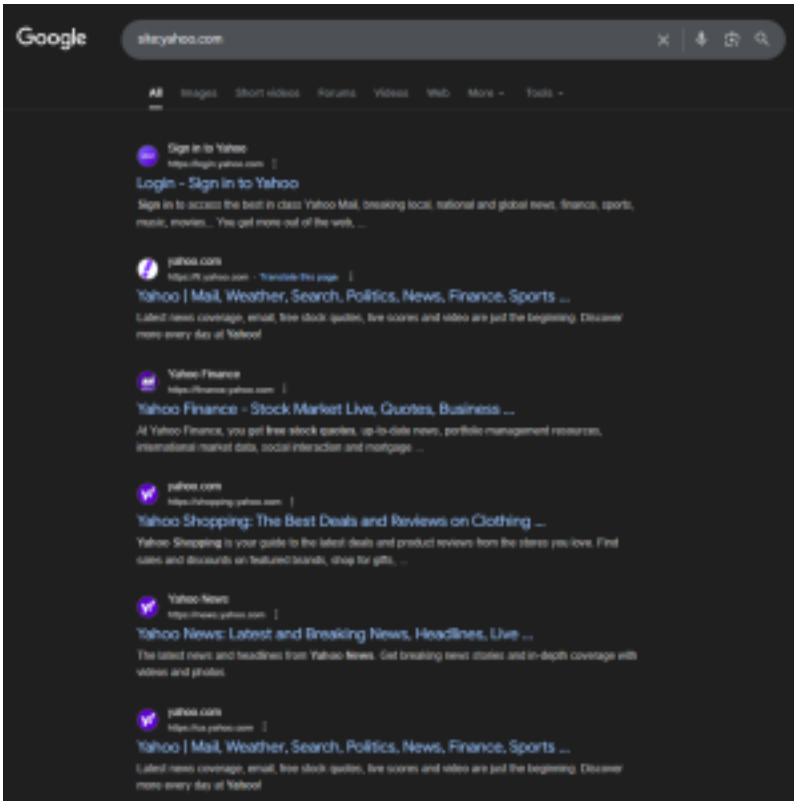
NTI – Ethical Hacking
Assignment 1 Report
Information gathering “reconnaissance”

Prepared by:
Hadeer Amr

For my first task, I was required to select a target and conduct an information gathering phase. This involved identifying a suitable domain, applying reconnaissance techniques, and collecting relevant data to better understand the target’s structure and potential attack surface.

My Target was : yahoo.com and tools I have used : sublist3r, amass,

theHarvester Also I tried searching in google and I have found subdomains

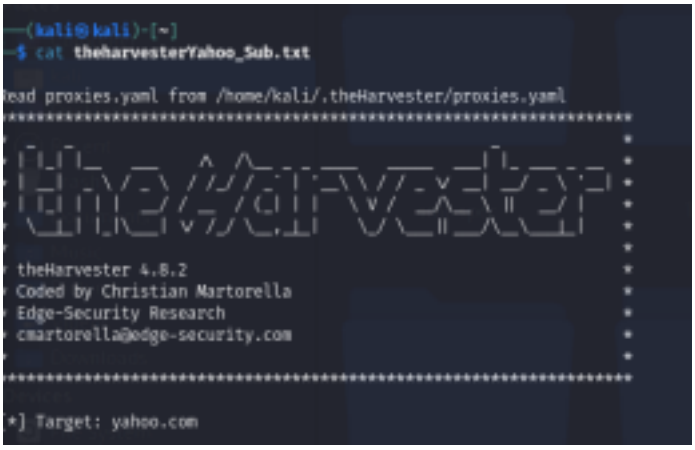


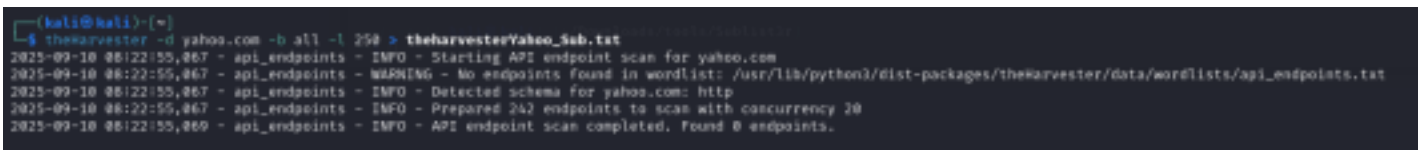
Commands:
theHarvester -d yahoo.com -b urlscan -l 200 > urlscanYahoo.txt
theHarvester -d yahoo.com -b all -l 250 > theharvesterYahoo_Sub.txt
python3 sublist3r.py -d yahoo.com > Sublist3r_subs.txt
python3 sublist3r.py -d yahoo.com -e urlscan
amass enum -passive -d yahoo.com > amassyahoo.txt “most one takes time”
cat theharvesterYahoo_Sub.txt | sort -u > cleaned.txt

