ETHICAL HACKING

INFORMATION GATHERING

FOOTPRINTING

Spiderfoot

Researcher by: Nin Kanong

SpiderFoot – A Automate OSINT Framework

Spiderfoot is a python script and can be run on any machine with Python installed. Using spiderfoot, we can gather information from almost any open source data source available. For this tutorial, we will be using Kali Linux as spiderfoot is installed by default on it. Spiderfoot has an embedded web server and hence has a web-based interface.

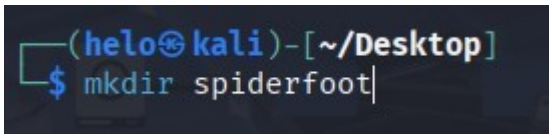
Installation Spiderfoot Framework :

1. Open your Kali Linux operating system. Move to the desktop using the following command. You have to move to Desktop because on desktop you have to create a directory into which you have to clone the tool. Use the following command to move to Desktop.

- kali> cd Desktop

2. Now you are on the desktop. Here you have to create a new directory called spiderfoot. In this directory, you have to clone the tool from Github. Use the following command to create a new directory.

- kali> mkdir spiderfoot

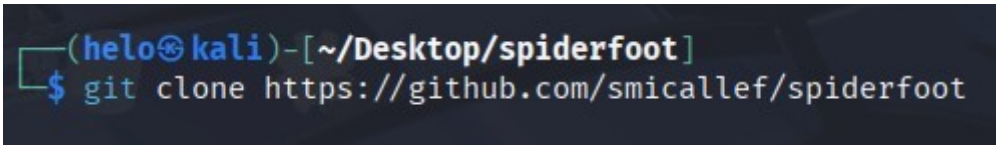


3. Now use the following command to move in the directory that you have created.

- kali> cd spiderfoot

4. Now you are in spiderfoot directory. In this directory, you have to clone the tool from GitHub.
Use the following command to clone the tool from GitHub.

- kali> git clone https://github.com/smicallef/spiderfoot



> One more

Clone the Repository:

- kali> git clone https://github.com/smicallef/spiderfoot.git

- kali> cd spiderfoot

Alternatively, download the stable version:

- kali> wget https://github.com/smicallef/spiderfoot/archive/v4.0.tar.gz

- kali> tar zxvf v4.0.tar.gz

- kali> mv spiderfoot-4.0 spiderfoot

- kali> cd spiderfoot

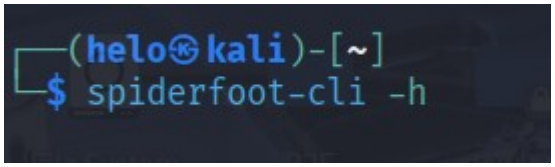
Install Dependencies:

- kali> pip3 install -r requirements.txt

Spiderfoot CML

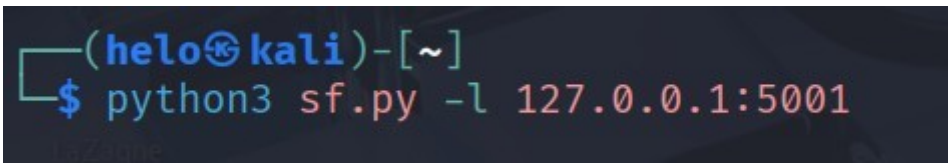
- kali> spiderfoot -h

- kali> spiderfoot-cli -h



Start the Web Server:

- kali> python3 sf.py -l 127.0.0.1:5001

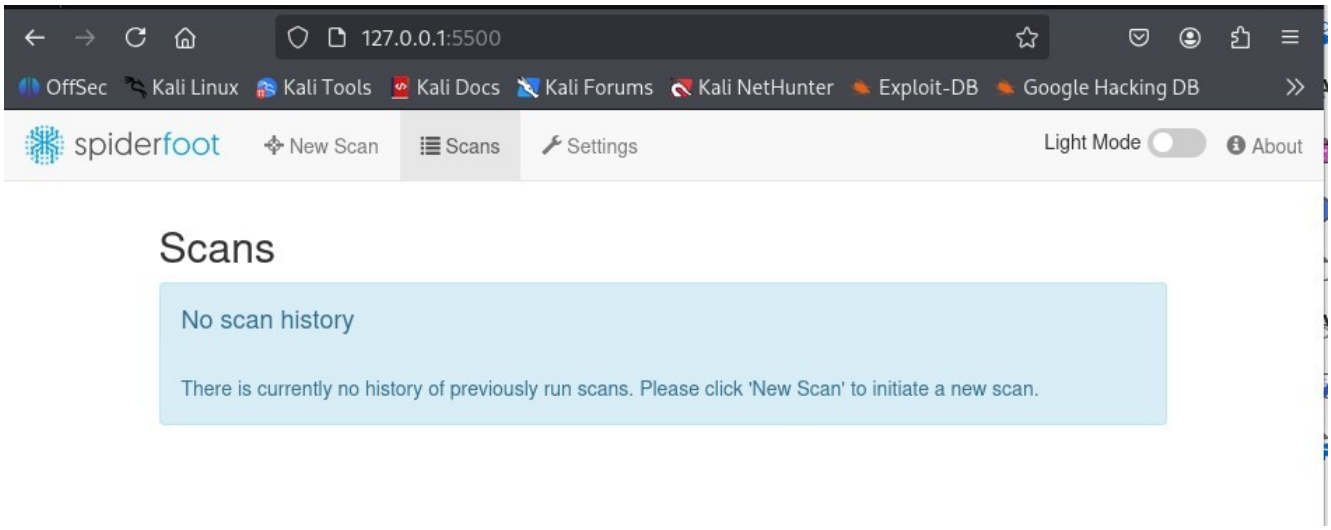


To start spiderfoot on Kali, all you have to do is use the “-l” option and then specify a IP address and port on which you want the web server to listen on. The “-l” option stands for listen. Here we have configured spiderfoot to listen on the port 5500 of localhost.

