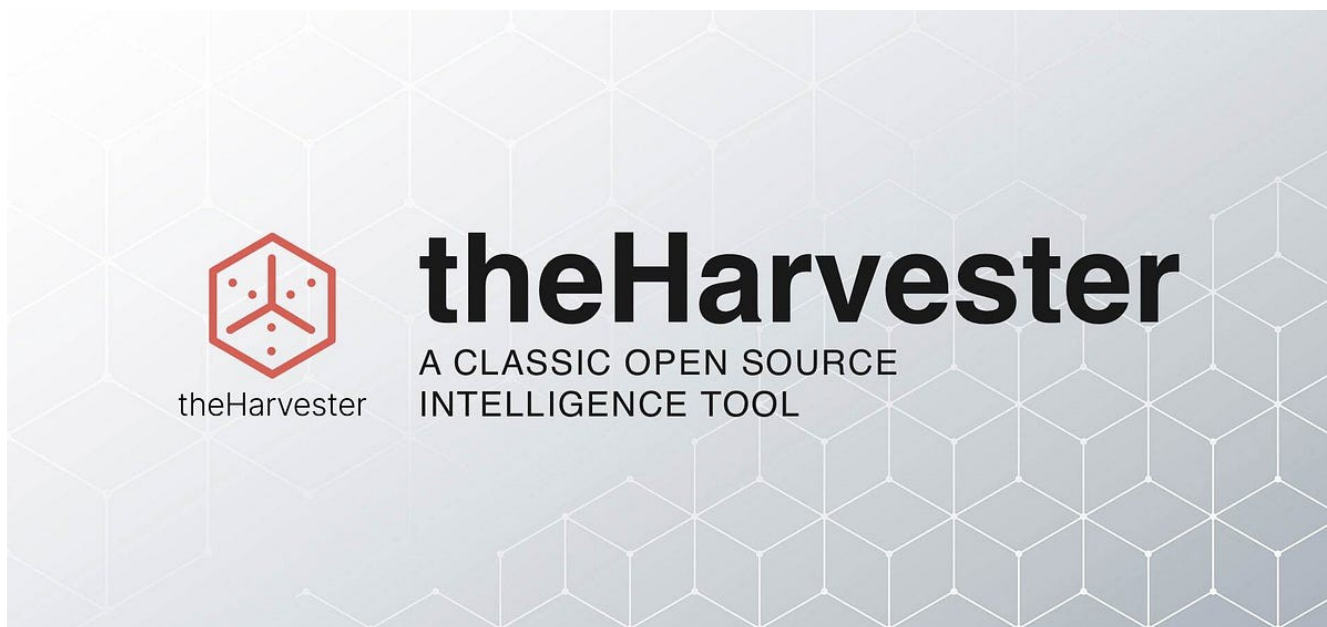




ETHICAL HACKING

FOOTPRINTING & RECONNAISSANCE

theHarvester



Researcher By: Nin Kanong

theHarvester Tool

What is TheHarvester?

TheHarvester is a powerful and widely used OSINT (Open Source Intelligence) tool that helps Ethical Hackers, Penetration Testers, and cybersecurity professionals collect information related to target domains and organizations from various search engines, databases, and other publicly available services. The primary function of theHarvester is to gather critical information about a target domain, such as:

- **Subdomains:** Alternative domain addresses linked to the target.
- **Emails:** Employee or organizational email addresses. One of the most valuable pieces of information you can collect during a reconnaissance phase is a list of email addresses. These emails can later be used for social engineering attacks or identifying potential weak points in security configurations.
- **IP Addresses:** TheHarvester can map domain names to associated IP addresses. These addresses can be used for further network scanning and vulnerability analysis.
- **Hostnames:** Additional domain names or services linked to the target.

This information is primarily used during the reconnaissance phase of penetration testing or when conducting security assessments. It gathers information without directly interacting with the target system, minimizing detection.

Main Options

Option	Description
-d	Target domain (e.g., example.com)
-b	Data source (e.g., bing, duckduckgo, hunter)
-l	Limit the number of search results (default: 100)
-s	Start with result number x (use with -l)
-f	Save output to an HTML file
-v	Verbose mode
-h	Show help message
-c	Save to a CSV file
-e	Use specified DNS server for resolving (e.g., 8.8.8.8)
-n	Do not resolve IPs
-t	Take a screenshot of discovered hosts
-r	Perform DNS resolution on found hosts
-p	Perform DNS reverse lookup
-S	Use SSL for web requests
-q	Do not print banner or warnings
-k	Use SHODAN API to search for discovered IPs
-screenshot	Take screenshots of IPs/hosts found (requires Selenium setup)

Step-by-Step Guide to Using TheHarvester

Step 1: Installing TheHarvester

The first step is to install theHarvester tool on your system. It is commonly available on Kali Linux, but you can also install it manually on other Linux distributions or even on Windows.