theHarvester
Assignment 1

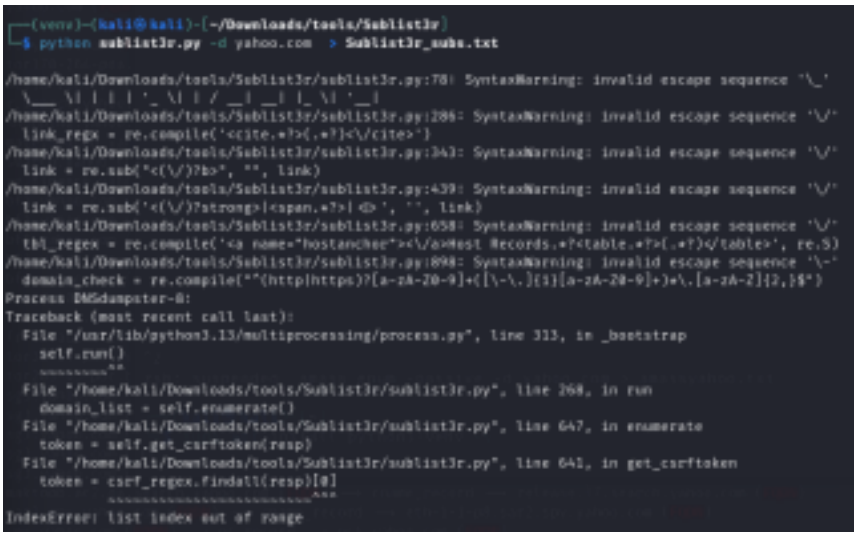
Frist, I started with discovering the three tools using **-h** option

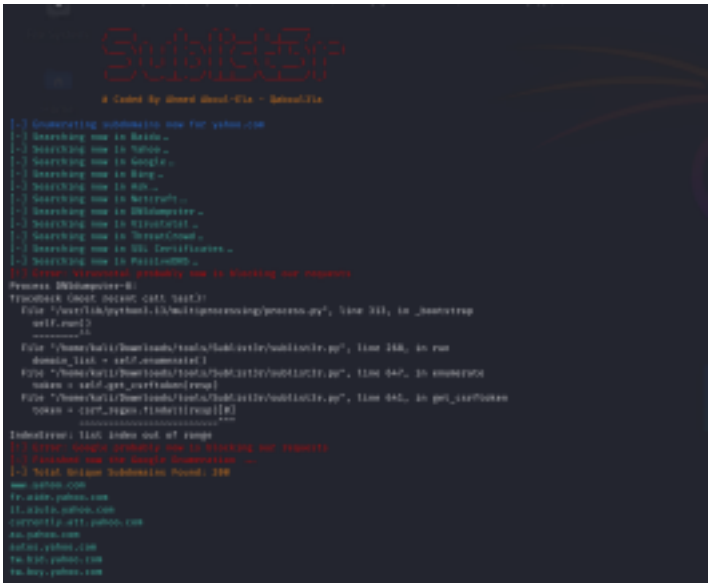
And those are my steps running the commands

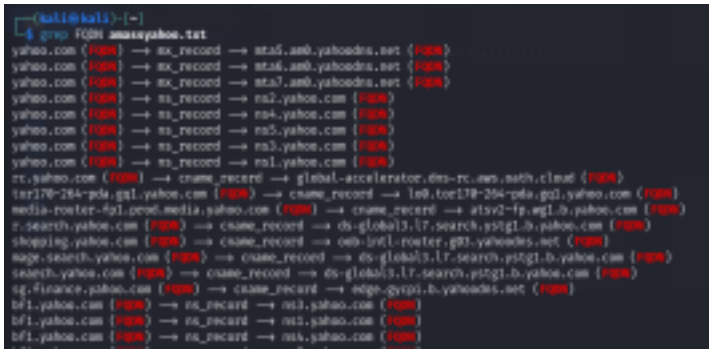




theHarvester
Assignment 1







Amass tool gave me a lot of information but was not organized and I found some duplicates to make it more easier I searched for script to automate the sorting and organizing and I used this one to gave me excel sheet and I attached it in the link that contain my findings in this stage