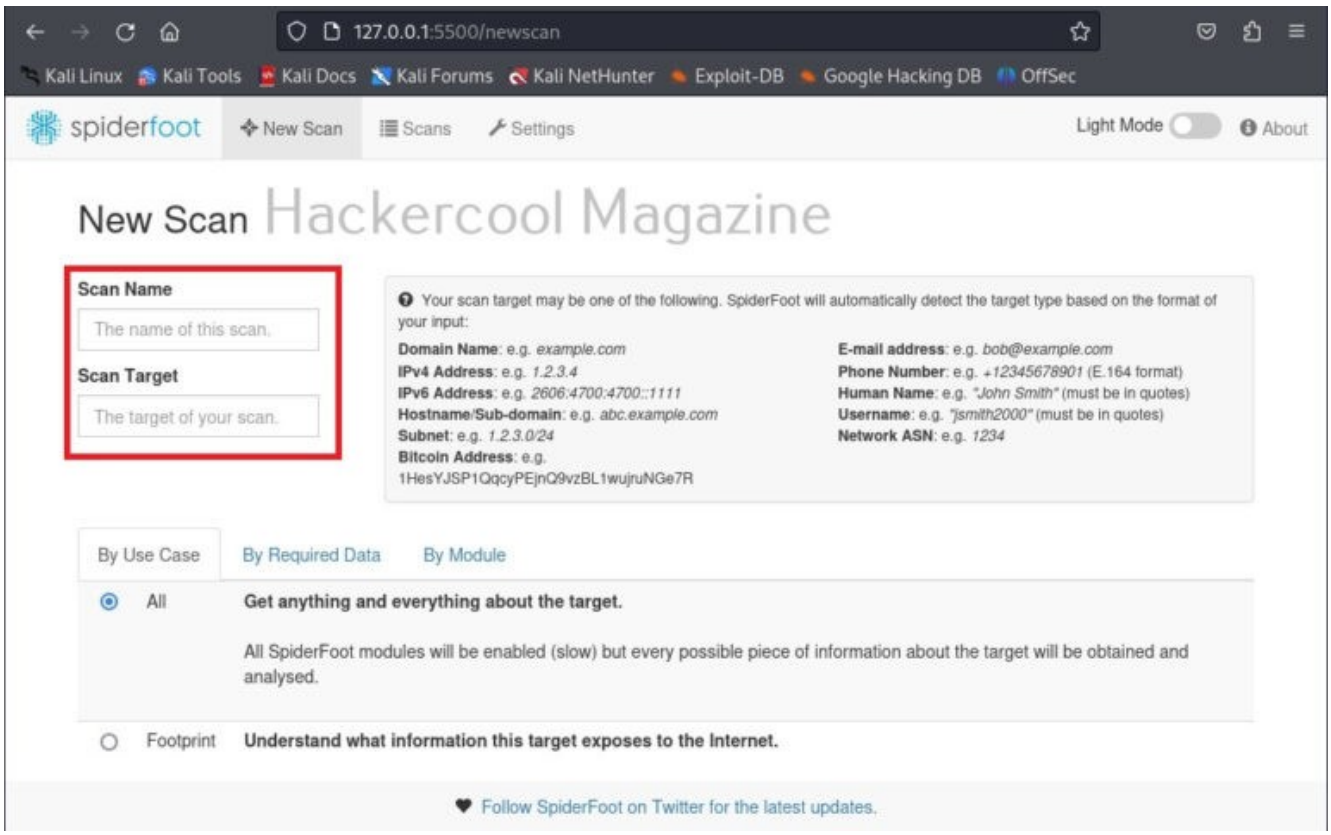
- kali> spiderfoot -l 127.0.0.1:5500

```
┌──(helo㉿kali)-[~]
└─$ sudo spiderfoot -l 127.0.0.1:5500

*************************************************************
 Use SpiderFoot by starting your web browser of choice and
 browse to http://127.0.0.1:5500/
*************************************************************

2025-07-05 11:22:56,162 [INFO] sf : Starting web server at 127.0.0.1:5500 ..
.
2025-07-05 11:22:56,168 [WARNING] sf :
*********************************************************************
Warning: passwd file contains no passwords. Authentication disabled.
Please consider adding authentication to protect this instance!
Refer to https://www.spiderfoot.net/documentation/#security.
*********************************************************************
```

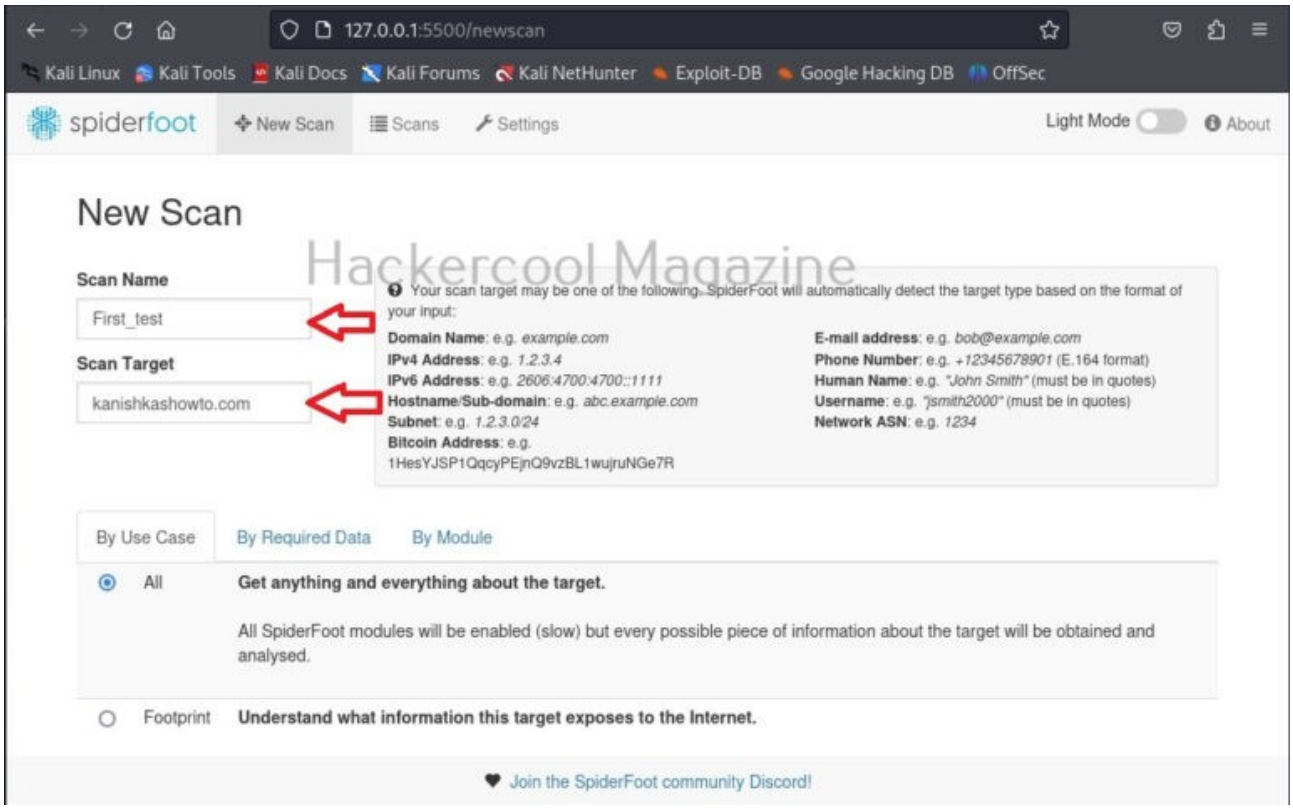
Now, browse to the above highlighted URL using your favorite browser. You should see this.



