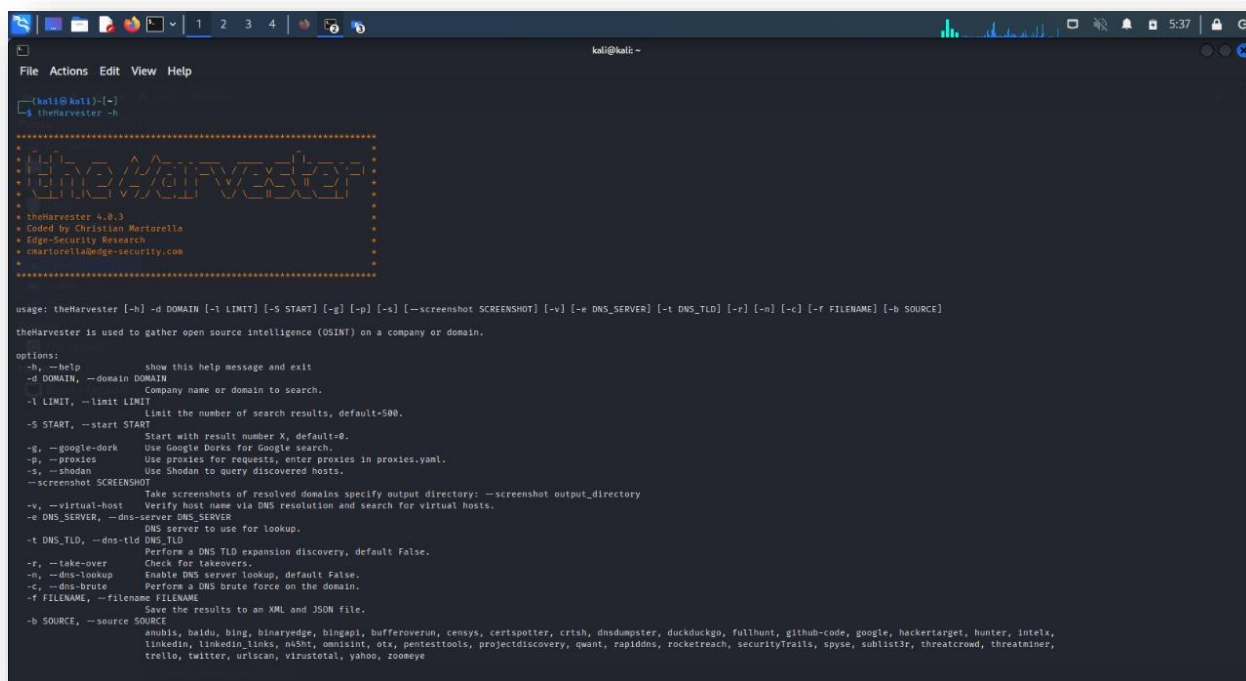


Report about The Harvester tool on Linux

Prepared by: Ahmed Moataz Abdelmonem Elrefaey

- Target Domain **Microsoft.com**
- The help menu of the Harvester tool
theHarvester -h



```
kali@kali:~$ theHarvester -h

theHarvester

theHarvester 4.0.3
Coded by Christian Martorella
Edge-Security Research
cmartorell@edge-security.com

usage: theHarvester [-h] -d DOMAIN [-l LIMIT] [-s START] [-g] [-p] [-s] [--screenshot SCREENSHOT] [-v] [-e DNS_SERVER] [-t DNS_TLD] [-r] [-n] [-c] [-f FILENAME] [-b SOURCE]

theHarvester is used to gather open source intelligence (OSINT) on a company or domain.

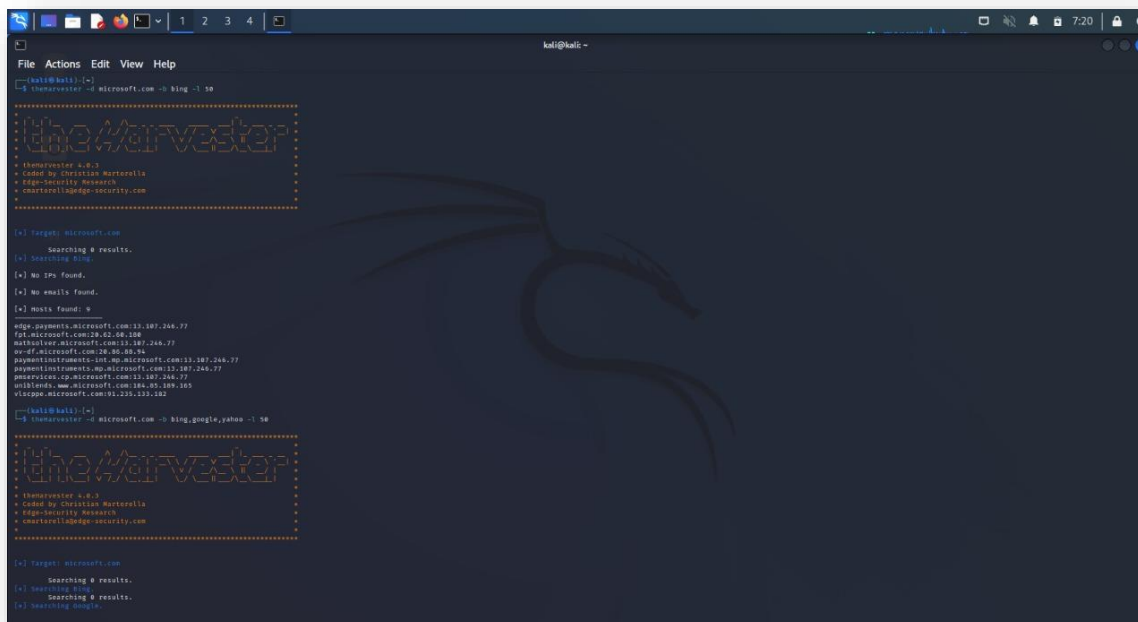
options:
  -h, --help            show this help message and exit
  -d DOMAIN, --domain DOMAIN
                        Company name or domain to search.
  -l LIMIT, --limit LIMIT
                        Limit the number of search results, default=500.
  -s START, --start START
                        Start with result number X, default=0.
  -g, --google-dork      Use Google Dorks for Google search.
  -p, --proxies          Use proxies for requests, enter proxies in proxies.yaml.
  -s, --shodan           Use Shodan to query discovered hosts.
  --screenshot SCREENSHOT
                        Take screenshots of resolved domains specify output directory: --screenshot output_directory