theHarvester
Assignment 1

```
┌──(kali㉿kali)-[~]
└─$ chmod 777 parse_dns.sh

┌──(kali㉿kali)-[~]
└─$ ./parse_dns.sh
[+] Done! Saved to dns_records.csv
```

```
  GNU nano 8.2                                  parse_dns.sh
#!/bin/bash
input="amassyahoo.txt"
output="dns_records.csv"

# Write CSV header
echo "Record Type,Domain,Target" > "$output"

# Parse lines
awk -F' ⟶ ' '{
    domain=$1
    record=$2
    target=$3
    gsub(/\(FQDN\)/,"",domain)
    gsub(/\(FQDN\)/,"",target)
    gsub(/ /,"",record)
    gsub(/^ +| +$/,"",domain)
    gsub(/^ +| +$/,"",target)
    print record "," domain "," target
}' "$input" >> "$output"

echo "[+] Done! Saved to $output"
```

	Record Type	Domain	Target
2	mx_record	yahoo.com	mta5.am0.yahoodns.net
3	mx_record	yahoo.com	mta6.am0.yahoodns.net
4	mx_record	yahoo.com	mta7.am0.yahoodns.net
5	ns_record	yahoo.com	ns2.yahoo.com
6	ns_record	yahoo.com	ns4.yahoo.com
7	ns_record	yahoo.com	ns5.yahoo.com
8	ns_record	yahoo.com	ns3.yahoo.com
9	ns_record	yahoo.com	ns1.yahoo.com
10	cname_record	rc.yahoo.com	global-accelerator.dns-rc.aws.oath.cloud
11	cname_record	tor170-264-pda.gq1.yahoo.com	lo0.tor170-264-pda.gq1.yahoo.com
12	cname_record	media-router-fp1.prod.media.yahoo.com	atsv2-fp.wg1.b.yahoo.com
13	cname_record	r.search.yahoo.com	ds-global3.l7.search.ystg1.b.yahoo.com
14	cname_record	shopping.yahoo.com	oob-intl-router.g03.yahoodns.net

The reconnaissance uncovered a wide range of Yahoo's digital footprint information I have found Subdomains & FQDNs (over 200 entries via Sublist3r and Amass), PTR Records (reverse DNS mappings), Email Infrastructure (mail server hosts), IPv4 and IPv6 Addresses, Netblocks & ASNs (infrastructure ownership and ranges), Emails (4 public addresses via theHarvester)

I utilized three tools Sublist3r, Amass, and theHarvester to enumerate Yahoo's public-facing infrastructure and gather OSINT data. The results are summarized as follows:

Using Sublist3r it Successfully enumerated 200+ subdomains associated with yahoo.com and these subdomains represent different parts of Yahoo's infrastructure (mail servers, CDN nodes, internal naming conventions, etc.).

And using Amass provided deeper enumeration and infrastructure mapping, including: FQDNs (Fully Qualified Domain Names)examples include oxy-oxygen-2001-4998-44-803f-1184.ne1.yahoo.com and 218.43.30.72.in addr.arpa (reverse DNS format) and PTR Records (Pointer Records): Mapping IPs back to hostnames example 218.43.30.72.in-addr.arpa → unknown.yahoo.com and Subdomains: Examples include loo.tor394-300-pda.bf1.yahoo.com and sonic312-7.consmr.mail.ir2.yahoo.com and Email-related Hosts: Such as sonic312-7.consmr.mail.ir2.yahoo.com (part of Yahoo Mail servers) and IPv4 and IPv6 addresses as IPv4: 72.30.30.237, 216.155.192.212 and IPv6: 2001:4998:ef99:209::2019 and Netblocks (IP Ranges): Examples include 72.30.0.0/19 and 2001:4998::/32 and ASNs Identified Autonomous System Numbers related to Yahoo infrastructure.

And finally using theHarvester I Collected 4 public email addresses linked to yahoo.com and unfortunately No employee names or LinkedIn profiles were discovered limited internal data exposure from this tool.