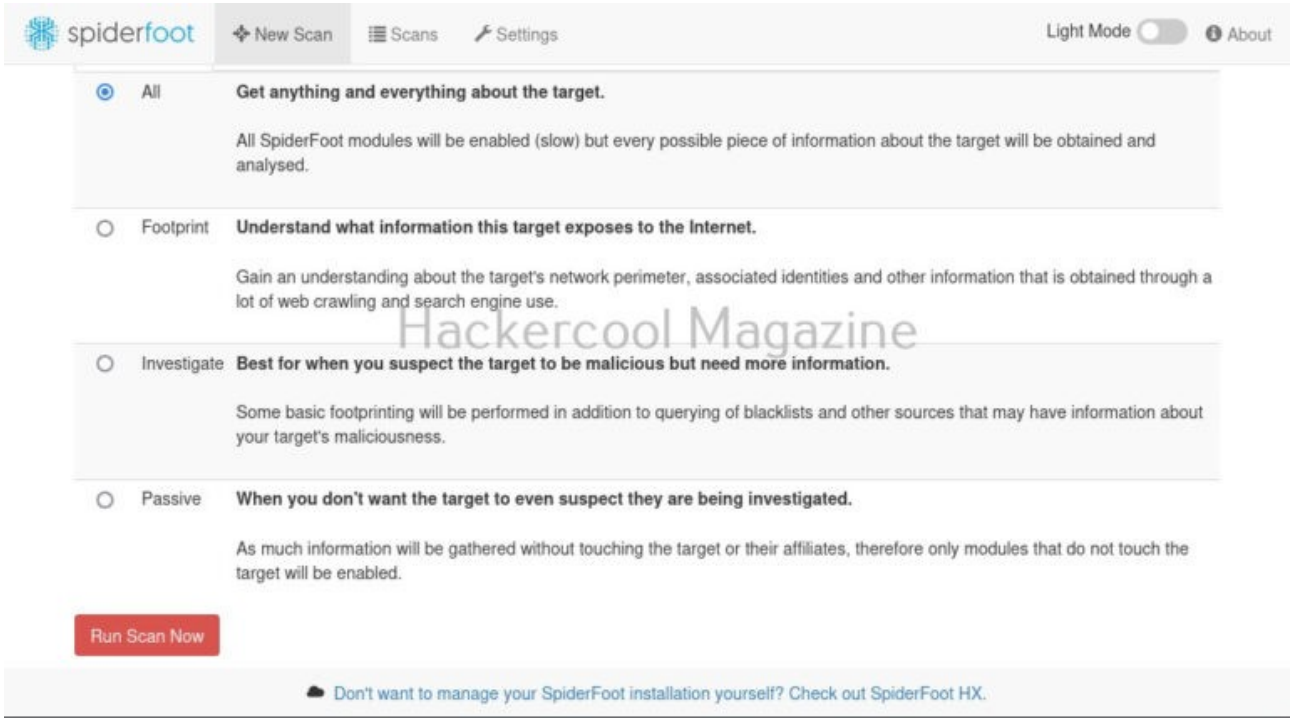
Since we have not yet performed any scans yet. There is no scan history. To start a new scan, click on “New scan”.



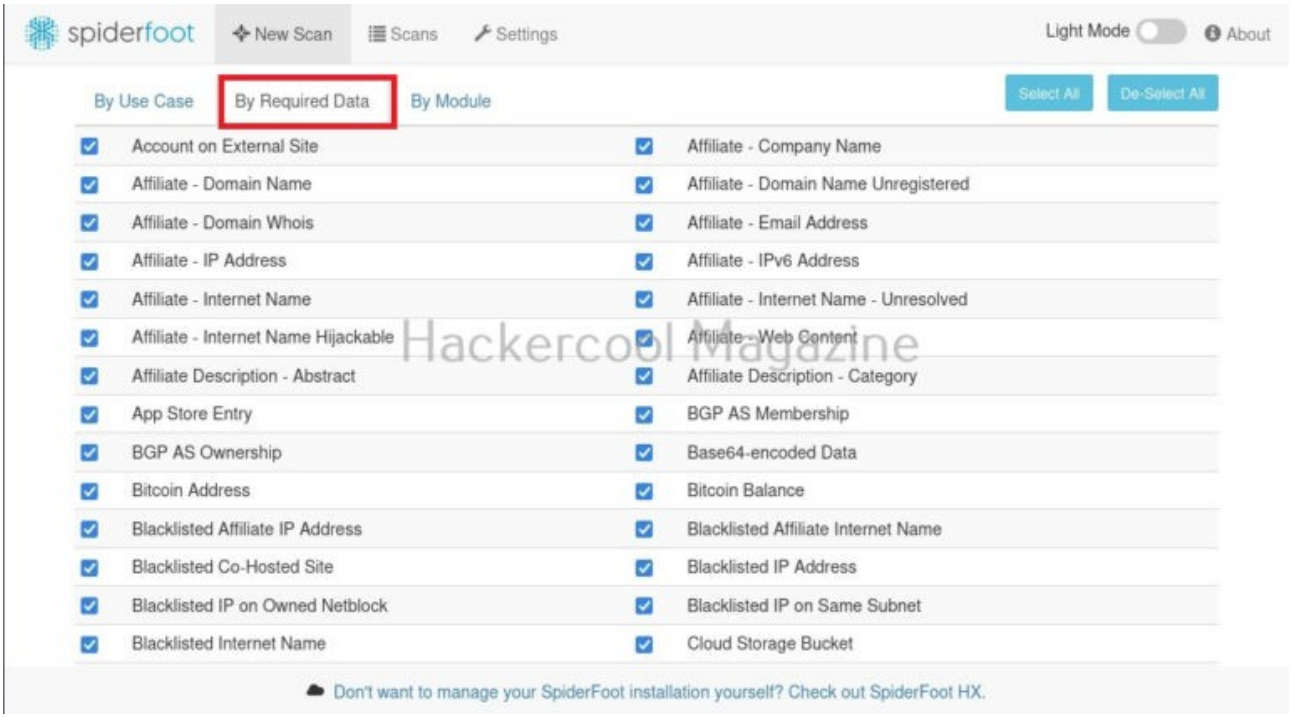
Spiderfoot can gather information from domain name, IPV4 or IPV6 address, host names, sub-domains, subnet, Bitcoin address, E-mail address, phone number, human names, usernames and networks. Let’s start our search with a domain name first.



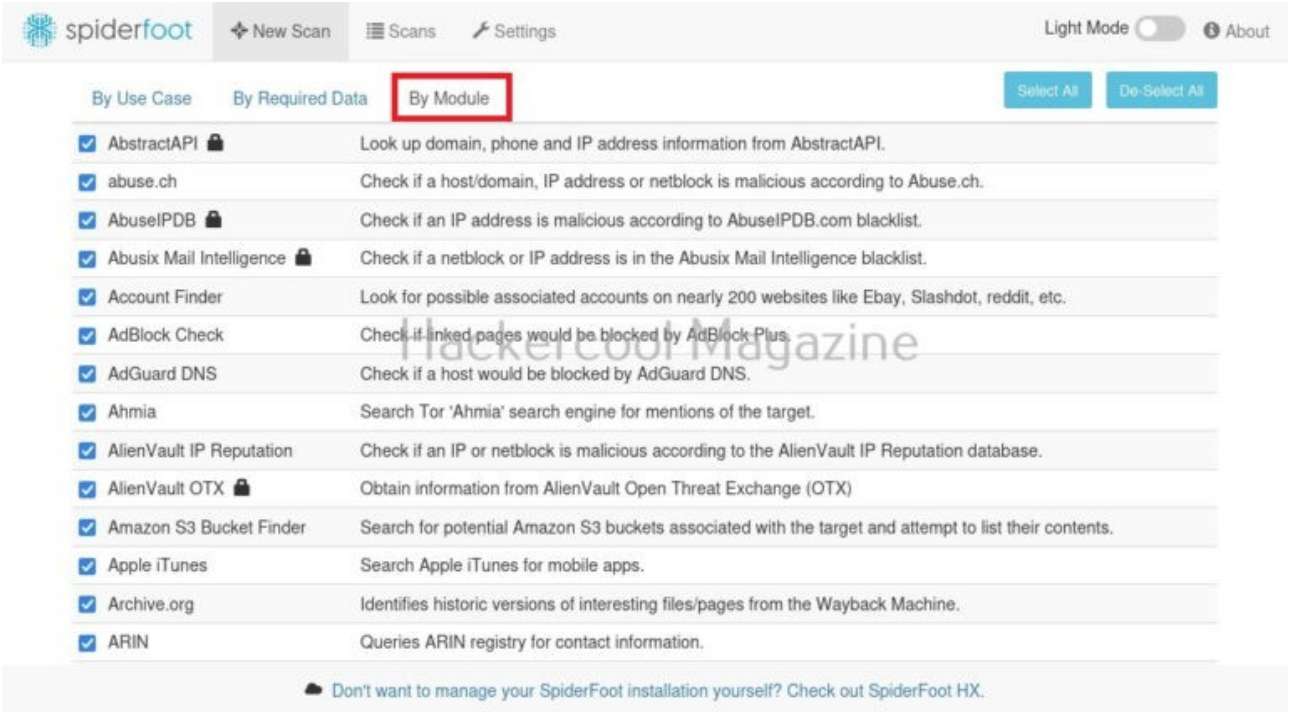
After entering the name of the scan and the scan target scroll down a bit.