1. **Kali Linux:** TheHarvester comes pre-installed on most penetration testing distributions like Kali Linux.

You can verify the installation by using:

[illegible]

If you wish to update to the latest version or need to install it, you can use the following commands:

```
- kali> sudo apt install theHarvester //used to install theHarvester in kali
```

2. Installing on Other Linux Distributions:

If you are using a non-Kali Linux OS, you can clone the official repository from GitHub and install it manually:

```
- kali> git clone https://github.com/laramies/theHarvester.git
```

```
- kali> sudo pip3 install -r requirements.txt
```

Once installed, you can check whether it's running properly by executing:

- kali> theHarvester -h

This will display the help menu showing the available options and arguments.

[illegible]

Step 2: Understanding the Syntax and Basic Options

theHarvester tool has a wide range of options that you can use to tailor your search.

Let's explore some common syntax and options:

Basic Command :

- kali> theHarvester -d -b

Common options:

Here's a breakdown of the most commonly used options in TheHarvester:

- **-d <domain>** : Specifies the domain to search.
- **-b <source>** : Defines the data source (e.g., yahoo, bing, shodan). You can specify multiple sources separated by commas.
- **-l <limit>** : Limit the number of results fetched from the data sources.
- **-f <filename>** : Save the output into a file (in HTML format).
- **-n** : Perform DNS enumeration using search results.
- **-t** : Perform DNS TLD expansion.
- **-s <start>** : Start with a specific result number (useful when you want to skip initial results).
- **-v** : Enable verbose mode for more detailed output.

[illegible]

What it does: This command collects information about the specified domain (infosectrain.com) from all supported search engines and data sources.

Step 3: Dig Deeper with Verbose Mode

- **kali> theHarvester -d infosectrain.com -b yahoo,bing -l 100 -v**

```
(root@kali)-[/home/kali]
└─$ theHarvester -d infosectrain.com -b yahoo,bing -f results
Read proxies.yaml from /root/.theHarvester/proxies.yaml
*****
*                                     *
* theHarvester                       *
*                                     *
* theHarvester 4.6.0                 *
* Coded by Christian Martorella      *
* Edge-Security Research             *
* cmartorella@edge-security.com      *
*                                     *
*****

[*] Target: infosectrain.com

Read api-keys.yaml from /root/.theHarvester/api-keys.yaml
[*] Searching Yahoo.
    Searching 0 results.
[*] Searching Bing.

[*] No IPs found.

[*] Emails found: 2
-----
sales@infosectrain.com

[*] Hosts found: 1
-----
lms.infosectrain.com

[*] Reporting started.
[*] XML File saved.
[*] JSON File saved.
```

What it does: This command searches for data using multiple sources and provides a more comprehensive result.

Step 4: Saving Results to a File

You can save your results for future reference by using the -f option followed by the filename.

- kali> theharvester -d example.com -b yahoo,bing -f results

```
(root@kali)-[~]
# theHarvester -d infosectrain.com -b yahoo,bing -f result
Read proxies.yaml from /etc/theHarvester/proxies.yaml
*****
*
* [ASCII Art]
*
* theHarvester 4.8.0
* Coded by Christian Martorella
* Edge-Security Research
* cmartorella@edge-security.com
*
*****

[*] Target: infosectrain.com

Read api-keys.yaml from /etc/theHarvester/api-keys.yaml
Searching 0 results.
[*] Searching Bing.
[*] Searching Yahoo.

[*] No IPs found.

[*] Emails found: 2
-----
name@infosectrain.com
sales@infosectrain.com

[*] No people found.

[*] Hosts found: 1
-----
lms.infosectrain.com

[*] Reporting started.
[*] XML File saved.
[*] JSON File saved.
```

What it does: This command outputs the results of the scan to files called **results.xml** and **results.json**. You can use this file for later analysis or report generation.

