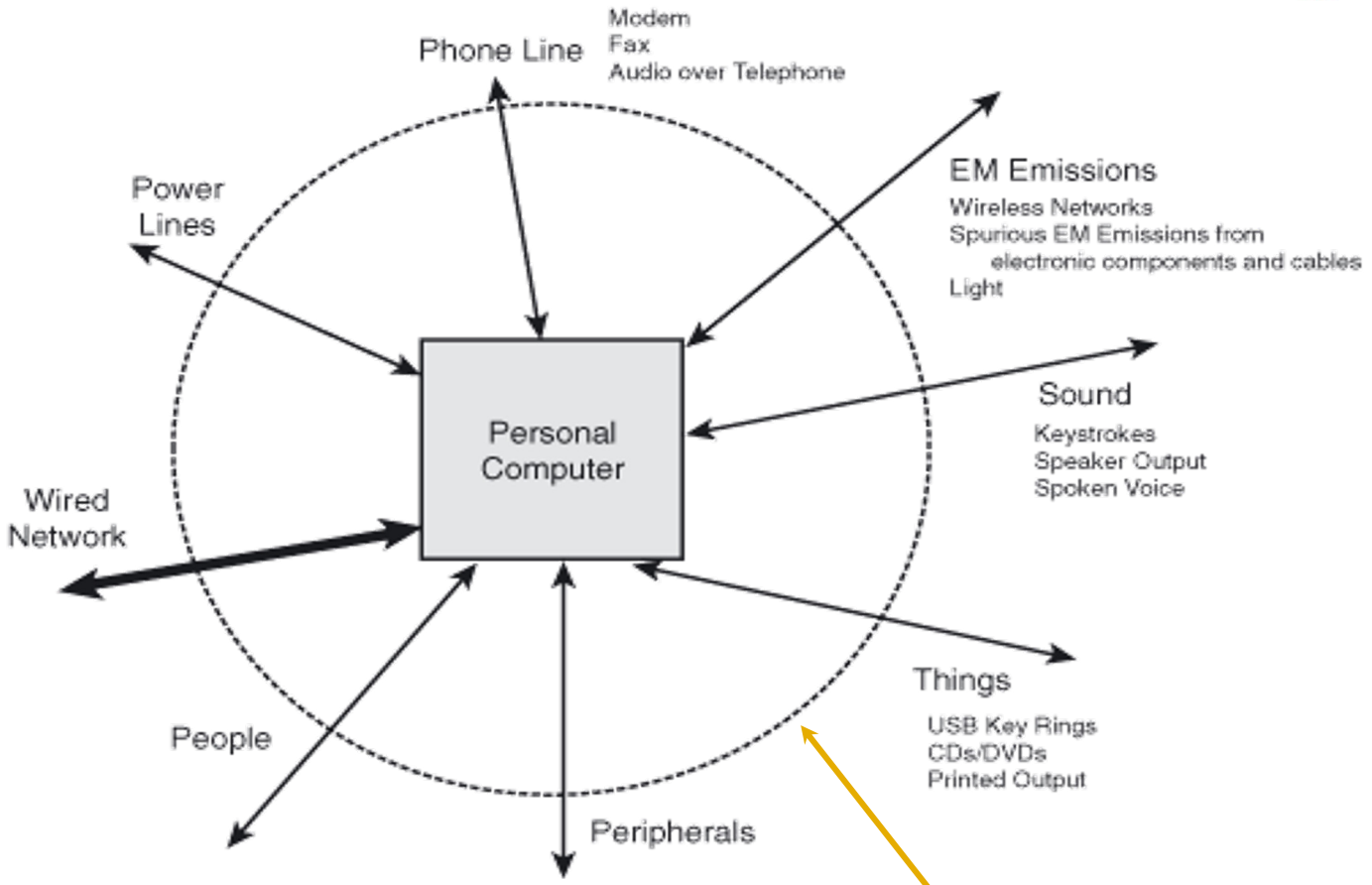
Information Flow

When you access the web, your information flows from your PC across the Internet until it finally reaches an online company

There are numerous points along the path where your information can be viewed or changed without your knowledge

- Individual Computer
- Network Communications
- Final Destination Website

Information Flow & Leakage



Imaginary line around your computer

Information Flow & Leakage

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Networks

- Wired, Wireless and Telephone networks
- Biggest threat vector for web-based information leakage
- Turning on your machine initiates a number of networking protocols and applications which reveal sensitive information
  - ARP, DNS, Email, AV software
- Wired networks require eavesdroppers to be physically present on the network. Attacker can use wiretaps and compromised machines to gain access
- Wireless machines broadcast their traffic so it is easier sometimes for an attacker to gain access. They do not need physical access

Information Flow & Leakage

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Peripherals

- Most computers today use USB, Bluetooth, serial, parallel, video, audio, FireWire and network ports
- They also communicate via buses (PCI/SCSI) which allow communications to external devices such as hard drives and scanners and also internal devices like graphic cards, network interface cards and modems
- These communications emit RF energy which is visible outside the device packaging or cabling itself
- Webcams and Microphones can be used to collect information
- Mobile phones can also be used since they contain microphones and cameras

Information Flow & Leakage

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

EM Radiation

- All electronic circuits emit some amount of electromagnetic energy
- These include radio bands, microwaves, infrared, gamma-rays
- Radio Waves are used in home wireless networking
- Some of these emissions can be detected from great distances
- Chips and Cabling emit RF energy
- LCD Monitors emit EM radiation which can reveal the contents of its screen
- Wireless keyboard and mice which broadcast every keystroke
- They can be detected with specialised interception equipment

Information Flow & Leakage

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Sound

- Advances in sensor technology has allowed for detection of sound from great distances
- Trojan horse programs or spyware on your machine can transmit sound from an external or internal microphone across the network to an attacker

Power Lines

- Another vector by which information can leave your machine
- It is possible to communicate information across electrical wiring

Information Flow & Leakage

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Humans and the Things They Carry

- Source of information leakage for a long time
- People can take analog and digital storage devices outside of the boundaries of your PC
- Media such as CDs, DVDs, USB keys and external hard drives were all invented for the purpose of moving data
- These devices are portable and can in some cases store lots of information

Data Communications On The Network

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Interactions with the servers of online companies require trusting a number of organisations

- Local ISPs
- Long-Haul Internet Backbone providers
- Domain Name Service
- The Online Company itself

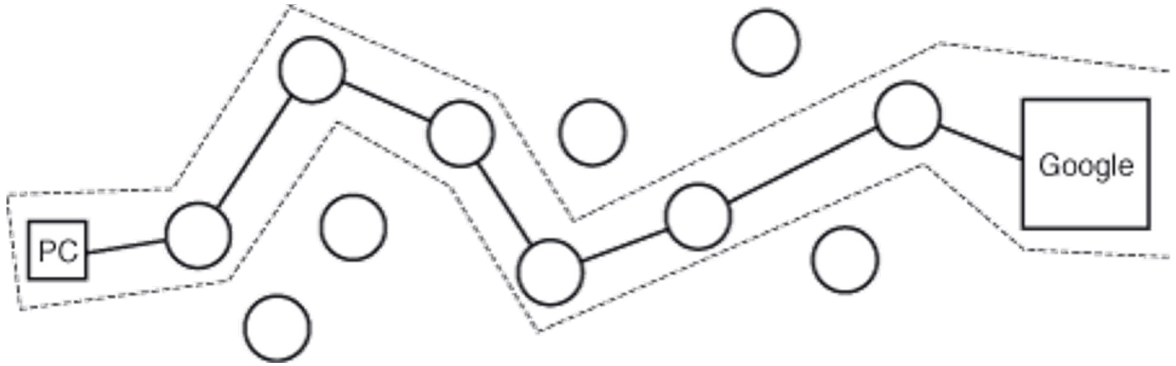
Even if you trust all of these organisations, an attacker can gain access to the communication link and threaten the confidentiality, integrity and availability of the communication stream

