

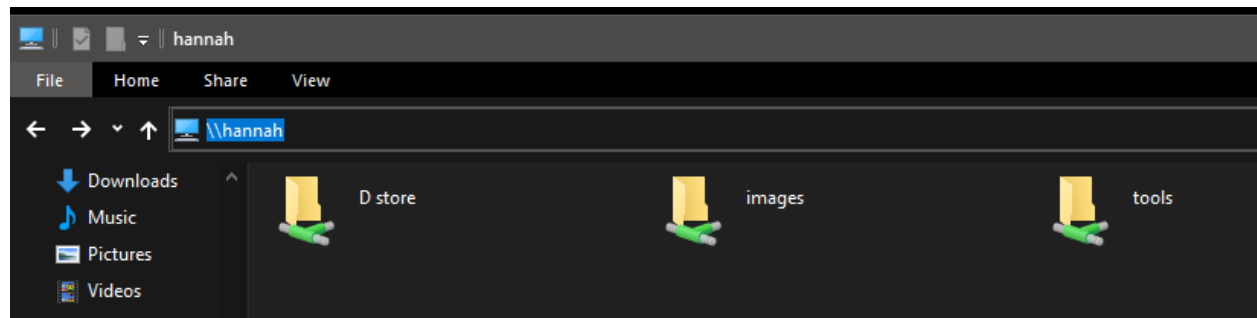


Footprinting, information gathering and OSINT Exercises.

Note 1: - When copying and pasting from this document, be careful. Sometimes Word will alter characters. For example two dashes are often altered to a single character.

Note 2: - The exercises are designed to give you experience of using a range of tools. You should evaluate the tools as you go along.

Note 3: - The location of the tools folder in the University labs is `\\hannah\tools\pentest`



Note that Information contained in this document is for educational purposes.

Contents

1	Introduction	1
1.1	Introduction To Footprinting	1
1.1.1	The Basics of Footprinting.....	1
1.1.2	What is OSINT?.....	1
1.1.3	Footprinting Methodology.....	1
2	Footprinting Exercises.....	3
2.1	Examine Authoritative Bodies.....	3
2.2	Web Based Tools.....	4
2.2.1	On-line Whois.....	4
2.2.2	Finding Sub-Domains	5
2.2.3	Finding Web-Servers	6
2.2.4	Finding Mail Servers.....	6
2.2.5	Google Hacking	7
2.2.6	Shodan Searching.....	10
2.2.7	Using 192.com.....	13
2.2.8	Accessing Archived Information	13
2.2.9	Multi-information Sites.....	13
2.2.10	Google Maps	13
2.1	Freedom of Information Act	14
2.2	Local Tools.....	15
2.2.1	OWASP Mantra	15
2.2.1	Maltego	16
3	Further Resources	21
3.1	Copycat Domain Names.....	21
3.2	Immersive Labs	21
3.3	Pimeyes	21
3.4	OSINT – VM (This is a major project).	22
3.5	Further Reading	22

1 INTRODUCTION

1.1 INTRODUCTION TO FOOTPRINTING

Footprinting is the process of accumulating data regarding an organization and their specific network environment, usually for the purpose of finding ways to intrude into the environment. Footprinting can reveal system vulnerabilities and improve the ease with which they can be exploited.

1.1.1 The Basics of Footprinting

Specific information about the organisation is gathered using non-intrusive methods. For example, the organizations own web page may provide a personnel directory, which may prove useful if the hacker wants to use social engineering to reach their objective. Information obtained may include: -

- Company locations and branches
- Other companies with which the target company partners or deals
- News - such as mergers or acquisitions
- Links to other company-related sites
- Operating systems and hardware being used
- IP addresses, E-mail addresses
- Newsgroups posts
- Phone numbers
- Policies and procedures (which may help identify the types of security mechanisms in place).

Footprinting may require a manual research, such as studying the company's web page and employee's social network accounts for useful information. However, there are many tools that can automate the information gathering process.

1.1.2 What is OSINT?

This stands for Open Source Intelligence which is scraping information from public sources. There is such a wealth of legally collectible OSINT available now thanks to social media and the prevalence of online activities that this may be all that is required to give an attacker everything they need to successfully profile an organization or individual.