This information provides a comprehensive view of Yahoo's external attack surface and forms the

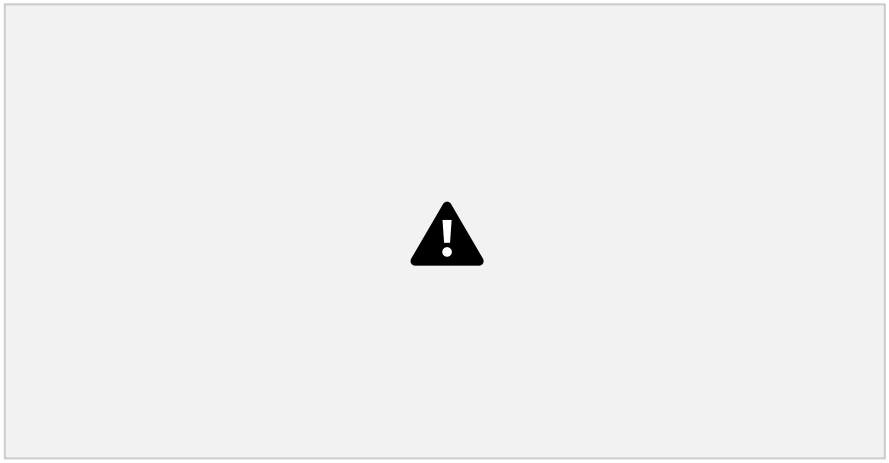
baseline for further vulnerability assessment.

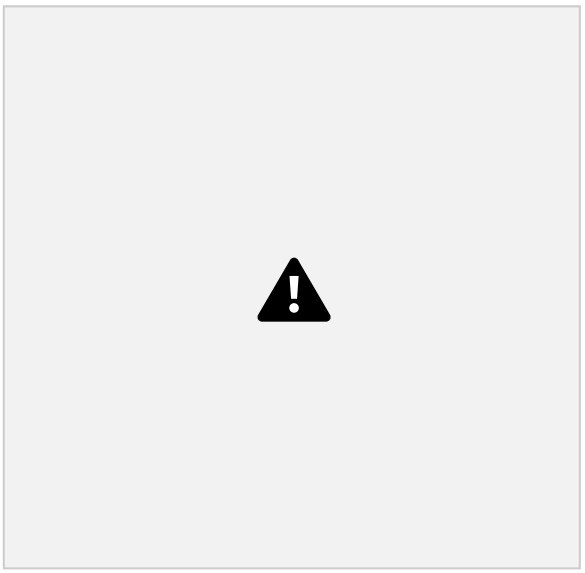
theHarvester
Assignment 1

This is a list of all the information I retrieved is uploaded in this link

https://drive.google.com/drive/folders/1TKoyzk-3jfsYPuZfYQEvvPrXrRStWs8W?usp=sharing

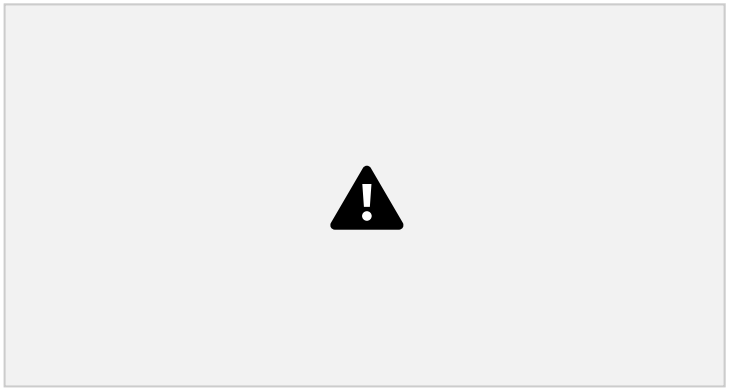
Using multiple sources truly affect the results for example using theHarvester -b all gave me 4 emails, 1 URL, 6171 hosts, 172 Ips and 19 ASNS while -b urlscan gave me 19 ASNS, 2 URLs, 60 Ips and 2 hosts and some resources does not gave me any results