It only takes one component to break the security of the link

Information Flows and Leakage on the Internet

When you try to access a website, your messages are routed through links between network devices, usually routers, to the end website

The communication begins with a connection from your machine to the first device in the chain, then through 10-20+ other devices before it reaches its final destination



Information Flows and Leakage on the Internet

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

An attacker with access to any of these devices can potentially eavesdrop, alter, redirect or block your communications

When you are performing activities online, your requests and responses are sometimes sent in the clear

Although it is possible to encrypt your information when it travels over the network (SSL/TLS), it is unencrypted at either endpoint and therefore still vulnerable.

Aside: Use Windows **tracert** command to see how many hops it takes to reach Google.com. In Linux, the equivalent is called **traceroute**

Footprints

When accessing information online, we all leave a distinct trail in the logs of each server we interact with

These logs can be aggregated, mined and used to create user profiles to help organisations with targeted advertising

We will look at the following:

- How logs are created
- The information stored in a log
- Profiling users
- Uniqueness and Behavioural Targeting

Web Interaction and Data Retention

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

In order to access a web page, you type in a uniform resource locator (URL) in to the address bar of your browser

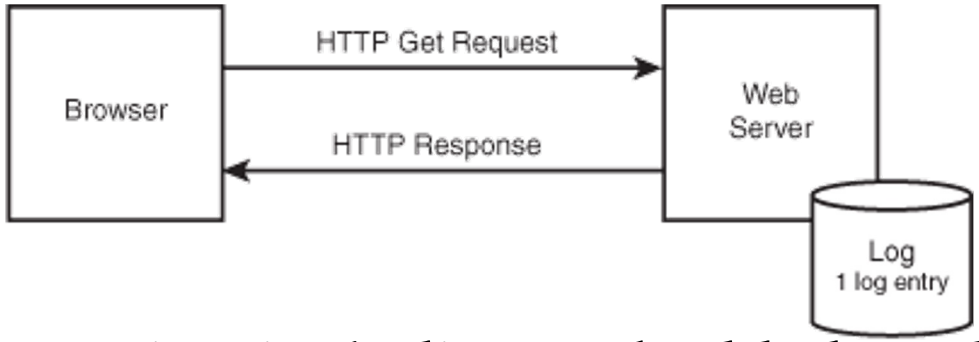
A URL is a global address to any publicly available document or other object on the web

Web Browsers and Web Servers communicate using HTTP (Hypertext Transfer Protocol) which defines the rules for requesting and receiving web content

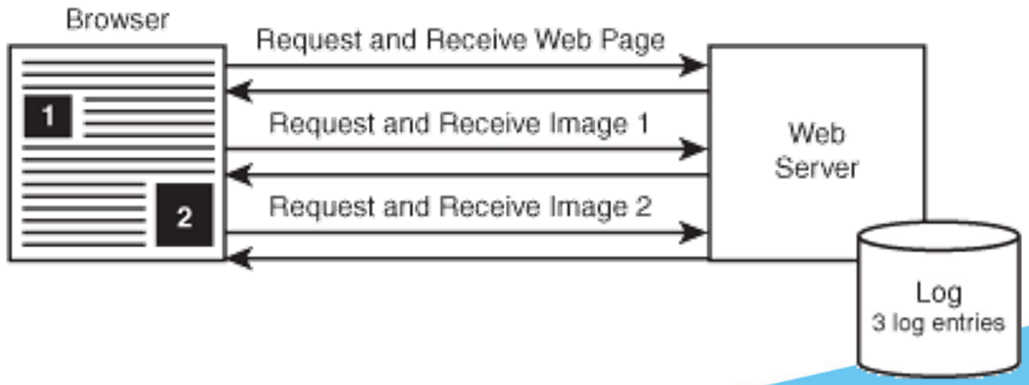
A Web Server is a software application which contains web pages, images or content and a private log

Web Interaction & Logs

Basic Interaction for a web page:

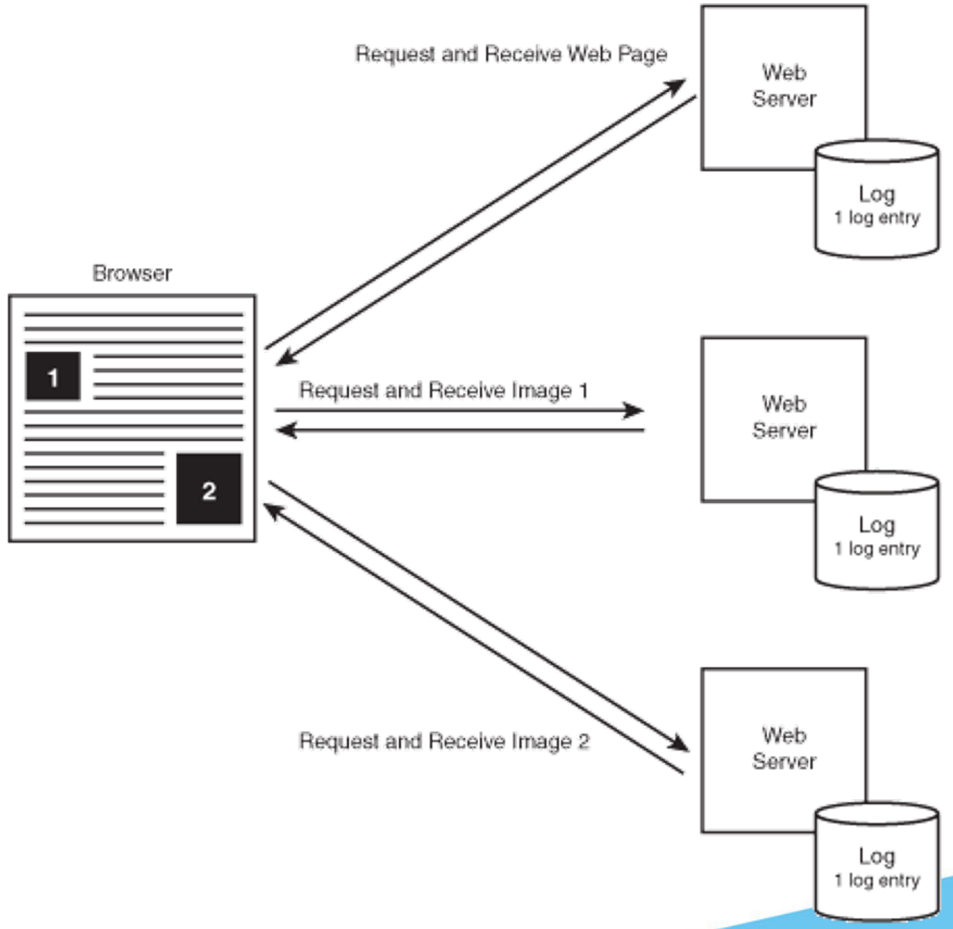


Normal Interaction including embedded media:



Web Interaction & Logs

Interaction with media from other 3rd Party sites



The Trail You Leave Behind

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Web Server Logs

- Each time a browser accesses a web page, image or object, an entry is placed in a web server log
- Example of a Log Entry:
- 86.X.X.X - - [27/Feb/2020:04:49:28 -0700] "GET / HTTP/1.1" 200 15384 "http://www.google.co.uk/search?hl=en&q=MALWARE &meta=" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/78.0.3904.108 Safari/537.36"

- <u>Breakdown:</u>
- 86.X.X.X –> IP Address of device making the request
- 27/Feb/2020:04:49:28 -0700 -> Date and time of request
- GET / HTTP/1.1" 200 15384 -> The HTTP request
- http://www.google.co.uk/search?hl=en&q=MALWARE &meta= -> The referring URL, the search term used and that the user speaks English.
- Version of the browser used, the client OS ➔ Chrome 78 on Windows 10

The Trail You Leave Behind

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

IP Addresses

- These were designed to identify computers on the Internet and are used to address packets that all Internet computers send and receive as they communicate
- Every time you communicate with a website, your browser and the destination web servers send and receive many packets which identify the address of the user and the destination
- With the growth in use, IP addresses became limited and it was decided that they should be shared
- Although this sharing provided some anonymity to users by breaking the direct link between user and IP Address, it is still possible to tie an IP to a particular machine

The Trail You Leave Behind

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

IP Addresses

- In order to tie an IP address to a machine you must:
  - Track the IP address back to the organisation which is allowed to use it.
  - Get access to the users details through the organisation.
- IP Addresses can be used to pinpoint a user's physical location

The Trail You Leave Behind

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Browser Header Fields

- Browsers disclose a lot of information as they visit web sites and not all of it is necessary for web transactions.
- See the **Browser Spy** website for extra details
- Headers sent from a Firefox Browser to the Google web server

GET / HTTP/1.1	The HTTP request.
Host: www.google.com	The host where the request is being made.
User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.8.1.11) Gecko/20071127 Firefox/2.0.0.11	Information about the application and **operating system** making the request.
Accept:text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;q=0.8, image/png,*/*;q=0.5	Media types the browser **will accept** in response.
Accept-Language: en-us,en;q=0.5	Preferred **language** of the requested content.

The Trail You Leave Behind

- Headers sent from a Firefox Browser to the Google web server

Accept-Encoding: gzip,deflate	Encoding formats the browser will accept
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7	Character sets the browser will accept
Keep-Alive: 300	The number of requests that the browser supports. Here, it is 300 requests per connection
Connection: keep-alive	A request to allow the client to make multiple requests over a given connection, to increase performance
Cookie:PREF=ID=0a0661ceb826a27d:TM=1199309328:LM=1199309328:S=ZoVTNgAwhwr1s3IY	Sending name/value pairs contained in a **cookie** Google issued during an earlier visit

The Trail You Leave Behind

Response from the Google Web Server to the Firefox Browser

HTTP/1.1 200 OK	The status code for the request. In this case, the code is 200, indicating success.
Cache-Control: private	The response is for a single user and shouldn't be placed in a shared cache.
Content-Type: text/html; charset=UTF-8	The **media type** of the response.
Content-Encoding: gzip	The type of encoding used on the object sent in response.
Server: gws	Information about the destination web server. In this case, gws probably stands for Google Web Server.
Content-Length: 2333	Size of the object sent in response to the request.
Date: Wed, 02 Jan 2020 21:35:03 GMT	Time of the response.

The Trail You Leave Behind

Cookies

- Web servers use cookies to mark web browsers with identifying information due to the fact that **HTTP is stateless** and cannot keep track of previous interactions.
- Two main types of cookies exist:
  - **Persistent cookies** which can exist for many years in a user's browser cache, repeatedly identifying the user to the issuing website across many visits.
  - **Session cookies** which exist only for the duration of a single online visit.
- Cookies are issued either by the web server of the site a browser is visiting (first-party cookies) or by a third party when a given page includes embedded content, such as an advertisement or video, provided by a third-party server.

The Trail You Leave Behind

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Cookies

- Cookies are small pieces of data that are passed to and stored by the browser by a given website. When needed, such as on a repeat visit, the cookie is passed back to the web server to uniquely identify the user
- Although it is possible to set a web browser to block cookies, doing so breaks many online applications
- Although, cookies can only be sent to the domain that issued them, there is a problem in that if you visit a website with 3rd party content then technically your browser visited that website and therefore they can issue you with a cookie

The Trail You Leave Behind

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Cookies

- Each browser on a given computer, such as Chrome, Safari, or Edge maintains a separate set of cookies
- In addition, most OS maintain a separate profile for each user of the browser, each with its own distinct set of cookies
- Cookies uniquely identify a browser/user account combination for a given individual
- The same cookie used from different IP addresses would allow an online company to map the networks of different domains that the user visits, perhaps including home, workplace, school, and travel destinations

The Trail You Leave Behind

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

HTTP Referer Data

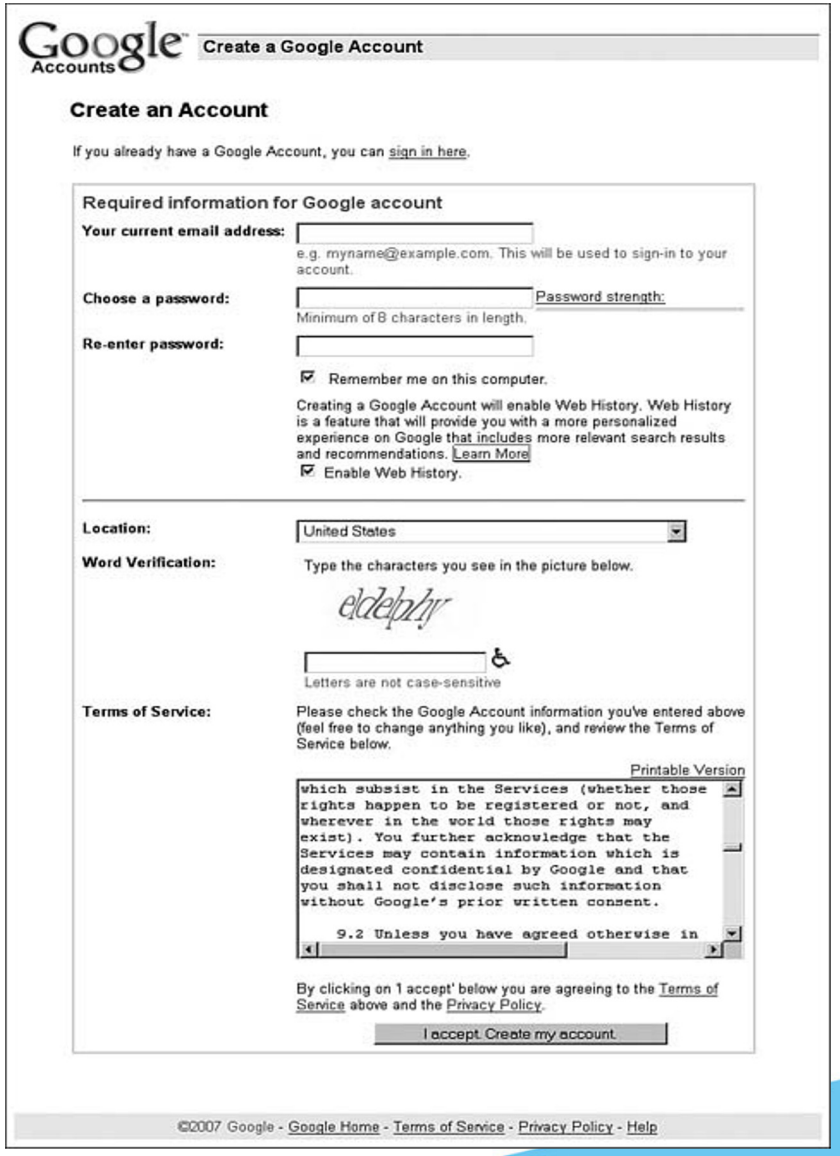
- When a user clicks a link in a web page, it makes a request to the appropriate web server for the desired object. It also passes a referer value, which indicates the page the user is browsing from
- They allow online companies to locate entry points to their web sites and link these sources to specific users via IP addresses and cookies
- More important, because Google embeds search queries in the URL that users click, the destination web site receives the search queries in addition to the user's IP address