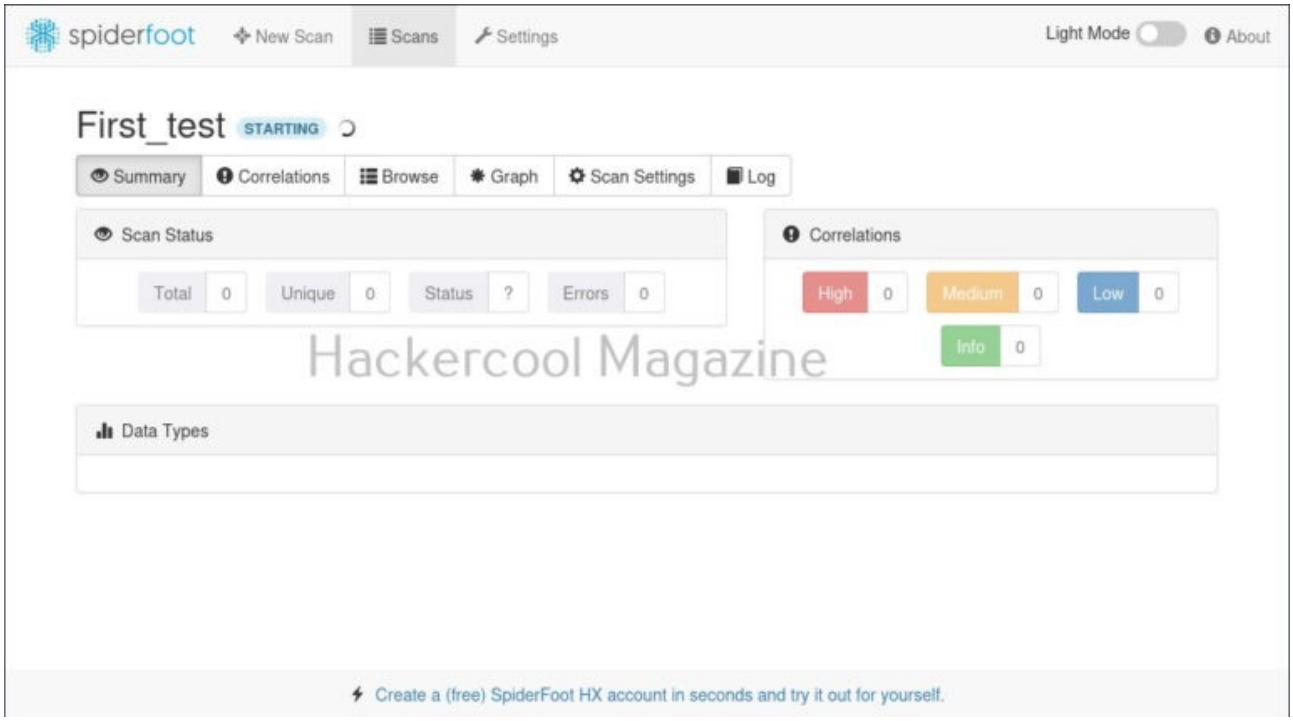
There are various ways you can search with for any target using SpiderFoot. You can also search based on what you require about the target.



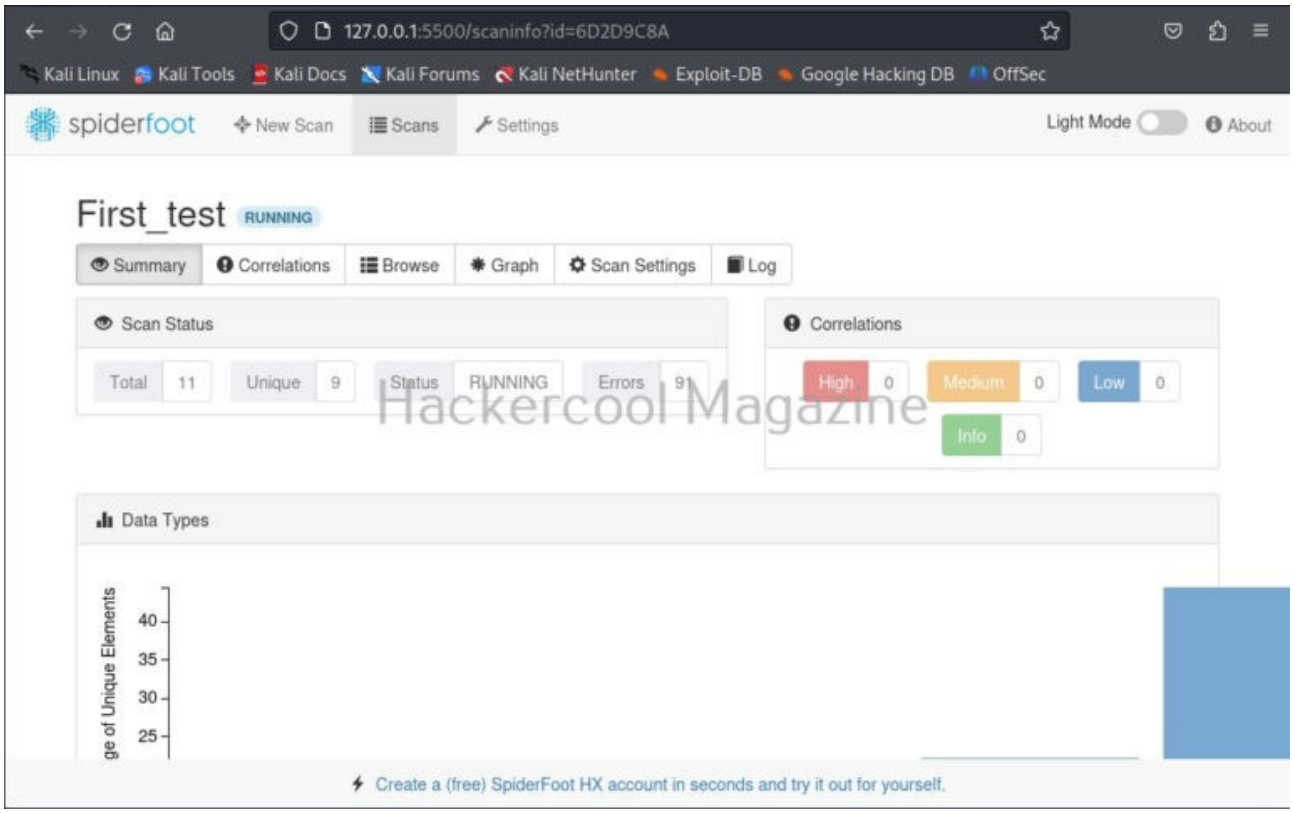
You can also search based on required module (more about modules later).