1.1.3 Footprinting Methodology

Most penetration testing companies will use a footprinting methodology. For example: -

1. Examine authoritative bodies (IANA, ICANN, RIPE etc.).
2. Examine web site(s).
3. Use web based tools
4. Use local tools (OS command or download)
5. Conduct an Internet Search.

6. Web site copying
7. Social Engineering
8. Dumpster diving
9. Other techniques

Separating the acts of “Footprinting” and “Scanning” is difficult. This document is mostly related to “passive” searching (i.e. we try not to send any information to the target servers so that they can’t trace us). The scanning exercises in the future will be related to “active” activities. Consider these exercises as the first phase of fact finding about a company network.

Note (1): - Professional penetration testers will perhaps spend 3 or 4 hours footprinting whereas hackers can spend months on the process.

Note (2): - Some of the applications below can take a long time to run and master (e.g. Maltego so the following exercises are mainly designed to introduce you to the tool.

2 FOOTPRINTING EXERCISES

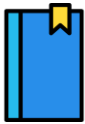
2.1 EXAMINE AUTHORITATIVE BODIES

Internet registry sites are the main Internet bodies and can give excellent information to a hacker and a security specialist alike.

The five regional registries and the areas they cover are: -

- RIPE - Europe, the Middle East, Central Asia
- AfriNIC - Africa, portions of the Indian Ocean
- APNIC - Portions of Asia, portions of Oceania
- ARIN - Canada, many Caribbean and North Atlantic islands, and the United States
- LACNIC - Latin America, portions of the Caribbean

Taking RIPE as an example, the RIPE Network Management Database (often called the "RIPE Database") contains information about IP address space allocations and assignments, routing policies, reverse delegations, contacts in the RIPE service region. In essence it is a database of the IP address owners in the RIPE region.



Take note!
These databases are normally termed the WHOIS databases.



Use the following WHOIS database to discover information regarding St Andrews University (st-andrews.ac.uk).
<https://apps.db.ripe.net/search/full-text.html>



Question.
Examine the information you discovered. You should refer to the lecture material to illustrate what is valuable.

2.2 WEB BASED TOOLS

There are numerous web based tools that make footprinting easier to perform.

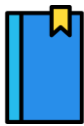
2.2.1 On-line Whois



These sites make searching the whois database easier.

- <http://whois.domaintools.com/>

Try looking for **abertay.ac.uk**

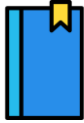


Take note! The related **TLDs** (top level domains) for the above site as shown in Figure 1. Find the link down the right hand side of the webpage.

View Screenshot History	
Available TLDs	
General TLDs	Country TLDs
The following domains are available through our preferred partners. Select domains below for more information. (3rd party site)	
■ Taken domain. ■ Available domain. ■ Deleted previously owned domain.	
AbertaY.com	View Whois
AbertaY.net	Buy Domain
AbertaY.org	Buy Domain
AbertaY.info	Buy Domain
AbertaY.biz	Buy Domain
AbertaY.us	Buy Domain

Figure 1: Example Output of Whois site.

What could be the use of this to a pen tester?



Take note! There are other tools for whois.domaintools.com as shown in Figure 2 (the web page layout may be different from that shown).

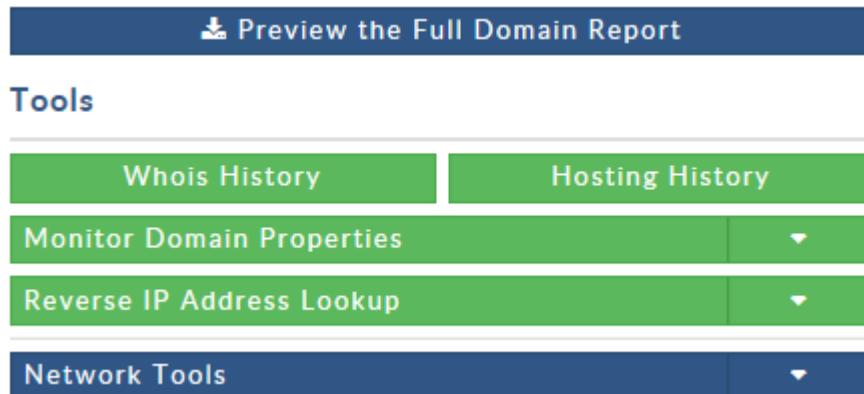


Figure 2: Other footprinting tools found on whois site.