Semantic Disclosures

Registered User Accounts

- Registering for a user account identifies you uniquely to an online company
- Many online services seek to find compelling reasons for you to go through the hassle of creating an account, such as by providing access to additional services or allowing you to personalise their interactions

Semantic Disclosures – Reg. User Accounts



Google Accounts — Create a Google Account

Create an Account

If you already have a Google Account, you can sign in here.

Required information for Google account

Your current email address: e.g. myname@example.com. This will be used to sign-in to your account.

Choose a password: Password strength: Minimum of 8 characters in length.

Re-enter password: Remember me on this computer.

Creating a Google Account will enable Web History. Web History is a feature that will provide you with a more personalized experience on Google that includes more relevant search results and recommendations. Learn More

Enable Web History.

Location: United States

Word Verification: Type the characters you see in the picture below.

eldelphjr

Letters are not case-sensitive

Terms of Service: Please check the Google Account information you've entered above (feel free to change anything you like), and review the Terms of Service below.

Printable Version

which subsist in the Services (whether those rights happen to be registered or not, and wherever in the world those rights may exist). You further acknowledge that the Services may contain information which is designated confidential by Google and that you shall not disclose such information without Google's prior written consent.

9.2 Unless you have agreed otherwise in

By clicking on 'I accept' below you are agreeing to the Terms of Service above and the Privacy Policy.

I accept. Create my account.

©2007 Google - Google Home - Terms of Service - Privacy Policy - Help

Semantic Disclosures – Reg. User Accounts

First, you are required to enter an **e-mail** address

Google verifies your address by sending a **follow-up confirmation** e-mail with a link to click before the account is activated.

This process ensures a **valid e-mail** address. (Remember, you've also disclosed your IP address and likely been tagged with a cookie)

Second, by leaving **Remember Me** on This Computer checked, I am automatically logged in as I use Google's services

Similarly, **Enable Web History** is selected by default. Web History is a Google Service that keeps track of a user's activity and helps personalise service

Semantic Disclosures - Web Site Navigation

Navigation between web sites (inter–web site navigation) can be tracked via third-party advertising networks (such as Google's **AdSense** and **DoubleClick** services), web analytics services (such as **Google Analytics**), and click-through monitoring that reveal browsing habits useful for targeted advertising and user profiling

Similarly, how a user navigates within a website (intra–web site navigation) is easily retrieved from **server logs** and provides insight into the information that user is seeking and the speed of the network connection, and perhaps helps identify that user because of unique behaviours

Likewise, Google Analytics and advertising networks might reveal what people are clicking on in a given site

Inter Website Navigation

Ways in which Inter Website Navigation can be tracked:

- HTTP referer data reveals the previously visited web site to the destination web site
- Third-party advertising networks reveal a user's visits to any site serving ads from the same advertiser
- Web-analytic services such as Google Analytics can track user's movements across any site in the analytics network
- Web bugs, small 1×1 transparent GIF images, placed by cooperative web masters, allow logging of user activities by third-parties
- Embedded third-party content, such as videos and maps, allows content providers to monitor a user's visits to any site including their content
- Click-through monitoring reveals which link a user clicked on a web page and, hence, that user's next destination

Inter Website Navigation

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Click-through Tracking

- JavaScript is used on web pages to provide interaction
- JavaScript can also be used on a webpage to allow the current web server to be able to detect and log which link a user clicked on a page
- This technique is used by Google which ties together search queries with the links that you click
- If JavaScript is turned on in your browser, the actual links Google returns point back to Google

Intra Website Navigation

Intra website navigation tracks a **user's browsing pattern** within a single domain

It has been used for **website analytics** for a number of years

Navigation within a particular website can reveal user's desired goals and could also be used to identify them

Most users exhibit unique or at least uncommon navigation behaviours that **uniquely identify them**

Uniqueness and Behavioural Targeting

So far, we have looked at the different ways in which a user discloses information to online companies either knowingly or not

- IP Addresses
- Cookies
- Browser Header Fields
- User Accounts
- Semantic Information

Large organisations then use this information to uniquely identify a user, profile their activities and possibly link all of this to discover their **real-world identity**

Behavioural Targeting

Also known as **User Profiling**, this uses web interaction data to categorise a user's interests and allow targeted advertising

It allows companies to increase the effectiveness of their **advertising campaigns** by capturing data generated by website and landing page visitors

When it is done without the knowledge of users, it may be considered a breach of browser security and illegal by many countries' privacy, data protection and consumer protection laws

Behavioural Targeting

The typical approach to this starts by using **web analytics** to break-down the range of all visitors into a number of discrete channels

Each channel is then analysed and a **virtual profile** is created to deal with each channel

These profiles can be based around **Personas** that gives the website operators a starting point in terms of deciding what content, navigation and layout to show to each of the different personas

Again, this behavioural data can be combined with known demographic data and a visitor's past purchase history in order to produce a greater degree of data points that can be used for targeting