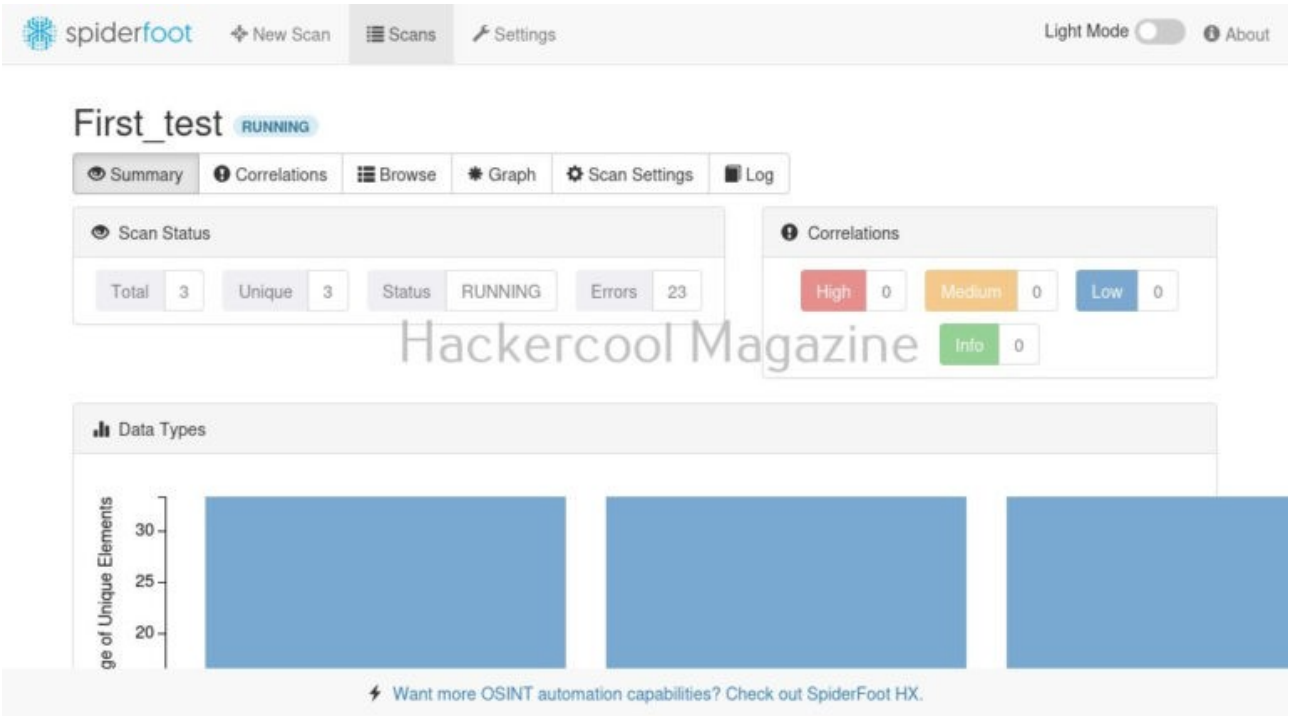


I select “All” and click on “Run scan now”. The scan starts and may look empty at the beginning.

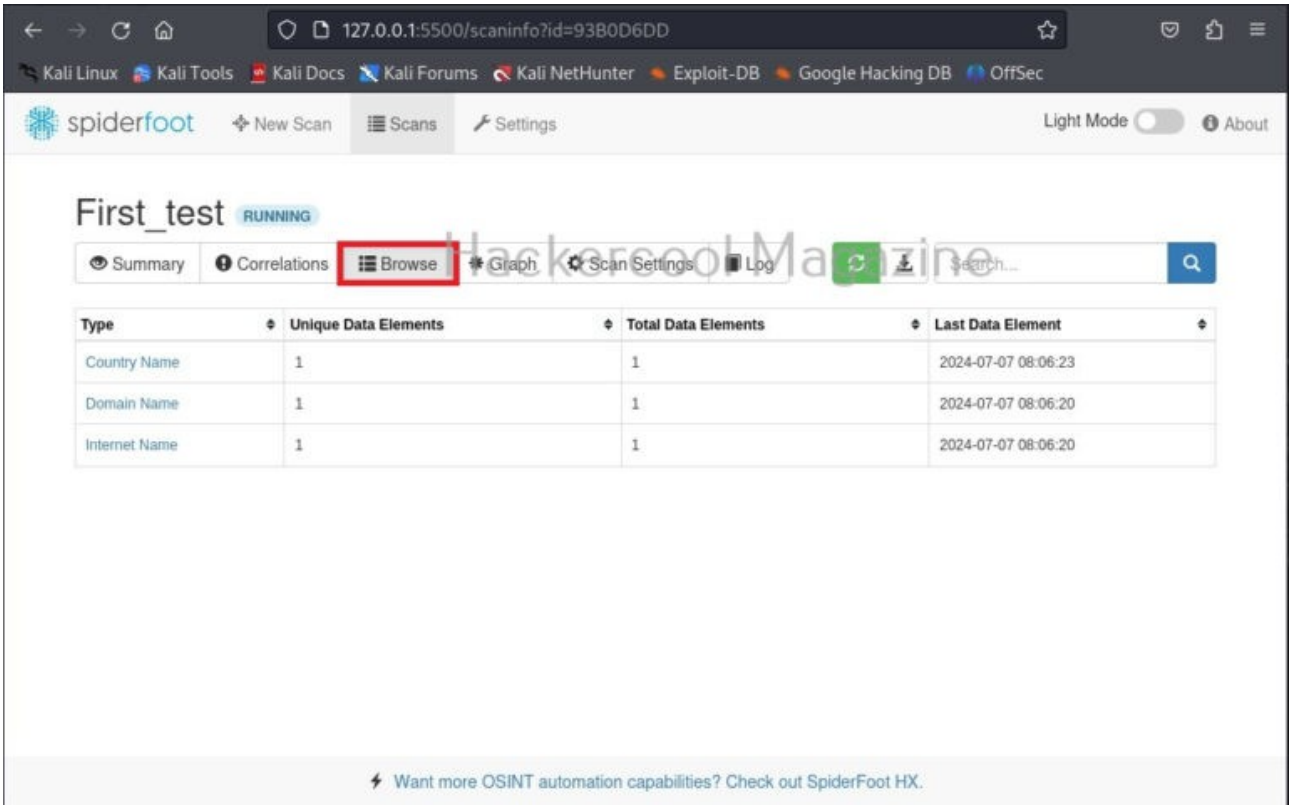


As the scan progresses, your screen will be filled with bars as shown below.