Examine the sample **Domain report example** on **MyLearningSpace**.

2.2.2 Finding Sub-Domains

It can be useful to find out about all the subdomains of a company.




Use the following website to get subdomain information about **St-andrews.ac.uk**.

<https://dnsdumpster.com/>

This site gives really useful information. E.g. Look at the example files on **MyLearningSpace**. (e.g. **st-andrews.ac.uk.png** and **st-andrews.ac.uk.xlsx**).

2.2.3 Finding Web-Servers

Finding a company website via Google (or other search engine is simple). The Netcraft website gives more useful information.



Try to identify which web servers are running for St Andrews University using <https://searchdns.netcraft.com/> as shown in Figure 3.

Search: [search tips](#)


site contains ▼

st-andrews

lookup!

example: site contains .netcraft.com

Figure 3: Netcraft query



Question.

Examine the site report link and answer the following questions.

What OS version is running? (look for OS).

What Web Server version is running? (Apache? They may try to hide it though).


2.2.4 Finding Mail Servers

To find a “receiving” E-Mail server (SMTP or ESMTP) is relatively simple using DNS queries (MX) record. Mxtoolbox allows a simple passive method of finding the IP addresses of a company’s mail server (and some more information).



Use the following website to find the **st-andrews.ac.uk** mail servers.

<https://mxtoolbox.com/index.aspx>



Also do a diagnostic test. This passively gives “banner” (i.e. it allows us to determine the mail server software and version) and other information (does it relay mail?)

<https://mxtoolbox.com/index.aspx>

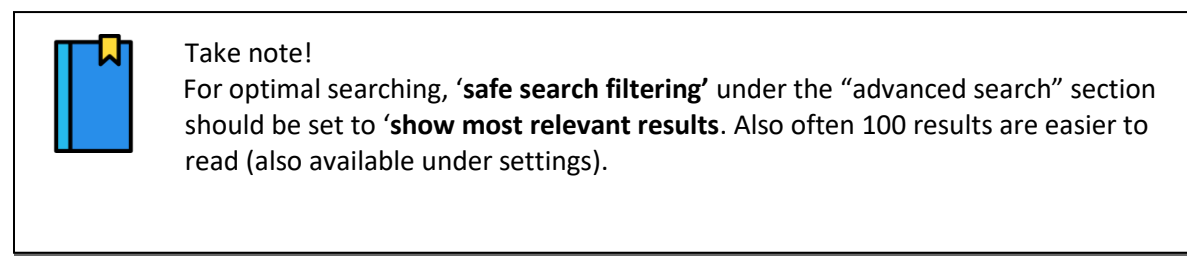
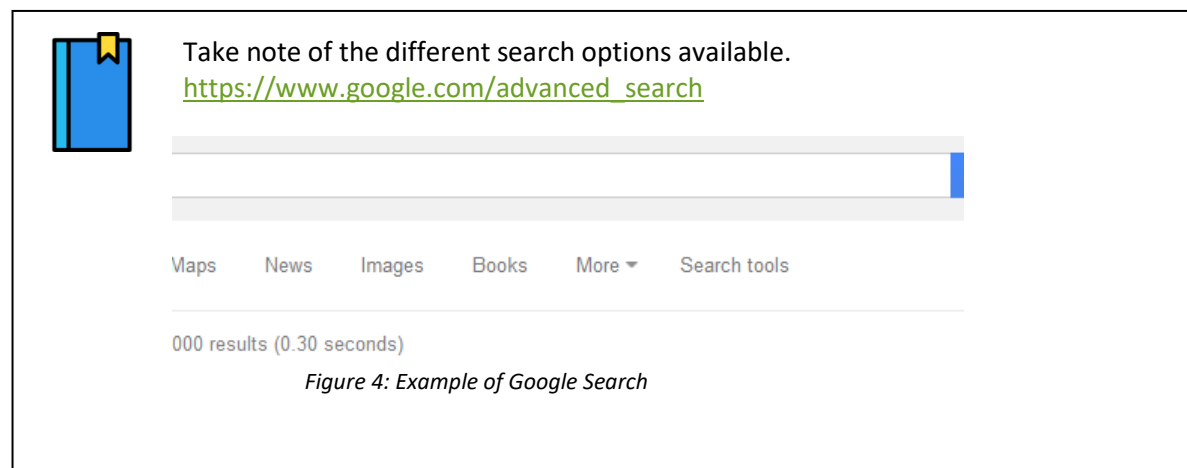
2.2.5 Google Hacking