Uniqueness

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Fingerprinting Scenario

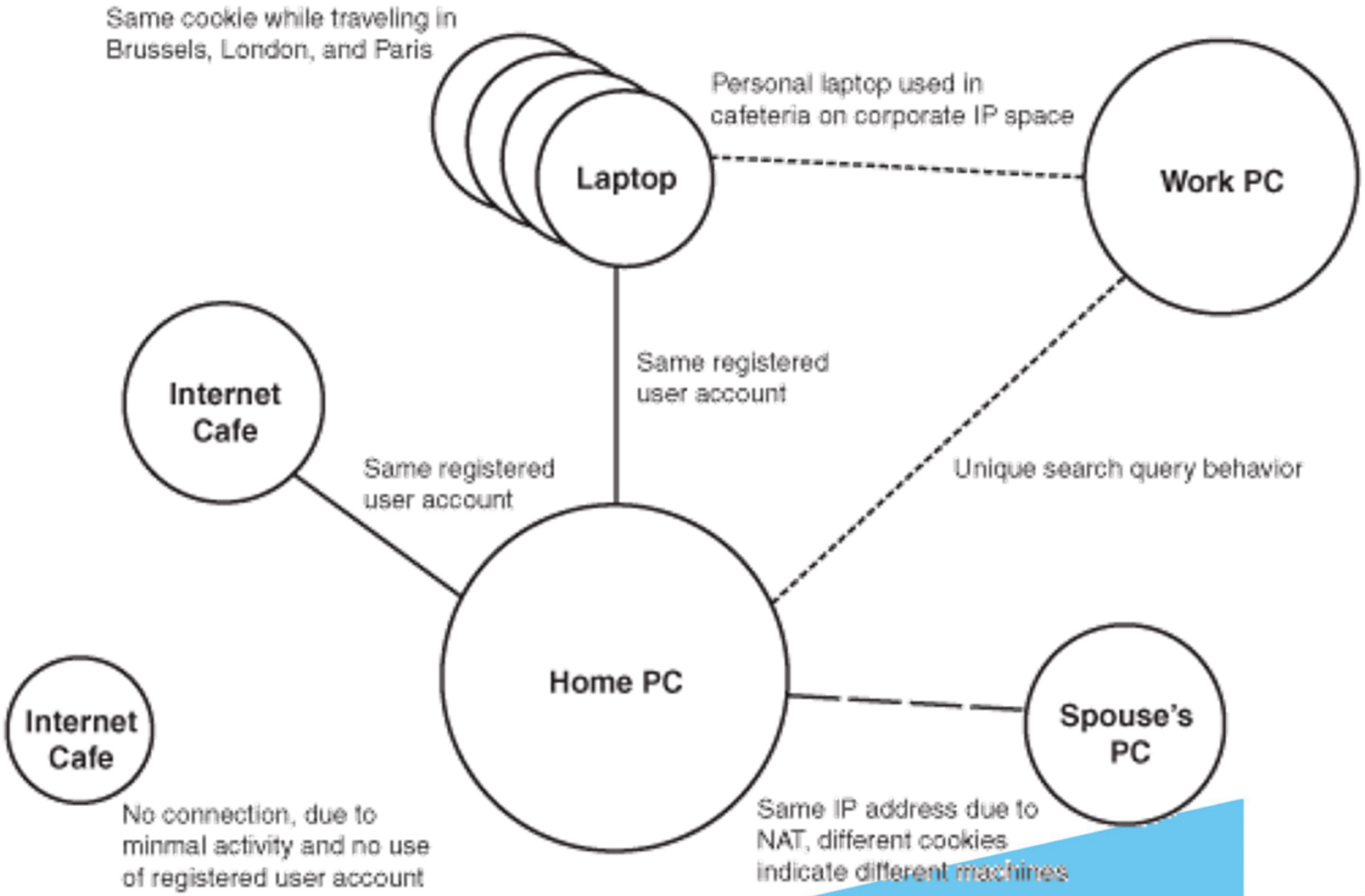
- A user employs many Google services on their home computer over time
- They then purchase a new laptop to use when travelling, this new laptop will have a different IP Address and different cookies to what is on their home computer
- Over time, the user will disclose significant information that will converge into unique forms of behaviour
- This will then identify the user of the laptop and the home computer as the same individual with an ever increasing degree of accuracy
- This uniqueness comes from the sum total of all online activity

Connections

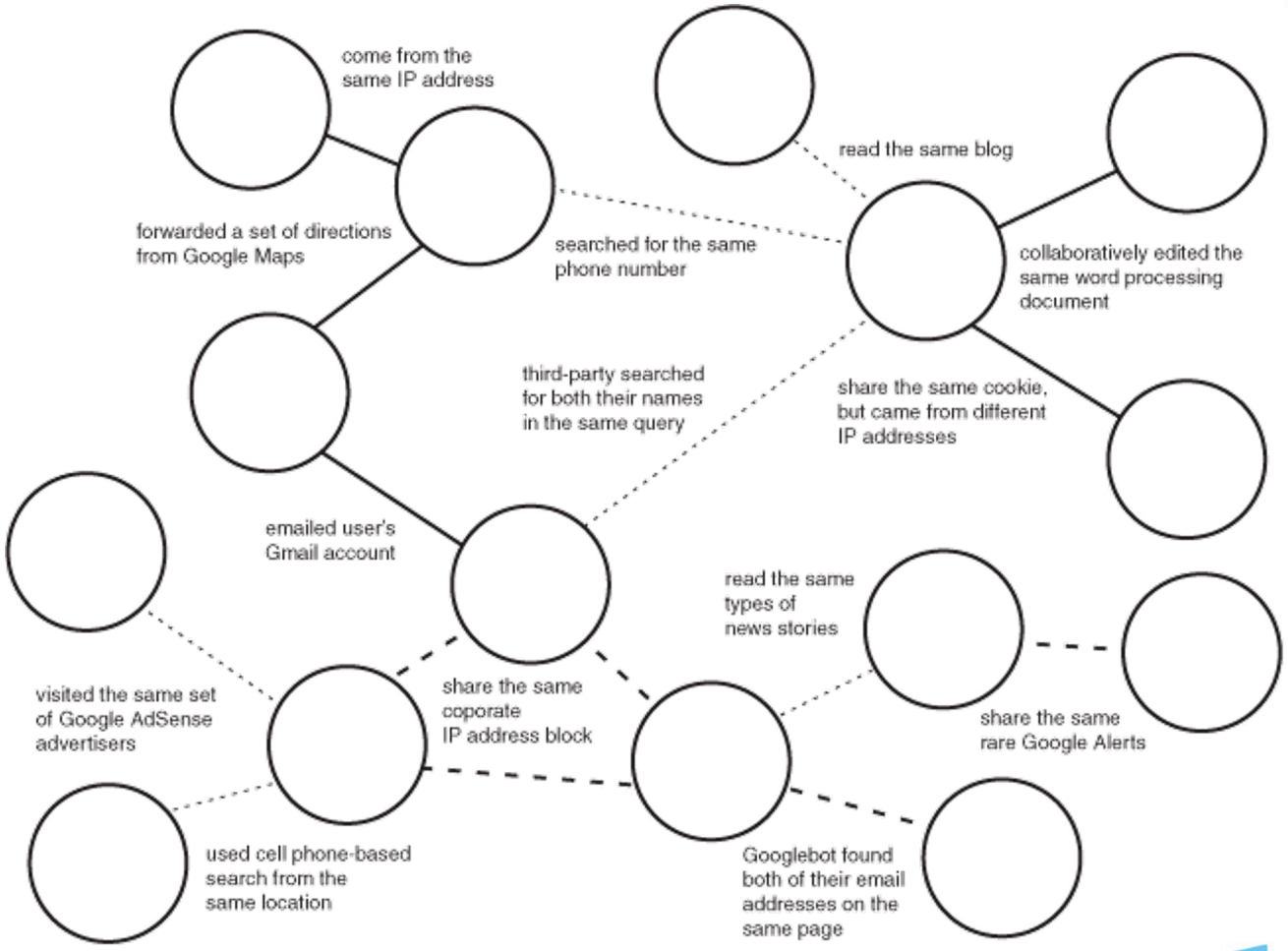
These traces that users leave behind allow organisations to link together individual users, and their profiles, across multiple computing platforms and networks as well as across communities of users who share the same attributes

Virtually any characteristic can be used to perform the linkage, including cookies, registered user accounts, browser header data, IP Addresses, physical location, computing platform, and semantic information, as well as information provided by other users and information collected by Googlebot from web pages

Connections – Linking the Same User



Connections – Linking Different Users



Searching the Web

The information-centric nature of the web demanded a way to locate web sites and relevant information

Enter the search engines but only one would rise to the top: **Google**

Today we can search for anything from anywhere

What's in a search?