  -v, --virtual-host     Verify host name via DNS resolution and search for virtual hosts.
  -e DNS_SERVER, --dns-server DNS_SERVER
                        DNS server to use for lookup.
  -t DNS_TLD, --dns-tld DNS_TLD
                        Perform a DNS TLD expansion discovery, default False.
  -r, --take-over        Check for takeovers.
  -n, --dns-lookup       Enable DNS server lookup, default False.
  -c, --dns-brute        Perform a DNS brute force on the domain.
  -f FILENAME, --filename FILENAME
                        Save the results to an XML and JSON file.
  -b SOURCE, --source SOURCE
                        shodan, baidu, bing, binaryedge, bingapi, bufferoverrun, censys, certspotter, crish, dnsdumpster, duckduckgo, fullhunt, github-code, google, hackertarget, hunter, intel,
                        linkedin, linkedin_links, n4sh1t, omnisint, otx, pentesttools, projectdiscovery, quant, rapiddns, rocketreach, securityTrails, spyse, sublist3r, threatcrowd, threatminer,
                        trello, twitter, urlscan, virustotal, yahoo, zoomeye
```

- What types of information did you find?

I found several subdomains, but I didn't find any emails because most companies remove the important emails for security reasons, but we can find some services emails using search engines like "hunter" using private API.