Google hacking is the term used when a hacker tries to find exploitable targets and sensitive data by using search engines. Although Google blocks some of the better known Google hacking queries, nothing stops a hacker from crawling your site and launching the Google Hacking Database queries directly onto the crawled content.

Used effectively, Google is an invaluable tool that can help a hacker to discover information about a target or identify potential victims. The following exercises introduce you to some of the Google search techniques and tips that you can use.

2.2.5.1 The Google Interface

First of all, make sure you are familiar with the Google interface. Note this changes regularly and may not look exactly as featured in figures below!

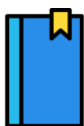


2.2.5.2 Google Hacking Basics

For reference, Table 1 below presents a list of the most common Google advanced operators.

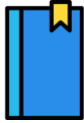
Table 1: Advanced Google Search Operator and Description

Operator	Description
Intitle	<ul style="list-style-type: none">Locates strings in the title of a pageCan be used with other operators
Allintitle	<ul style="list-style-type: none">Searches for all terms in the title of a pageNot very successful with other operators
Inurl	<ul style="list-style-type: none">Locates strings in the URL of a pageCan be used with other operators
Allinurl	<ul style="list-style-type: none">Searches for all terms in the URL of a pageNot very successful with other operators
Filetype	<ul style="list-style-type: none">Locates specific types of files based on the file extensionSynonymous with extBest used with other search information, e.g. filetype:pdf inurl:www.abertay.ac.ukCan be used with other operators
Allintext	<ul style="list-style-type: none">Locates all search terms in the text of a page
Site	<ul style="list-style-type: none">Searches within a specific site or domainCan be used with other operators or alone
Link	<ul style="list-style-type: none">Locate links to a specific site or URL (can be useful to find out which external sites are affiliated with your target)Not very successful with other operators
Inanchor	<ul style="list-style-type: none">Locates strings in the descriptive part of linksCan be used with other operators
Daterange	<ul style="list-style-type: none">Finds pages that have been Google indexed within a specific date rangeBest used with an additional search term
Numrange	<ul style="list-style-type: none">Locates a number within a particular rangeCan be used with other operators
Cache	<ul style="list-style-type: none">Displays the cached copy of a pageCan't be used with other operators
related	<ul style="list-style-type: none">Return sites that are related to a specified site or URLNot very compatible with other operators
Info	<ul style="list-style-type: none">Returns summary information about a pageNot very compatible with other operators
Author	<ul style="list-style-type: none">Returns group postings by a specific authorBest used as a group search
Group	<ul style="list-style-type: none">Searches within group namesBest used as a group search
Insubject	<ul style="list-style-type: none">Locates strings within a group postBest used as a group search
msgid	<ul style="list-style-type: none">Searches for a particular message ID in group listingBest used as a group search
define	<ul style="list-style-type: none">Returns definitions of a word or search termnot very compatible with other operators



Take note!

The syntax for using these operators is **operator:search**.



Take note!

If there are spaces in the search term that you want included in an advanced operator, use the period character which Google translates as a single wildcard character. For example, **intitle:Ethical Hacking** will return pages with 'ethical' in the title and the word hacking appearing anywhere on the page.

Intitle:ethical.hacking on the other hand will only return those pages with Ethical Hacking appearing in the title.



Questions.