We enter what we are looking for into the search box and Google does the rest

Perhaps one query will not reveal that much information about a user but what about **all of the queries** which we enter over time

Searching the Web

Information which we reveal in our searches

- Locations
- Religious Affiliations
- Medical Conditions
- Business Plans
- Hobbies and Interests
- Stage of Life
- Hopes and Dreams

Communications

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Email

- One of the oldest communication mechanisms on networked computers
- Approx. 206 billion emails sent every day in 2020 (business & consumer users)
- Gmail has over half a billion users
- Gmail is nominally free but is subsidised by advertising
- By analysing the e-mails and inserting context-specific advertising, Google has discovered a profitable business model

Communications

Gmail Privacy Policy extract

- "All e-mail services scan your e-mail. They do this routinely to provide such popular features as spam filtering, virus detection, search, spellchecking, forwarding, auto-responding, flagging urgent messages, converting incoming e-mail into cell phone text messages, automatic saving and sorting into folders, converting text URLs to clickable links, and reading messages to the blind."

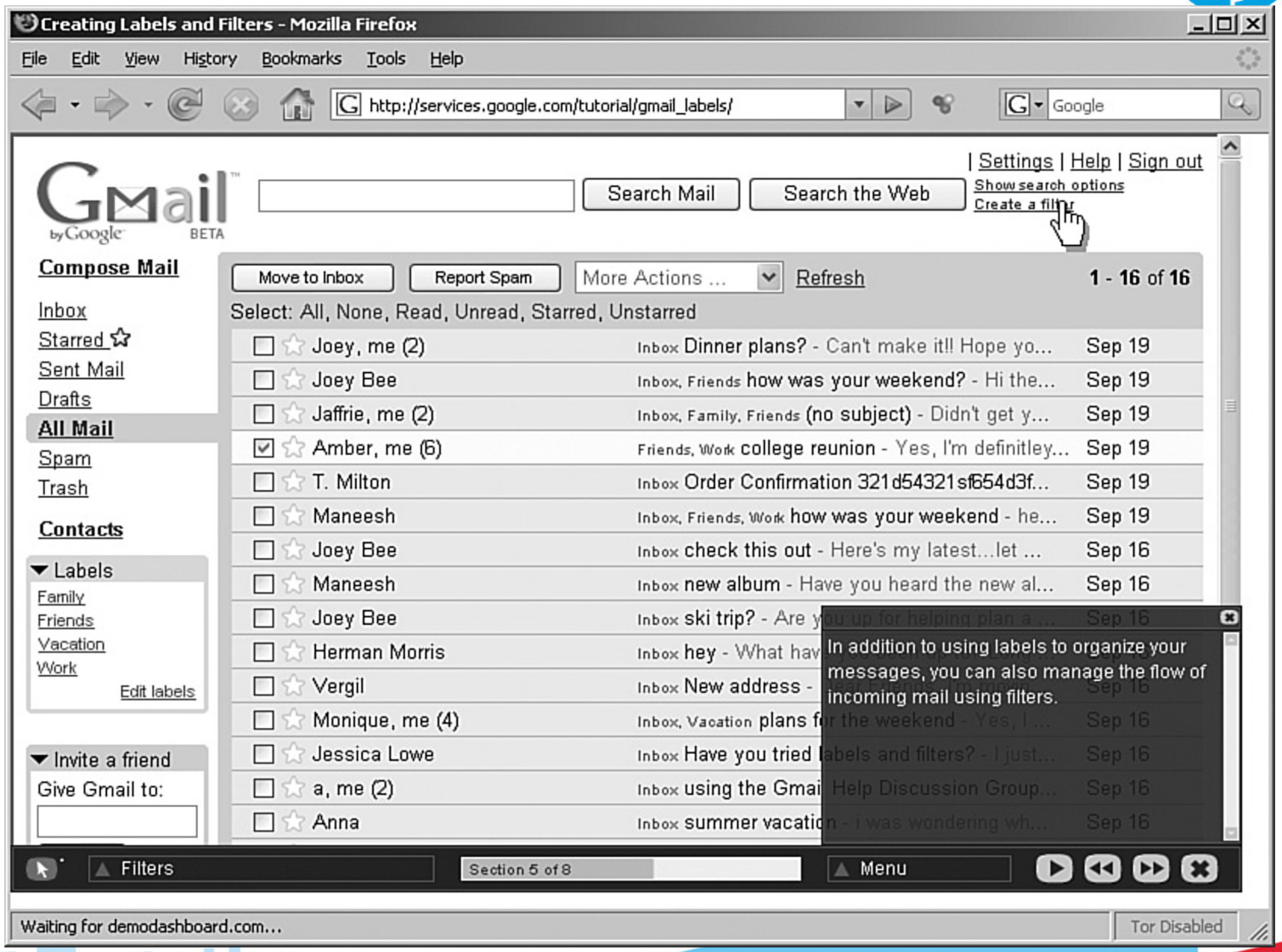
Google's Privacy Policy Analysis

Excerpt	Analysis
"You need a Google Account to access Gmail. Google asks for some personal information when you create a Google Account, including your alternate contact information and a password, which is used to protect your account from unauthorized access. A Google Account allows you to access many of our services that require registration."	A Gmail account uniquely identifies you as you use Gmail and other Google services. Alternate contact information provides a way to link activities with other non-Google accounts.
"Gmail **stores, processes, and maintains** your messages, contact lists, and other data related to your account in order to provide the service to you."	Google **stores copies** of your messages and contact lists on its servers and uses your IP address to assist in creating effective advertising.
"Google maintains and processes your Gmail account and its contents to provide the Gmail service to you and to improve our services. The Gmail service includes relevant advertising and related links based on the IP address, **content of messages**, and other information related to your use of Gmail."	
"Google's computers process the information in your messages for various purposes, including formatting and displaying the information to you, delivering advertisements and related links, preventing unsolicited bulk e-mail (spam), backing up your messages, and other purposes relating to offering you Gmail."	The e-mails **will be examined** by Google's machine processors.

Google's Privacy Policy Analysis

Excerpt	Analysis
"Residual copies of deleted messages and accounts may take up to **60 days to be deleted** from our active servers and may remain in our offline backup systems."	Deletion of e-mail from "offline" backup systems isn't guaranteed.
"We do not sell, rent, or otherwise share your personal information with any third parties except in the limited circumstances described in the Google Privacy Policy, such as when we believe we are required to do so by law."	Google can provide personal information to its subsidiaries, affiliated companies, and trusted businesses or persons for the purpose of processing personal information on their behalf. Information transfer can occur if Google becomes involved in a merger or acquisition. Information can also be shared as required to satisfy any applicable law, regulation, legal process, or enforceable government request.
"When you use Gmail, Google's **servers automatically record certain information** about your use of Gmail. Similar to other web services, Google records information such as account activity (including storage usage, number of log-ins), data displayed or clicked on (including UI elements, ads, links), and other log information (including browser type, IP-address, date and time of access, cookie ID, and referrer URL)."	Google can record log-in activity, IP addresses, cookie IDs, referer URLs, and user interface elements or advertisements the user clicks.