- subdomains I retrieved with IPs hosting that service.

theHarvester -d microsoft.com -b bing -l 50



```

kali@kali:~$ theHarvester -d microsoft.com -b bing -l 50
=====
theHarvester
=====
theHarvester v.2.0.3
  coded by Christian Martorella
  eugen-security Research
  cmartorell@eugen-security.com

[*] target: microsoft.com
    Searching # results.
[*] Searching bing.
[*] No IPs found.
[*] No emails found.
[*] Hosts found: 9
edge.payments.microsoft.com:13.107.246.77
fpt.microsoft.com:20.62.60.180
mathsolver.microsoft.com:13.107.246.77
ov-df.microsoft.com:20.86.88.94
paymentinstruments-int.mp.microsoft.com:13.107.246.77
paymentinstruments.mp.microsoft.com:13.107.246.77
prmservices.cp.microsoft.com:13.107.246.77
uniblends.www.microsoft.com:184.85.189.165
vlspepe.microsoft.com:91.235.133.182

kali@kali:~$ theHarvester -d microsoft.com -b bing,google,yahoo -l 50
=====
theHarvester
=====
theHarvester v.2.0.3
  coded by Christian Martorella
  eugen-security Research
  cmartorell@eugen-security.com

[*] target: Microsoft.com
    Searching # results.
[*] Searching bing.
[*] Searching google.
  
```

[*] Hosts found: 9

```

-----
edge.payments.microsoft.com:13.107.246.77
fpt.microsoft.com:20.62.60.180
mathsolver.microsoft.com:13.107.246.77
ov-df.microsoft.com:20.86.88.94
paymentinstruments-int.mp.microsoft.com:13.107.246.77
paymentinstruments.mp.microsoft.com:13.107.246.77
prmservices.cp.microsoft.com:13.107.246.77
uniblends.www.microsoft.com:184.85.189.165
vlspepe.microsoft.com:91.235.133.182
  
```

- Which source engine gave the most useful results?

I found that Bing is useful if I want to search for subdomains but for searching about emails the hunter is the most useful for this.

- How did using multiple sources affect the results?

theHarvester -d microsoft.com -b bing,google,yahoo -l 50

A screenshot of a Kali Linux terminal window. The terminal shows a user at the kali@kali- host running a command to search for information about 'TheHarvester'. The output displays search results from various sources like GitHub, Stack Overflow, and other websites, listing IP addresses and domains associated with the tool. A large, faint watermark of a cat's face is visible in the background of the terminal output.

```
kali@kali:~$ sudo theharvester -d microsoft.com --bing google,yahoo --port 8080
```

```
theHarvester
```

```
theHarvester v.0.9.0  
Created by Christian Martello  
FPGA Security Research  
cmartello@fpga-security.com
```

```
[+] Target: microsoft.com
```

```
Searching & results.
```

```
[+] Searching Bing.
```

```
Searching & results.  
Searching new results.  
Google is blocking your ip and the workaround, returning  
Searching new results.  
[+] Searching Google.
```

```
[+] No IPs found.
```

```
[+] No emails found.
```

```
[+] Hosts found: 22
```

```
account.microsoft.com|206.42.175  
answers.microsoft.com|13.107.248.77  
blogs.microsoft.com|161.193.213.21, 161.193.213.24  
com-dynmedia-3.microsoft.com|93.281.111.134, 93.281.111.144  
developer.microsoft.com|2.20.48.227  
edge-payment.microsoft.com|17.187.248.77  
ftp.microsoft.com|20.62.48.108  
learn.microsoft.com|2.20.62.214  
mathsolvers.microsoft.com|13.107.248.77  
microsoft.com|13.107.248.77  
myacctauth.microsoft.com|20.106.177.81, 20.106.177.145, 20.106.167.38, 20.106.167.34, 20.106.167.35, 20.106.177.16, 20.106.177.17, 20.106.167.39  
mysignin.microsoft.com|20.106.177.145, 20.106.167.38, 20.106.167.34, 20.106.167.35, 20.106.177.18, 20.106.177.17, 20.106.167.39, 20.106.177.83  
paytmicrosoft.com|20.80.48.6  
payments.microsoft.com|13.107.248.77  
paymentinstrumentss.ms.microsoft.com|13.107.248.77  
smarterprice-up.microsoft.com|13.107.248.77  
support.microsoft.com|13.107.248.77  
techcommunity.microsoft.com|161.193.213.21  
enkibooks.www.microsoft.com|161.193.213.21  
visualstudio.microsoft.com|23.106.106.48  
vscode.microsoft.com|13.107.248.77  
www.microsoft.com|2.20.42.126
```