Using Google, conduct some searches using the above operators to familiarise yourself with the syntax and likely results.

intitle:"live view / -AXIS"

inurl:admin/backup

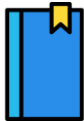
inurl:login.cfm

filetype:php inurl:"webeditor.php"

intitle:admin intitle:login

"adding new user" inurl:addnewuser -"there are no domains"

filetype:php inurl:"logging.php"



Take note!

- You can mix different operators and be imaginative with your search terms to discover some interesting information.
- Try searching for organisations that may use hardware or software with known vulnerabilities.
- Take a note of any search terms that you find particularly successful.

2.2.5.1 The Google Hacking Database

The Google Hacking Database (GHDB) is a database of queries that identify sensitive data. "Johnny iHackStuff" maintains the database.

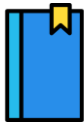


Examine the Google HACKING database from the following link (*you may have to copy and paste the link*).

<https://www.exploit-db.com/google-hacking-database>

2.2.6 Shodan Searching

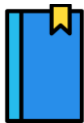
Shodan.io is a search engine that lets the user find specific types of computers (routers, servers, etc.) connected to the internet using a variety of filters. Some have also described it as a search engine of service banners, which are meta-data the server sends back to the client. This can be information about the server software, what options the service supports, a welcome message or anything else that the client can find out before interacting with the server.



Take note you will require an account on Shodan to fully use the features of the website. Thankfully, Shodan provides a free account upgrade if you register with an academic email (ac.uk). So when registering at the link below provide your University email.

<https://account.shodan.io/register>

2.2.6.1 Basic Searching



Note that SHODAN does what Google does but outputs much more specific data. Similar to Google, quotation marks can narrow a search and Boolean operators + and – can be used.

A simple search of **iis** will present many results of the IIS webserver header in various countries. Try a search for all Great Britain IIS sites by using **iis country:"GB"**.

SHODAN Explore Downloads Pricing [iis country:"GB"](#)

TOTAL RESULTS
286,092

TOP CITIES

London	191,459
Manchester	6,189
Cardiff	4,475

[View Report](#) [Download Results](#) [Hist](#)

New Service: Keep track of what you have cor

Document Moved [↗](#)

51.140.244.162
hiddenbeaches.co.uk
audley.com
discoveraudleytravel.co
m
photocomp.audleytravel.
com

SSL Certificate

Issued By:
|- Common Name:
R3
|- Organization:

H1
CC
CC
DE
SE
1.2

Figure 5: Example Shodan Output

It is also possible to search for specific versions of software such as old ones containing vulnerabilities. For example search for an older version of IIS like **iis/5.0**.

2.2.6.2 Search Filters

Some filters that can be used to narrow down a search include those featured in the Table 2.

Table 2: Excerpt Shodan Search Filters

Filters	Descriptions
Country	Filters results by two letter country code.
Hostname	Filters results by specified text in the hostname or domain.
Net	Filter results by a specific IP address, range or subnet.
OS	Search for specific operating systems
Port	Narrow the search for specific services.



Questions.

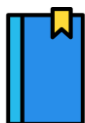
Conduct the following searches and note the results.

1. Apache
2. Apache country:GB
3. Apache hostname:.ac.uk
4. Default password

2.2.6.3 Some Shodan Searches

Below presents a list of various searches that can be used in Shodan to enumerate information on Web servers, CMS's, Network services and others.

Shodan Searches
IIS+6.0
Joomla
WordPress
Linksys
Linksys+CIT400
Cisco
port:23+ list+of+built-in+commands
. port:80+iisstart.html
Server: SQ-WEBCAM



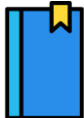
Experiment by searching for the words password, passwd, user, username etc.

Note that the + operator can be used to combine results.



Question.

Thorntons solicitors in Dundee is your target. Investigate related Shodan output, what possible paths exist to hack into it?



Select the explore tab as shown below and examine the current interesting links and search terms.

The screenshot shows the Shodan website interface. At the top, there is a navigation bar with tabs: Shodan, Maps, Images, Monitor, Developer, and More... Below this is a dark header with the Shodan logo, 'Explore' (selected), 'Downloads', and 'Pricing' with an external link icon. A search bar is on the right. The main content area is titled 'Explore' and features a grid of categories: 'Industrial Control Systems' (with an image of a factory), 'Databases' (with a glowing network diagram), and 'Network Infrastru' (with a network diagram). Below the categories, there is a 'RESEARCH' section with a card for 'Shodan 2000' and a 'BROWSE SEARCH DIRECTORY' section with a search bar labeled 'Search shared queries...'.