Gmail Labelling



Advertising and Embedded Content

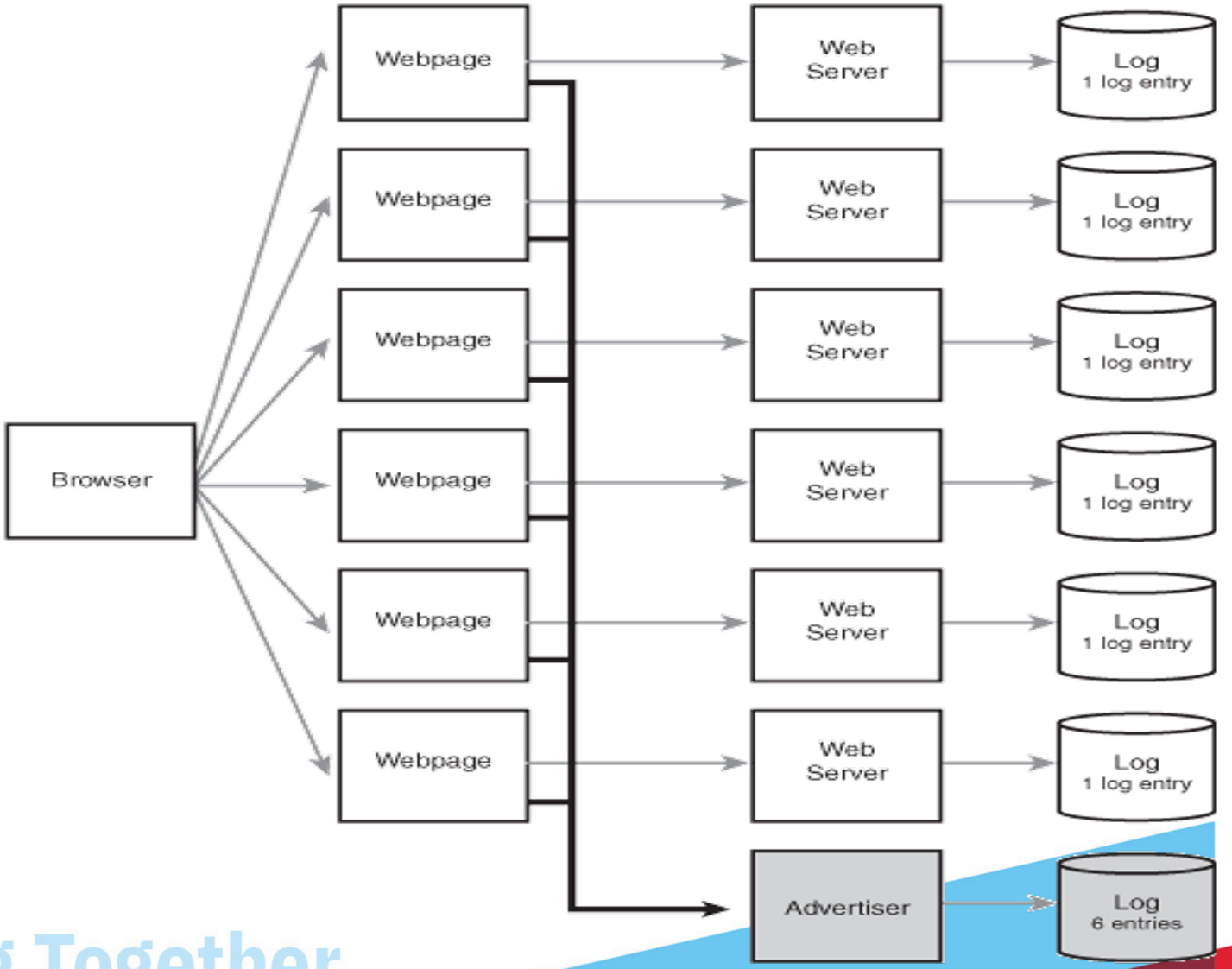
Publishing information is the backbone of web content

Bloggers and webmasters frequently rely on embedded content from companies such as Google to enhance the quality of their sites

Unfortunately, embedding third-party content is the equivalent of planting a web bug in web pages, alerting the source of the embedded content to a user's presence on a given site and facilitating logging, profiling and fingerprinting

The source of the embedded content can aggregate these single instances and track users as they browse

Cross Site Tracking



Cross Site Tracking

In the previous slide we see an example of Cross Site Tracking

- A user visits 6 distinct websites each hosting content from a single advertiser
- In turn, the user's visits create one set of log entries on each of the six legitimate servers
- However, because each visit contained an advertisement from a single advertiser, the advertiser is able to log all six visits

Advertising

Advertising is the fuel behind virtually all free tools online

It is also the means for **tracking** your web activity across the internet

Web Advertisements are big business with Google earning almost **$134.3 billion** in 2019 (Google's total earnings in 2019 = $160 billion)

Some of the largest services are offered by Google.

- AdSense
- AdWords
- DoubleClick

Advertising

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

AdSense

- Service which allows webmasters to earn advertising revenue by hosting AdSense ads
- The revenue from hosting these ads can range from a few hundred dollars a month to $50,000 a year
- It is an extremely popular service
- The ads are context-sensitive textual ads served by Google based on the hosting site's content
- Unfortunately, just visiting these websites allows Google to collect a user's IP address and log their visit

Advertising

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

AdWords

- Every time a user conducts a search on Google, the company makes money
- AdWords allow advertisers to bid on search terms that are displayed as part of the user's search results. The better the placement the higher the cost
- AdWords poses both information-disclosure risk and other security risks
- Attackers have used AdWords to serve malicious websites to users
- Google AdWords partners are a significant information disclosure risk as they can send searches from their sites to Google

Advertising

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

DoubleClick

- Online advertising service taken over by Google in 2008
- Excels in display advertising, such as flashy banner ads and video advertisements, which reach 80-85% of the web population
- The end result is a broad net that permits Google to track a user's searches and website visits, with the potential to impact privacy interests

Advertising Risks

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Malicious Ad Serving

- Advertising networks can serve as a malware attack vector
- Attackers can create misleading advertisements as a means to draw traffic to a malware serving or other malicious website
- The users' trust of the advertisement company and the hosting web site increases their trust of the advertisements, leaving web surfers more vulnerable to such an attack
- Virus writers have used the Google AdWords service to serve text ads that appeared to link to legitimate destination sites, but silently infected vulnerable web surfers by routing users through an intermediate, malicious site

Advertising Risks

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Malicious Interfaces

- There are times when advertising interface designers trick users into viewing their ads
- Examples:
  - Fake hyperlinks that pop up advertisements
  - Giant advertisements that cover the text of articles
  - Banner ads with fake buttons that appear to be part of the interface
  - Distracting advertising videos that begin playing the moment a page is viewed