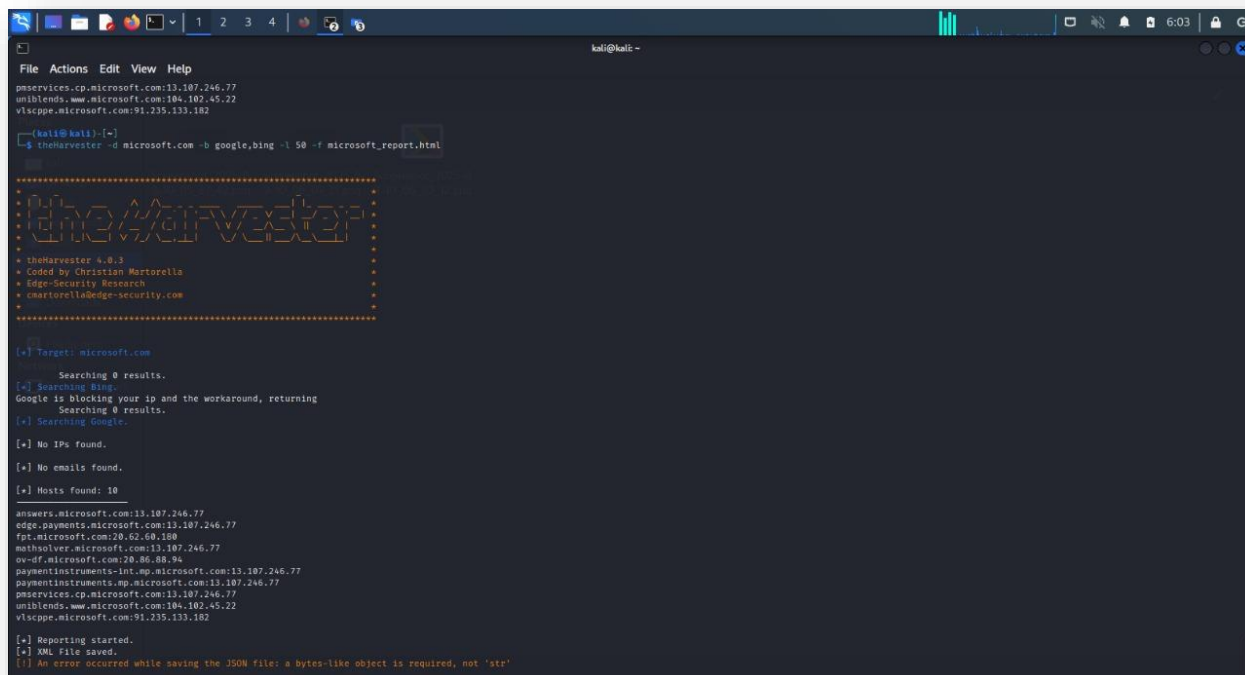
```
kali@kali:~$
```

[*] Hosts found: 22

account.microsoft.com:2.20.42.73
answers.microsoft.com:13.107.246.77
blogs.microsoft.com:141.193.213.21, 141.193.213.20
cdn-dynmedia-1.microsoft.com:95.101.114.130, 95.101.114.144
developer.microsoft.com:2.20.40.227
edge.payments.microsoft.com:13.107.246.77
fpt.microsoft.com:20.62.60.180
learn.microsoft.com:2.20.41.214
mathsolver.microsoft.com:13.107.246.77
msrc.microsoft.com:13.107.246.77
myaccount.microsoft.com:20.190.177.81, 20.190.177.145, 20.190.147.38, 20.190.147.34, 20.190.147.35, 20.190.177.18, 20.190.177.17, 20.190.147.39
mysignins.microsoft.com:20.190.177.145, 20.190.147.38, 20.190.147.34, 20.190.147.35, 20.190.177.18, 20.190.177.17, 20.190.147.39, 20.190.177.81
ov-df.microsoft.com:20.86.88.94
paymentinstruments-int.mp.microsoft.com:13.107.246.77
paymentinstruments.mp.microsoft.com:13.107.246.77
pmservices.cp.microsoft.com:13.107.246.77
support.microsoft.com:13.107.246.77
techcommunity.microsoft.com:184.85.177.101
uniblends.www.microsoft.com:184.85.189.165
visualstudio.microsoft.com:23.194.30.68
vlsccpe.microsoft.com:91.235.133.182

- To save output in XML or JSON file

theHarvester -d microsoft.com -b google,bing -l 200 -f microsoft_report.html



```
kali@kali:~$ theHarvester -d microsoft.com -b google,bing -l 200 -f microsoft_report.html

theHarvester
=====
* theHarvester 5.0.3
* Coded by Christian Martorella
* Edge-Security Research
* cmartorellabedge-security.com
*

[*] Target: microsoft.com
[*] Searching 0 results.
[*] Searching Bing:
Google is blocking your ip and the workaround, returning
[*] Searching 0 results.
[*] Searching Google:
[*] No IPs found.
[*] No emails found.
[*] Hosts found: 10
answers.microsoft.com:13.107.246.77
edge.payments.microsoft.com:13.107.246.77
fpt.microsoft.com:20.62.60.180
methodizer.microsoft.com:13.107.246.77
ov-df.microsoft.com:20.86.88.94
paymentinstruments-int.mp.microsoft.com:13.107.246.77
paymentinstruments.mp.microsoft.com:13.107.246.77
pservices.cp.microsoft.com:13.107.246.77