Step 5: Additional Useful Queries

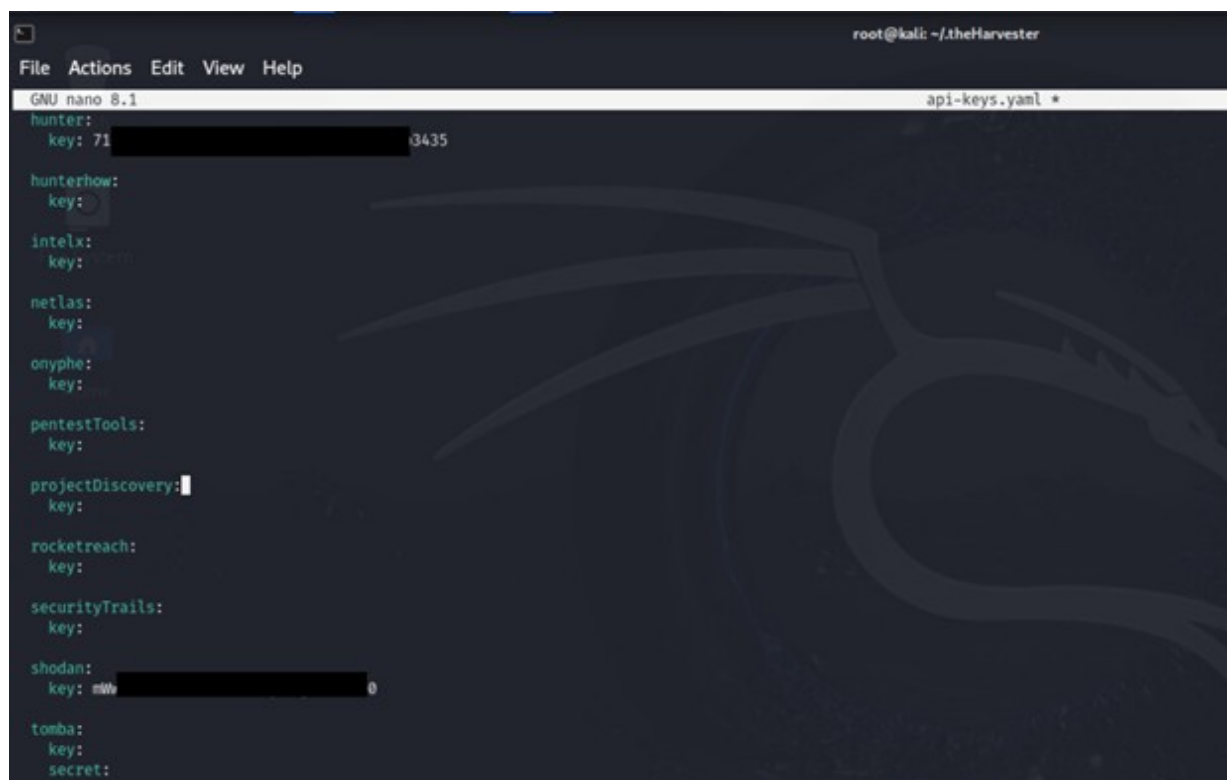
Using APIs for Better Results

If you have API keys for certain services, such as **Hunter.io**, you can improve your results by including them in the query.

1. Configure the API keys

Before using API-based services, you need to configure the keys in

~/theHarvester/api-keys.yaml



```
root@kali: ~/theHarvester
File Actions Edit View Help
GNU nano 8.1 api-keys.yaml *
hunter:
  key: 71[REDACTED]3435
hunterhow:
  key:
intelx:
  key:
netlas:
  key:
onyphe:
  key:
pentestTools:
  key:
projectDiscovery:
  key:
rocketreach:
  key:
securityTrails:
  key:
shodan:
  key: mW[REDACTED]0
tomba:
  key:
  secret:
```

Using Hunter.io to gather emails

-kali> theHarvester -d exxample.com -b hunter

```
(root@kali)-[~/theHarvester]
# theHarvester -d infosectrain.com -b hunter
Read proxies.yaml from /root/.theHarvester/proxies.yaml
*****
*
* [ASCII Art]
*
* theHarvester 4.6.0
* Coded by Christian Martorella
* Edge-Security Research
* cmartorella@edge-security.com
*
*****

[*] Target: infosectrain.com

Read api-keys.yaml from /root/.theHarvester/api-keys.yaml
[*] Searching Hunter.

[*] No IPs found.

[*] Emails found: 10
```

Step 6: Analyzing the Output

After running TheHarvester, you will get a list of results that might look something like this:

//Emails found

For Example

john.doe@example.com

jane.smith@example.com

- List entries

theHarvester can list a lot of entries as part of performing OSINT. You can specify the limit to the number of entries you want to be displayed.