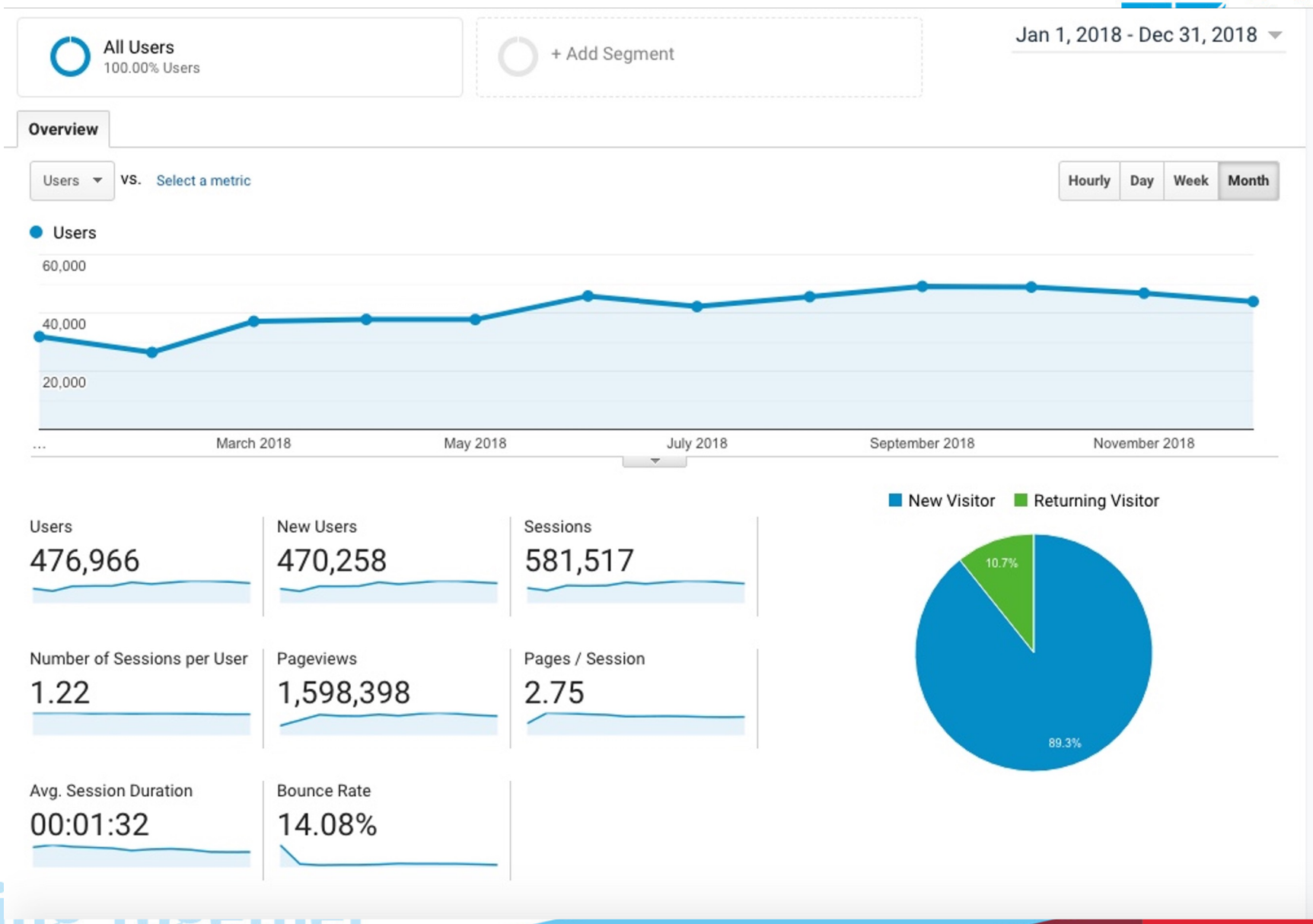
Advertising Risks

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Google Analytics

- Free tool for webmasters that provides a powerful and intuitive interface for analysing web log data
- It provides statistical and graphical analyses of web visitor activity based on web server log data and on data gained via cookies placed on users' computers, web bugs and JavaScript code
- A large number of websites today use Google Analytics
- If you combine this with sites participating in AdSense and AdWords then it can be intrusive in to a user's activity

Advertising Risks – Google Analytics



Teicneolaíochta na Mumhan
Technological University

Counter Measures

No clear solution exists to protecting your identity online

It is necessary to employ a number of different techniques but you must decide which ones suit you as locking down your information completely can hinder your web browsing experience

Techniques

- Patching Users
- Technical Protection

Counter Measures – Patching Users

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Raised Awareness

- Realise that a problem exists

Know what you are divulging

- Some things you need to consider
  - Consider each new tool carefully and understand the implications of the personal information that is disclosed through its use
  - Consider how much of what you are sharing you would ever share with your parents, partner, or co-workers
  - Read what privacy policies say directly and what they say between the lines
  - Disclose the minimum information to accomplish a task
  - Think in terms of years' or decades' worth of disclosures, not single instances or even days

Counter Measures – Patching Users

Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Usable Security

- Protection against web-based information disclosure demands usable security
- Usable security seeks to help people remain secure by creating systems that are designed to be effective, efficient, understandable, and easy to use
- Such systems don't waste users' time and attention; instead, they seek to disturb users only when they need to make a decision
- A trade-off exists between usability and privacy using today's technology—each current approach comes at a cost

Counter Measures – Technical Protection

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Controlling Cookies

- Reduce or eliminate cookies that accumulate on your computer
- Modern browsers provide intelligent cookie management
- Check out your browsers cookie management capability and what it sees as best practise

Counter Measures – Technical Protection

Diffusing or Eliminating your Disclosures

- Content Filtering
  - Employ your own proxy (Privoxy)
  - A proxy acts as an intermediary between your web browser and the destination website
  - Proxies have the capability to change both your outgoing and incoming traffic. For example, Privoxy can block web bugs, banner ads, cookies, and referer values
  - Because Privoxy is under your direct control, you can use it to tweak your outgoing and incoming traffic to your heart's content. It enables an extreme level of customisation, such as blocking advertisements and third-party content
  - A local installation of Privoxy can't provide full anonymous surfing; you need to add an anonymising proxy or network

Counter Measures – Technical Protection

Diffusing or Eliminating your Disclosures

- Self Monitoring
  - When you surf the web, your browser is capable of giving you only limited information on your past activities via its history function
  - What if you had the capability to track the information you disclosed during weeks, months, or years
  - Check the Google Web History service
  - Although not complete, it does allow you to view your online activity over time for some of its services

Counter Measures – Technical Protection

Diffusing or Eliminating your Disclosures

- Search Term Chaffing
  - Chaffing seeks to include false or misleading activity in this stream, to conceal your real activity and intent
  - An example is TrackMeNot by New York University's Daniel Howe and Helen Nissenbaum. TrackMeNot is a Firefox (& Chrome) browser plug-in that attempts to address many of the technical issues surrounding search query chaffing. Most notably, it seeks to mimic human search queries by intelligently timing spurious queries and dynamically evolving queries to mirror real searches, instead of using a static word list

Counter Measures – Technical Protection

Protect your network address

- Along with cookies, your **network IP address** identifies you to online companies every time you visit
- Your address might also be included in the header information of each email you send. Blocks of IP addresses are allocated to ISPs, companies, organizations, and educational institutions
- Perhaps the best hope for anonymous web browsing is the **anonymising proxy (TOR)**. Anonymising proxies act as intermediaries between you and the destination website. They make requests on your behalf, filter some identifying information (such as cookies, browser header fields, etc.), replace your IP address with theirs, and pass responses back to you

MTU
Ollscoil Teicneolaíochta na Mumhan
Munster Technological University

Thank You!