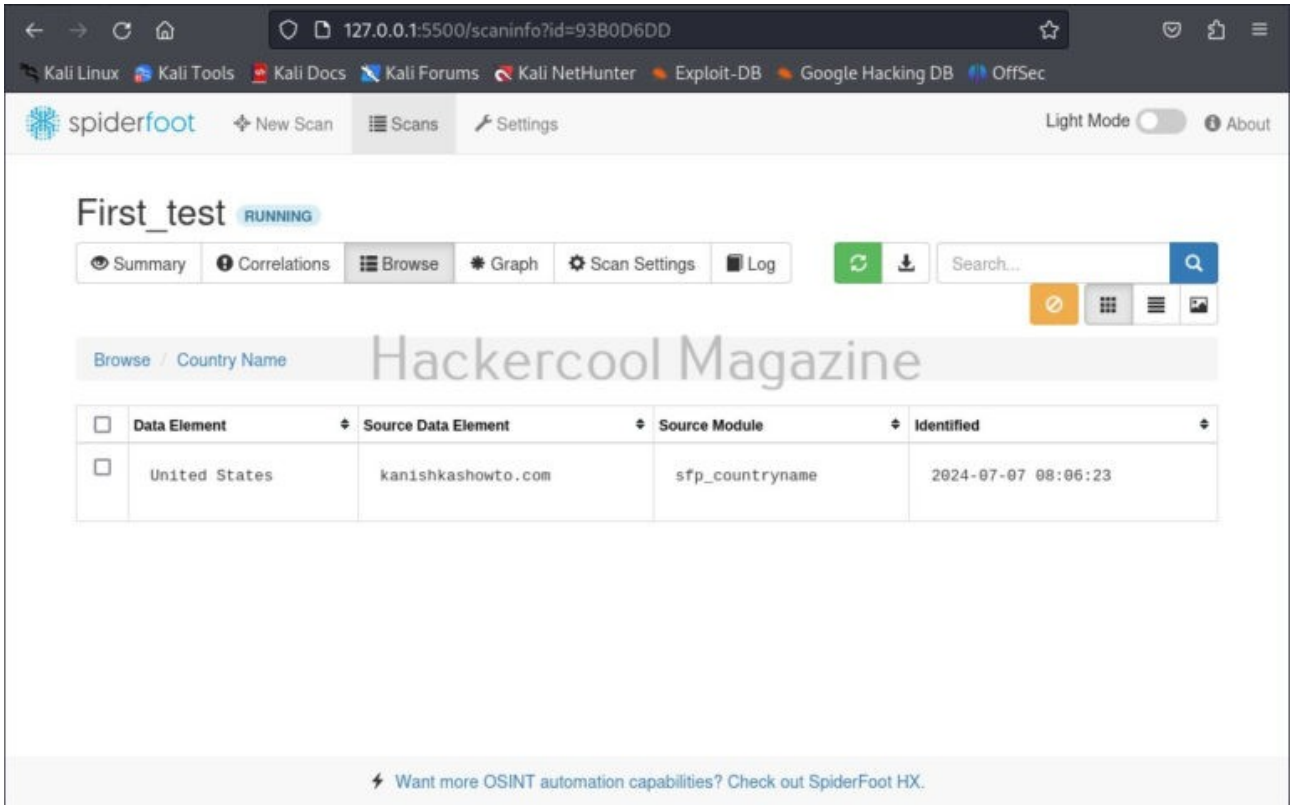




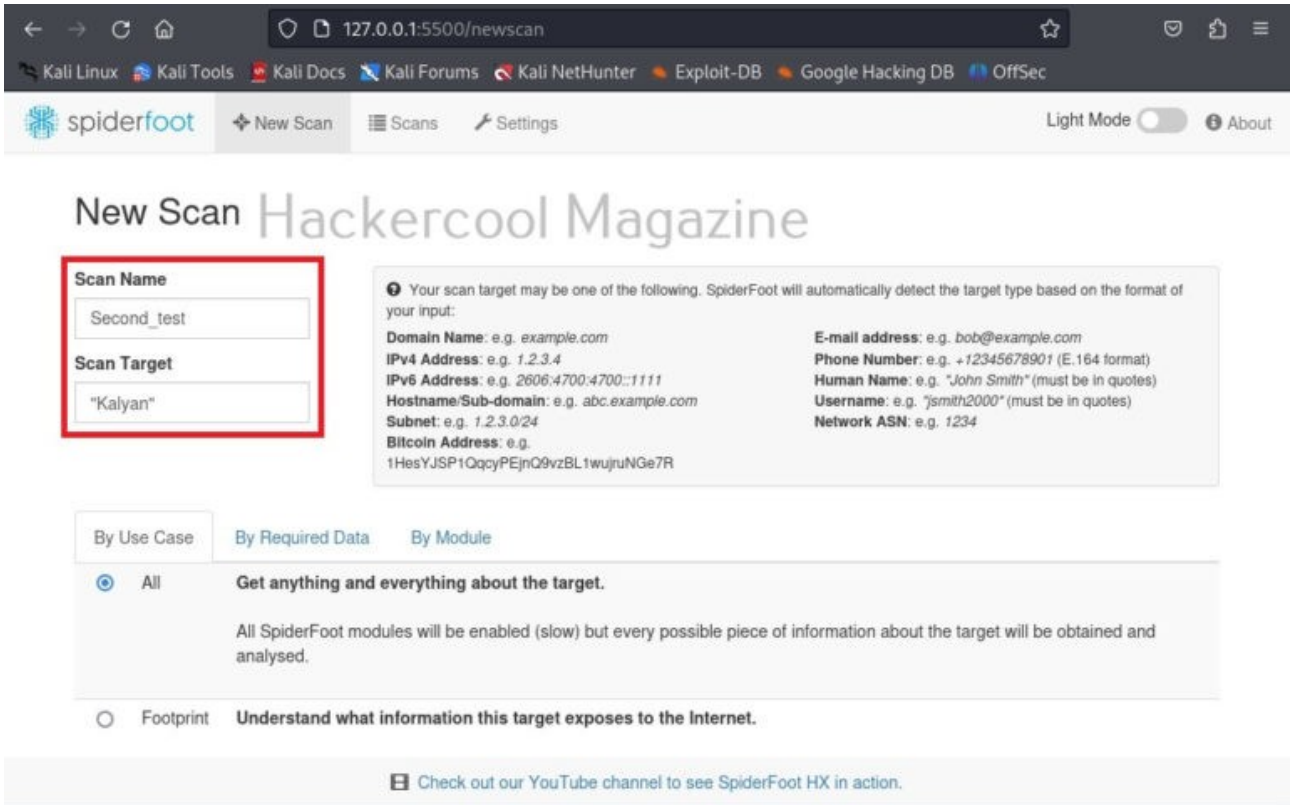
While the scan is still running, you can view the findings of the scan by going to the “Browse” tab as shown below.