-kali> theHarvester -d owasp.org -b bing -l 10

```
(root@kali)-[~]
# theHarvester -d owasp.org -b bing -l 10
```

The -S flag is used to enable SSL/TLS when making requests, not to set a result limit.

If your goal was to limit the number of results, use the -l option (lowercase L).

- Querying

theHarvester can list a lot of entries as part of performing OSINT. You can specify the limit to the number of entries you want to be displayed.

- kali> theHarvester -d owasp.org -b bing -l 10

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b bing -l 10
```

-shodan

theHarvester also has a option called “-shodan” that queries the [Shodan](#) search engine for any open ports or banners from discovered hosts. However, this requires api keys.

-kali> theHarvester -d owasp.org -b shodan

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b shodan
```

- Search for IP addresses related to owasp.org
- Use your Shodan API
- Show open ports, services, banners, and more

To save the result in HTML:

- kali> theHarvester -d owasp.org -b shodan -f owasp_shodan_report

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b shodan -f owasp_shodan_report
```

To CSV:

- kali> theHarvester -d owasp.org -b shodan -c owasp_shodan.csv

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b shodan -c owasp_shodan.csv
```

Advanced Usage of theHarvester with Shodan

🔑 1. Make Sure the API Key Works

Check if your key is valid by testing directly in terminal:

curl "https://api.shodan.io/api-info?key=YOUR_API_KEY"

You should see a JSON output like:

```
(root@kali)-[~]
# curl "https://api.shodan.io/api-info?key=YOUR_API_KEY"

<html>
<head>
  <title>401 Unauthorized</title>
</head>
<body>
  <h1>401 Unauthorized</h1>
  This server could not verify that you are authorized to access the document you requested. Either you
  supplied the wrong credentials (e.g., bad password), or your browser does not understand how to supply
  the credentials required.<br/><br/>
</body>
</html>
```

2. Run theHarvester with Shodan Backend

-kali> theHarvester -d owasp.org -b shodan

```
(root@kali)-[~]
# theHarvester -d owasp.org -b shodan -f owasp_shodan
```

- kali> theHarvester -d owasp.org -b shodan -f owasp_shodan

```
(root@kali)-[~]
# theHarvester -d owasp.org -b shodan -f owasp_shodan
```

4. Combine Shodan with Other Data Sources

You can run multiple sources one after another like:

- kali> theHarvester -d owasp.org -b bing,shodan,crtsh

```
(root@kali)-[~]
# theHarvester -d owasp.org -b bing,shodan,crtsh
```

- Find subdomains from Bing
- IPs and banners from Shodan
- SSL certificates from crt.sh

5. Smart Options

Option	Description
-l 50	Limit results (for Bing, DuckDuckGo, etc.)
-f <name>	Save report as HTML/XML
-c <file.csv>	Save results as CSV
-r	Try to resolve hostnames to IPs
-n	Disable DNS resolution
-S	Force HTTPS
-t	Take screenshots (with Selenium)

Full combo:

- kali> theHarvester -d owasp.org -b bing,shodan,crtsh -l 100 -f owasp_report -r -S

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b bing,shodan,crtsh -l 100 -f owasp_report -r -S
```

6. Troubleshooting Tips

Problem	Solution
✗ No results from Shodan	Check API key in /etc/theHarvester/api-keys.yaml
✗ Invalid API	Make sure the key is not expired or used up
✗ Command hangs or slow	Shodan may rate limit; use -l to limit data
✗ Permission denied	Run with sudo if access error

7. Output Location

After you run:

- kali> theHarvester -d owasp.org -b shodan -f owasp_shodan

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b shodan -f owasp_shodan
```

-screenshot

This command allows theHarvester to take screenshots of subdomains that are found.

```
(root@kali)-[~] * CSV output (if you use -o owasp.csv)
# theHarvester -d owasp.org --screenshot /home/kali/desktop
```

1. Basic Search with Google and Screenshots

- kali> theHarvester -d owasp.org -b google -l 50 --screenshot /home/kali/Desktop/screenshots

```
(root@kali)-[~]
# theHarvester -d owasp.org -b google -l 50 --screenshot /home/kali/Desktop/screenshots
```

2. Multi-Source Search with Screenshots

- kali> theHarvester -d owasp.org -b google,bing,dnsdumpster -l 100 --screenshot /home/kali/Desktop/screenshots

```
(root@kali)-[~]
# theHarvester -d owasp.org -b google,bing,dnsdumpster -l 100 --screenshot /home/kali/Desktop/screenshots
```