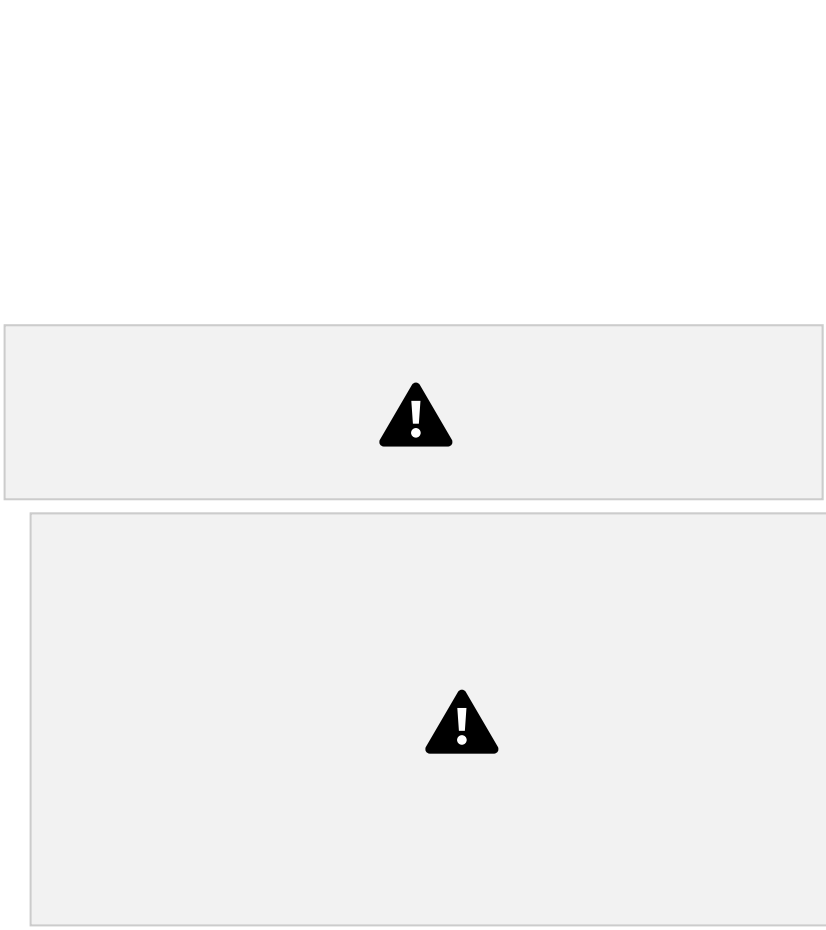


The source engine gave me the most useful results was amass the global default option and theHarvester -b all option gave me wider information and results that could help me more in my upcoming phases, but I thought it will differ based on the target.

I have found some duplicate entries across engines especially using amass tool after some thinking I thought about using sort using this ***cat theharvesterYahoo_Sub.txt | sort -u > cleaned.txt*** also another idea to enhance my script to make amass results better



or use this ***awk -F',' '!seen[$2]++' dns_records.csv > cleaned.csv*** also I can use feature in excel to remove the duplicates and I can gather all the information from different tools and engines and but in one file then perform the cleaning to make results clearer.



Based on the data I have found an attacker can use some of this data in phishing or social engineering attacks as the four e-mails I have found using theHarvester and the Email-related Hosts I found using amass tool or he could find LinkedIn of employees could help in getting more information to use in
phishing or social engineering attacks or any attacks related to what he could found also based on the subdomain he can perform some searching to find vulnerabilities and got more access

To sum up what I have learned from this task I have learned that using different tools in reconnaissance is very important because each one shows different parts of the target. For example, Sublist3r was very good at finding many subdomains quickly, Amass gave me a deeper view of the infrastructure, and theHarvester helped with OSINT, like e-mails. I also learned that some data was repeated across the tools, but this was useful because it confirmed the information and made it more reliable. At the same time, I found unique results that only one tool was able to discover. Through this task, I understood better how domains, subdomains, FQDNs, PTR records, and netblocks are all connected, and how they reveal the structure of a company's online presence. Even though theHarvester gave me only four emails and no employee names, I realized that even a small amount of data can still be used in attacks like

phishing.
Another thing I noticed is how big companies like Yahoo manage very complex infrastructures. I found more than 200 subdomains, and this showed me the importance of organizing and cleaning the data so it can be used effectively.
Finally, this task gave me a strong foundation for the next steps in penetration testing, such as vulnerability scanning and exploring the attack surface in more detail.

Disclaimer
This report is for educational purposes only.