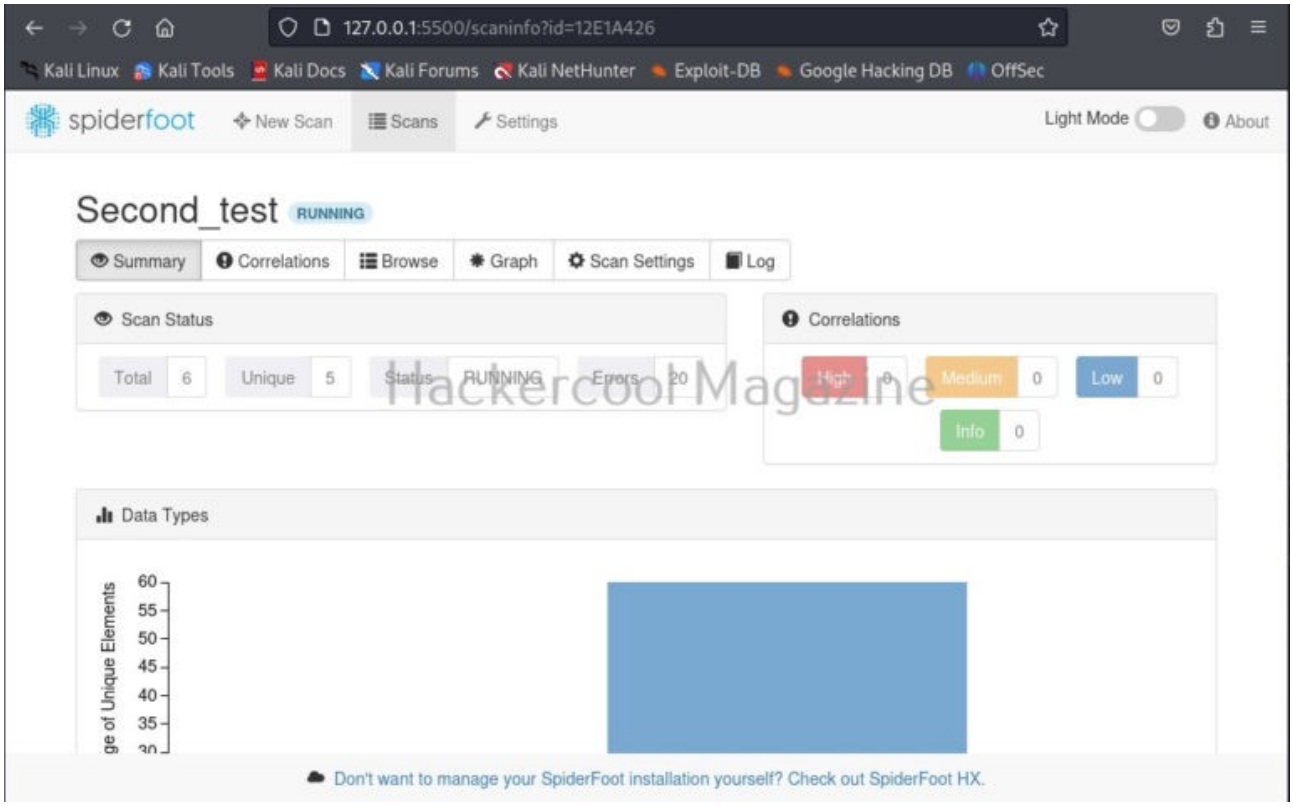


You can view each of the entries to find out what spiderfoot has detected.

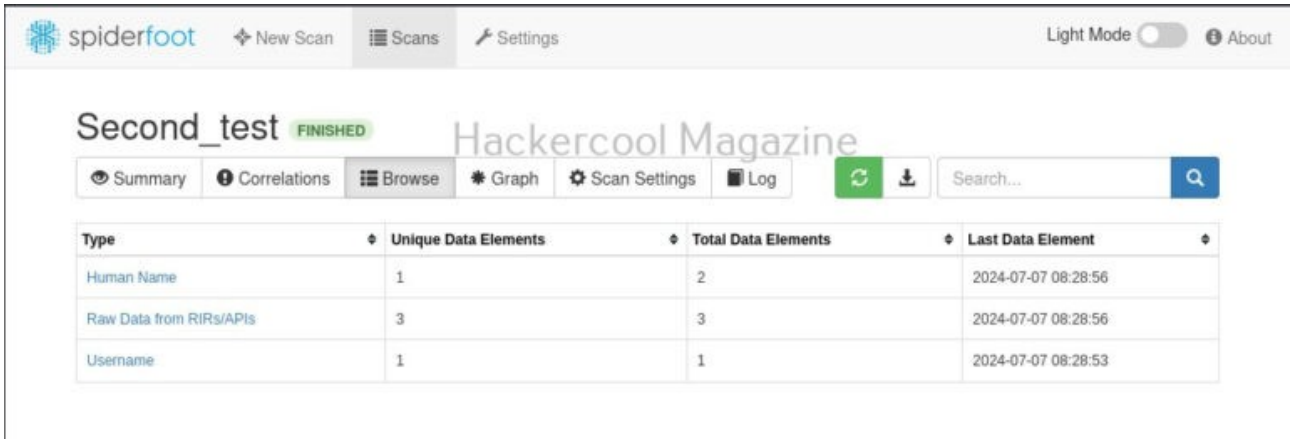


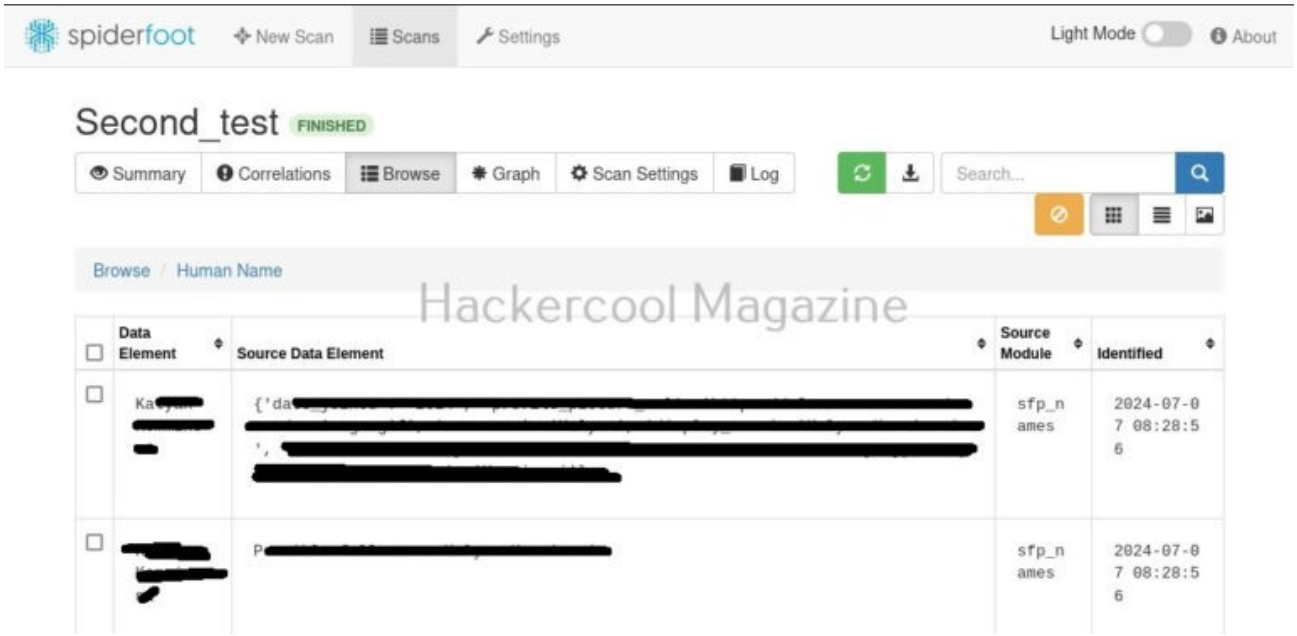
For example, in this case, the target website is hosted in USA. Now, let's search for a "Name" say "kalyan". The good thing about spiderfoot is that it will automatically detect the type of target based on format of your input.



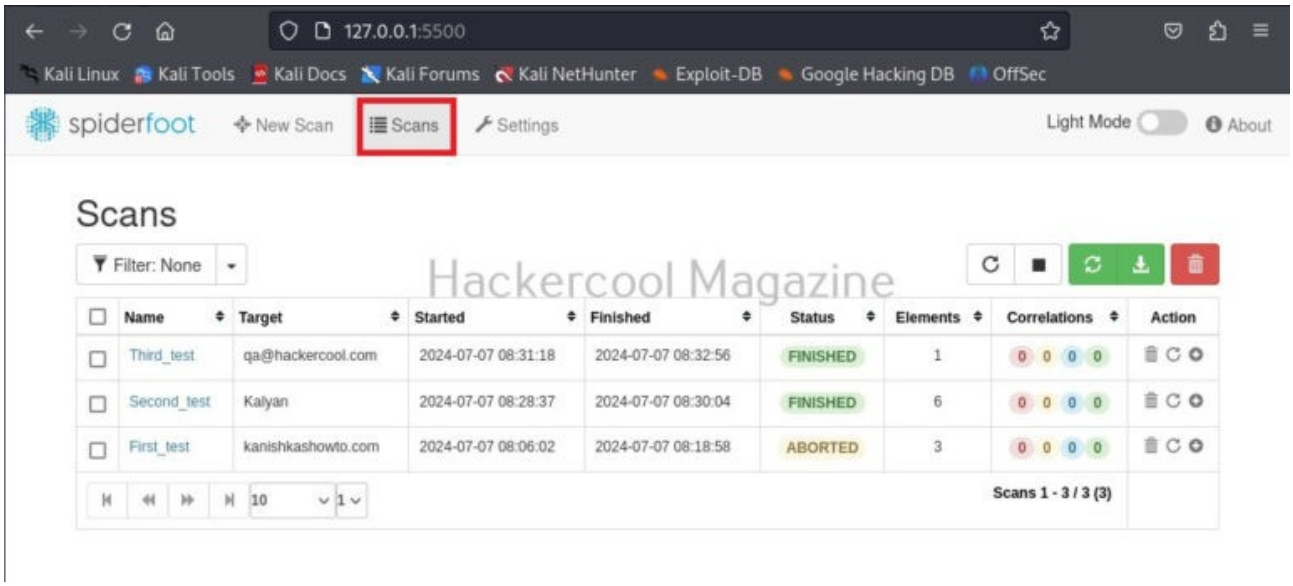


Here’s the result.