3. DNS Brute-Forcing with Screenshots

- kali> theHarvester -d owasp.org -b dnsdumpster -c --screenshot /home/kali/Desktop/screenshots

```
(root@kali)-[~]
# theHarvester -d owasp.org -b dnsdumpster -c --screenshot /home/kali/Desktop/screenshots
```

4. Shodan Integration with Screenshots

```
(root@kali)-[~]
# theHarvester -d owasp.org -b shodan -h --screenshot /home/kali/Desktop/screenshots
```

5. All Sources with HTML Output and Screenshots

- kali> theHarvester -d owasp.org -b all -l 200 -f /home/kali/Desktop/results.html --screenshot /home/kali/Desktop/screenshots


```

(root@kali)~[~]
# theHarvester -d owasp.org -b all -l 200 -f /home/kali/Desktop/results.html --screenshot /home/kali/Desktop/
screenshots
Read proxies.yaml from /etc/theHarvester/proxies.yaml
*****
*                               *
* [ASCII ART]                   *
*                               *
* theHarvester 4.8.0            *
* Coded by Christian Martorella *
* Edge-Security Research        *
* cmartorella@edge-security.com *
*                               *
*****

[*] Target: owasp.org

Read api-keys.yaml from /etc/theHarvester/api-keys.yaml

[!] Missing API key for bevigil.
Read api-keys.yaml from /etc/theHarvester/api-keys.yaml
Read api-keys.yaml from /etc/theHarvester/api-keys.yaml
Read api-keys.yaml from /etc/theHarvester/api-keys.yaml

```

-dns-brute

As the command explains, you can brute force DNS servers using this option.

theHarvester Commands with DNS Brute-Forcing (-c)

Below are corrected **theHarvester** commands using -c for DNS brute-forcing, targeting owasp.org. Since --dns-brute is invalid in the standard version, I'll use -c. If you're using a custom fork where --dns-brute is required, let me know, and I'll tailor further. Each command includes:

- **Command:** The exact command.
- **Purpose:** What it does.
- **Expected Output:** Sample terminal output based on typical results for owasp.org.
- **Notes:** Tips or potential issues.

1. DNS Brute-Forcing with DNSDumpster

- kali> theHarvester -d owasp.org -b dnsdumpster -c

```

(root@kali)~[~]
# theHarvester -d owasp.org -b dnsdumpster -c

```

2. DNS Brute-Forcing with Multiple Sources

- kali> theHarvester -d owasp.org -b google,bing,dnsdumpster -c -l 100

```

(root@kali)~[~]
# theHarvester -d owasp.org -b google,bing,dnsdumpster -c -l 100

```

3. DNS Brute-Forcing with Shodan

- kali> theHarvester -d owasp.org -b shodan -c -h

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b shodan -c -h
```

4. DNS Brute-Forcing with All Sources and HTML Output

- kali> theHarvester -d owasp.org -b all -c -l 200 -f /home/kali/Desktop/results.html

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b all -c -l 200 -f /home/kali/Desktop/results.html
```

-source

theHarvester uses many public sources to collect information. Some of them are anubis, baidu, bing, brave, censys, etc. We can even ask it to use a specific source using the “-source” command.

theHarvester Commands with DNS Brute-Forcing (-c)

Below are corrected **theHarvester** commands using -c for DNS brute-forcing, including Censys and other sources, targeting owasp.org. I’ll use -b (standard syntax) instead of --source, but note that your command used --source censys, which is equivalent to -b censys in **theHarvester 4.8.0**. No --screenshot is included, as per your request. Each command includes:

- **Command:** The exact command.
- **Purpose:** What it does.
- **Expected Output:** Sample terminal output based on typical results for owasp.org.
- **Notes:** Tips or potential issues.

1. DNS Brute-Forcing with Censys

- kali> theHarvester -d owasp.org -b censys -c

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b censys -c
```

2. DNS Brute-Forcing with DNSDumpster

- kali> theHarvester -d owasp.org -b dnsdumpster -c

```
(root@kali)-[~]  
# theHarvester -d owasp.org -b dnsdumpster -c
```

Reference

1. hackercool: <https://www.hackercoolmagazine.com/beginners-guide-to-theharvester-tool/>
2. infosectrain: <https://www.infosectrain.com/blog/step-by-step-guide-for-theharvester-tool/>

Contact:

Gmail: ninkanong200620@gmail.com