2.2.7 Using 192.com

Allows us to find information about a person from telephone numbers and also census information. Much of this is available free.



Go to 192.com and search for information about people.
<https://www.192.com/>

2.2.8 Accessing Archived Information

Archive.org is basically a library of Internet history. It is also a library of useful information (e.g. job adverts which reveal operating systems, programming languages etc).



Visit the website and examine St Andrew's or Abertays's web pages over the years. <https://www.archive.org/index.php>



Question.

Examine the archived vacancies over the years pages of St Andrews. How could this information be useful?

2.2.9 Multi-information Sites

There are many third party websites that give huge amounts of information.



Have a look at the tools below and search for St Andrews University and verify that the information you obtained from previous searches is consistent.

- <https://www.robtex.com/>
- <https://viewdns.info/>
- <https://centralops.net/co/>
- <https://network-tools.com>

2.2.10 Google Maps

Google maps (in particular street view) can be used for example to find information about the layout of a company, places to capture wireless without being seen or the location of the “skip” for “Dumpster Diving”.



Use Google Maps and examine Abertay University, and try to locate the skip (dumpster), routes in and out of the building(s).



Questions.

Why could this information be useful?
If tasked to monitor the wireless, where would you do so?

2.1 FREEDOM OF INFORMATION ACT

A relatively new and unusual method of finding information is by using the Freedom of Information Act. This act applies to any government associated body include government departments, local authorities, the NHS, Universities, Colleges, state schools and police forces. Much of the world has similar legislation.



The following link contain FOI requests against Abertay.
<https://www.whatdotheyknow.com/search/abertay/all>



Search for St Andrews University on this same FOI website, examine some requests to discover any useful information.

You may also wish to search for others e.g. NHS related FOIs.

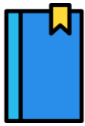
2.2 LOCAL TOOLS

The following is a selection of tools that are available from different operating systems.

2.2.1 OWASP Mantra

OWASP Mantra is a portable version of Firefox that comes loaded with different plug-ins. OWASP Mantra can be downloaded from the **tools folder** or from <https://sourceforge.net/projects/getmantra/>

- Download, and run it to install.



Take note!

You must close any Firefox windows that are open before launching Mantra.



Browse to the site you want to footprint and examine the information gathering menu, as seen in Figure 7.

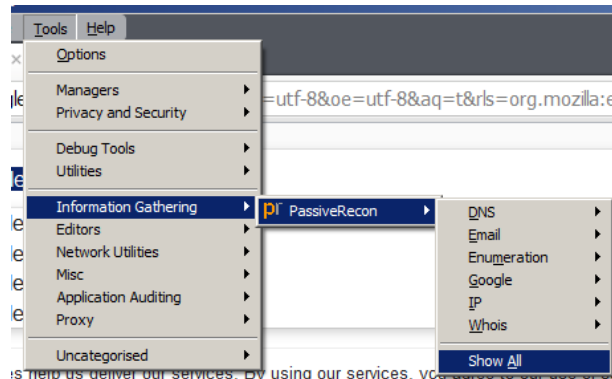
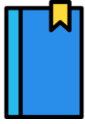


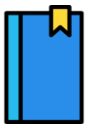
Figure 6: Example of Information Gathering menus.

2.2.1 Maltego

Maltego is an information gathering, visualisation and linking tool. Maltego is a large tool with many possible uses and can take a long time to master. The Community version that we will use has limitations (e.g. the amount of information that can be retrieved and exporting of the discovered data).



To use Maltego you must first Register an account. You can do so at the link below.
<https://www.paterva.com/web7/community/community.php>

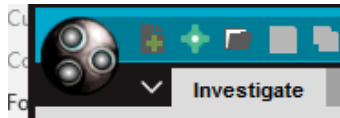


Take note that to use Maltego, you must have a valid e-mail account. You can use your Abertay account or for anonymity, you could create an e-mail address from <https://protonmail.com/> that does not require (intrusive) validation.