uniblends.www.microsoft.com:104.102.45.22
uniblends.www.microsoft.com:104.102.45.22
viscppe.microsoft.com:91.235.133.182

[*] Reporting started.
[*] XML File saved.
[!] An error occurred while saving the JSON file: a bytes-like object is required, not 'str'
```

- Are there duplicate entries across engines? How can you filter or clean that?

Yes, there are duplicate entries across engines, but we can filter or clean it using this command:

sort results.txt | uniq > clean_results.txt

“clean_results.txt” this is the file after filtering.

“results.txt” this file is the file I stored all hosts in (before filtering).

I stored the output of any command using this command:

Cat > results.txt

- How much internal data (like employee names or emails) was exposed?

No important internal data because emails and subdomains which out is services data (public data) but we should focus on the mails comes to these emails because any hacker can use these emails to send phishing mail and hackers can use subdomains and this is the first point to attack using social engineering.

- What did you learn?

By using theHarvester I learned how OSINT tools can collect information about a target domain from public sources. The tool helped me find some subdomains and see which IP addresses they resolve to. I also noticed that emails are not always easy to find because many search engines block automated queries. Overall, I understood that theHarvester is useful for basic recon, but it has limits and should be combined with other tools like Sublist3r or Amass to get a more complete picture.

- Compare theHarvester results with another tool like Sublist3r or amass. Do they match?

theHarvester: depends on search engines and some APIs, so it usually finds only a small number of subdomains.

Sublist3r: a bit better than theHarvester, it can find more subdomains but still limited.

Amass: the most powerful tool, it can discover hundreds or even thousands of subdomains because it uses many different sources and ASN enumeration.