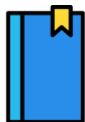


Install Maltego by downloading from the **tools folder** (or <https://www.maltego.com/downloads/>), and then run Maltego **selecting Maltego CE** before (registering and) logging in.

- Select **open a blank graph to play around** from the wizard or Select **New** from the Global button at the top left, seen in Figure 8.



Think of a **New graph** as a new investigation.

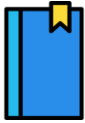


Here are some **IMPORTANT** directions and editing features that are useful:-

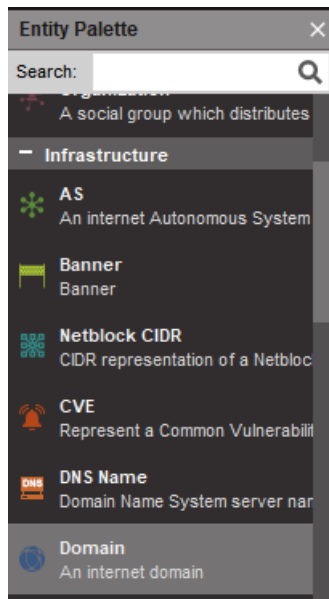
- To add an entity to the graph, **drag on to the graph**.
- To change properties, **double-click on the entity**.
- **Right-Click and drag** to move around.
- The **wheel** Zooms in/out

2.2.1.1 Information Gathering Techniques against a Target

The following tutorial will illustrate how we can gather information about a website using Maltego.



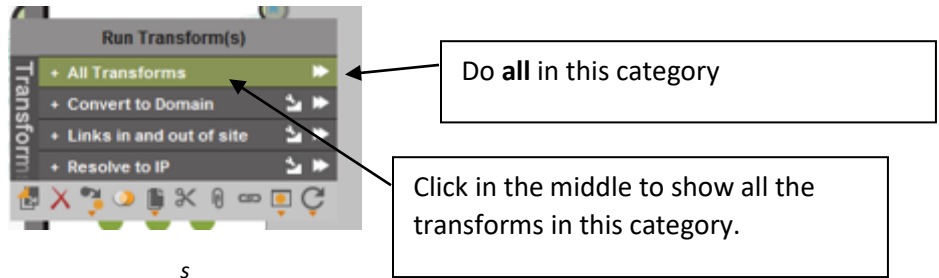
From the **Entity Palette** menu on the left, Select **Infrastructure**, then find **Website (it's down a little)**. Now drag a **Website** entity on to the graph and rename to **www.st-andrews.ac.uk**.



- **Right-Click** on the and the **Run Transforms** menu should appear as shown

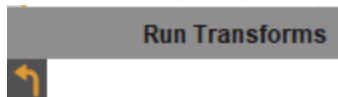


- To select a transform **right-click** on the website entity.



5

- Select **Resolve to IP** and run the transform to get the IP address of the website.
- Also, right-click on the **Website entity**, go up a level and then select **Convert to Domain**. Note the Back arrow must be used to get back to choose this.

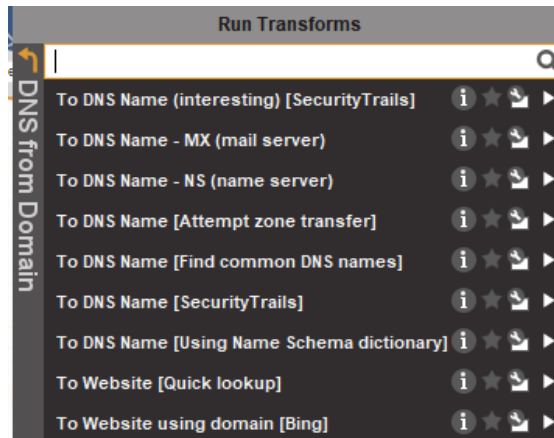


We should now have three entities that we can investigate. I.e. **The website, the IP address and the TLD** (top level domain).

- Right-Click on the **IP address** entity and Select the **All Transforms menu** as shown above and select the **IP Owner detail but this time select all**.



- From the domain **entity** (st-andrews.ac.uk), right-click and you should see the Run Transforms as shown below: -



Try the following: -

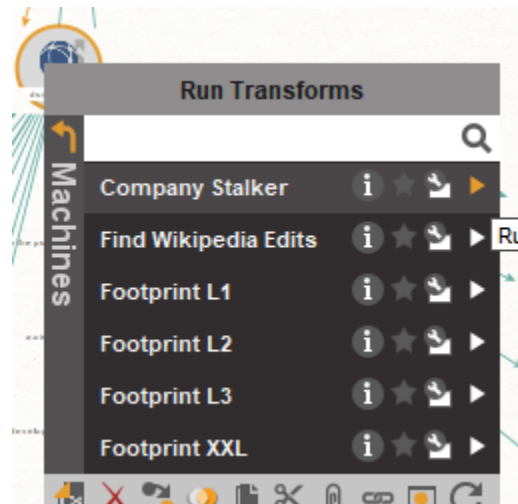
- Get the MX servers (i.e. the mail servers)
- Get the NS servers (i.e. the DNS name servers)
- Experiment with the menu shown below against the domain entity to see what information can be obtained. in.

- Select the **domain** and **right-click**. From the top level menu (i.e. Go back), select **Machines**.



Machines will run multiple transforms,

- Select **Company Stalker**!



Note that the machines can take some time to run.

- **Now experiment with Maltego.**

Some related links.

Geolocating using Maltego <https://www.maltego.com/blog/precise-geospatial-link-analysis-using-maltego/>

Maltego tutorial play list - https://www.youtube.com/watch?v=GwBiJqa_nEc&list=PLfRX-xJAc2yz6CjQVQuogJeCBoy8HbCOR&index=4&ab_channel=Maltego

3 FURTHER RESOURCES

This section includes additional reference material and extracurricular exercises, allowing for eager students to apply the knowledge gained in class outwith the scope of the module.

3.1 COPYCAT DOMAIN NAMES



Use the following link to look for copycat domain names for Abertay.

<https://research.domaintools.com/phisheye/>

3.2 IMMERSIVE LABS

Some content found on the Immersive Labs platform (via dca.immersivelabs.online) shares similar material to that covered in the lecture and lab exercises today. You must register for this.

- Geolocation
- Domain Intel
- Reverse Image Search
- EXIF
- Shodan.io
- Investigator Operations Security
- Cached and Archived Websites.
- OSINT: Deleted Tweet
- OSINT: Boarding Pass

3.3 PIMEYES

Pimeyes is a reverse image search engine. You can submit a picture to TinEye to seek out wherever it came from. TinEye uses neural networks, pattern recognition, machine learning, and image recognition technology instead of keywords or metadata.

Link: <https://pimeyes.com/en>

Very similar to this is tineye. Link: <https://www.tineye.com>

3.4 OSINT – VM (THIS IS A MAJOR PROJECT).

The Trace Labs team created a specialized OSINT VM specifically to bring together the most effective OSINT tools and customized scripts we saw being used during our Search Party CTF's. The Trace Labs OSINT VM is a quick way to get started and have access to the most popular OSINT tools and scripts all neatly packaged under one roof.

Link: <https://www.tracelabs.org/initiatives/osint-vm>

3.5 FURTHER READING

Beginner's Guides - ICANN. Iann.org. (2020). Retrieved 11 August 2020, from <http://www.icann.org/en/about/learning/beginners-guides>.

Better Whois: The WHOIS domain search that works with all registrars.. Betterwhois.com. (2020). Retrieved 11 August 2020, from <http://www.betterwhois.com/>.

DNS Oversimplified: How to check your DNS. Rscott.org. (2020). Retrieved 11 August 2020, from <http://www.rscott.org/dns/>.

GeekTools. Geektools.com. (2020). Retrieved 11 August 2020, from <http://geektools.com/whois.php>.

Ptgmedia.pearsoncmg.com. (2020). Retrieved 11 August 2020, from http://ptgmedia.pearsoncmg.com/images/9780789735317/samplechapter/0789735318_CH03.pdf.