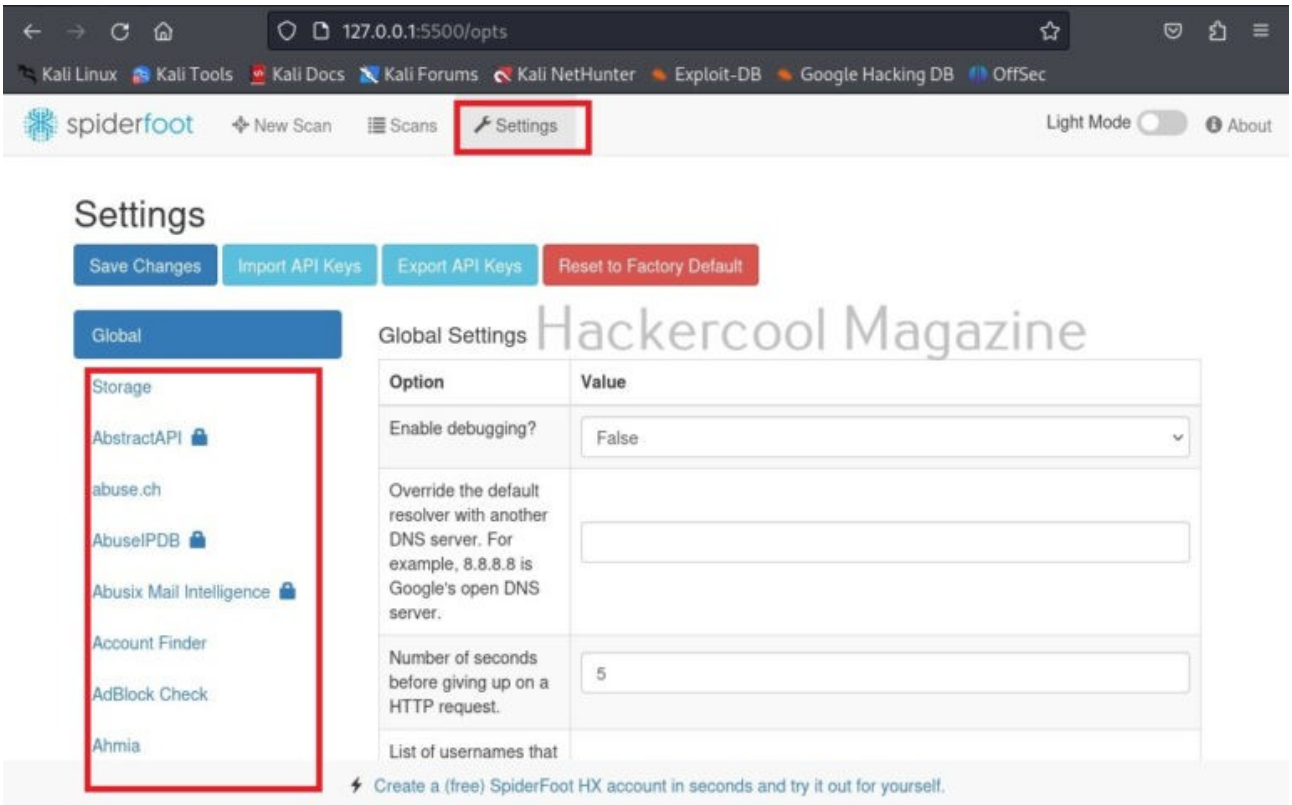


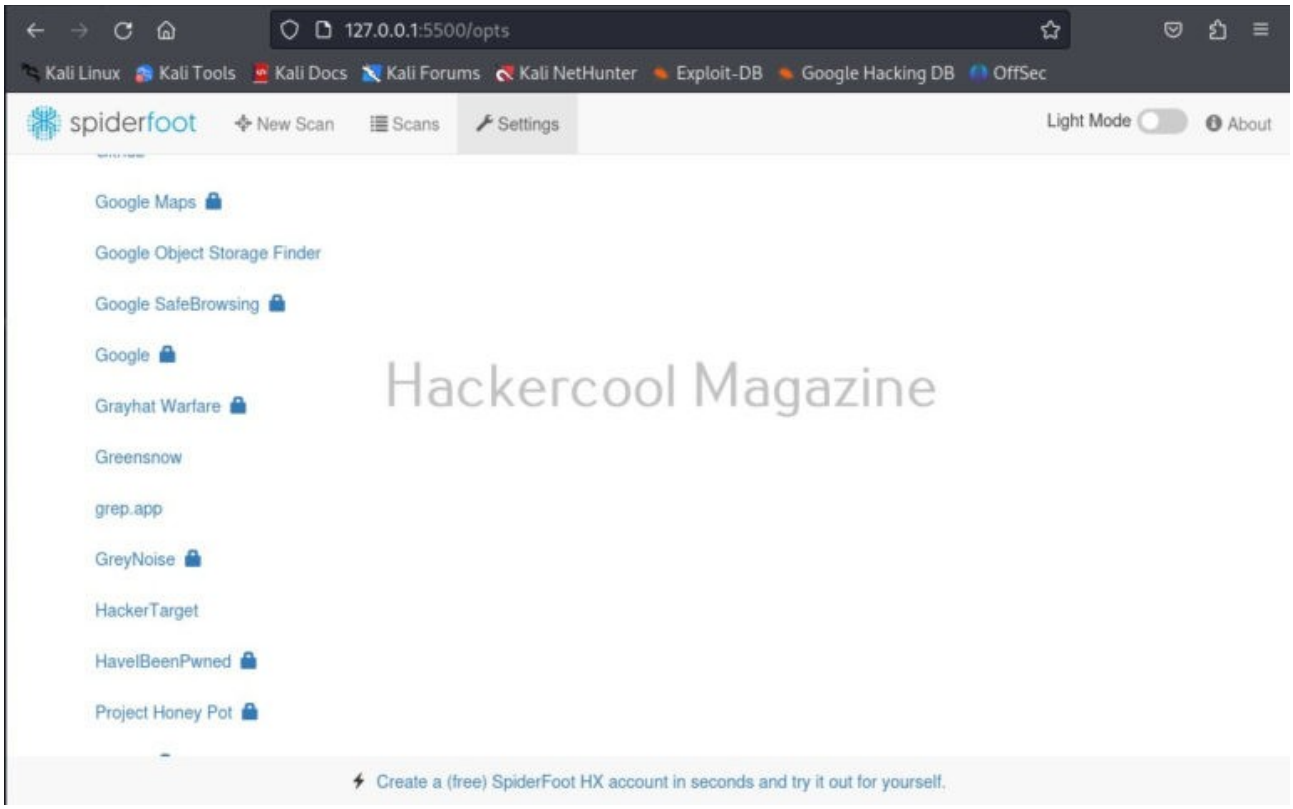


You can see all the scans you performed in the “scans” section.

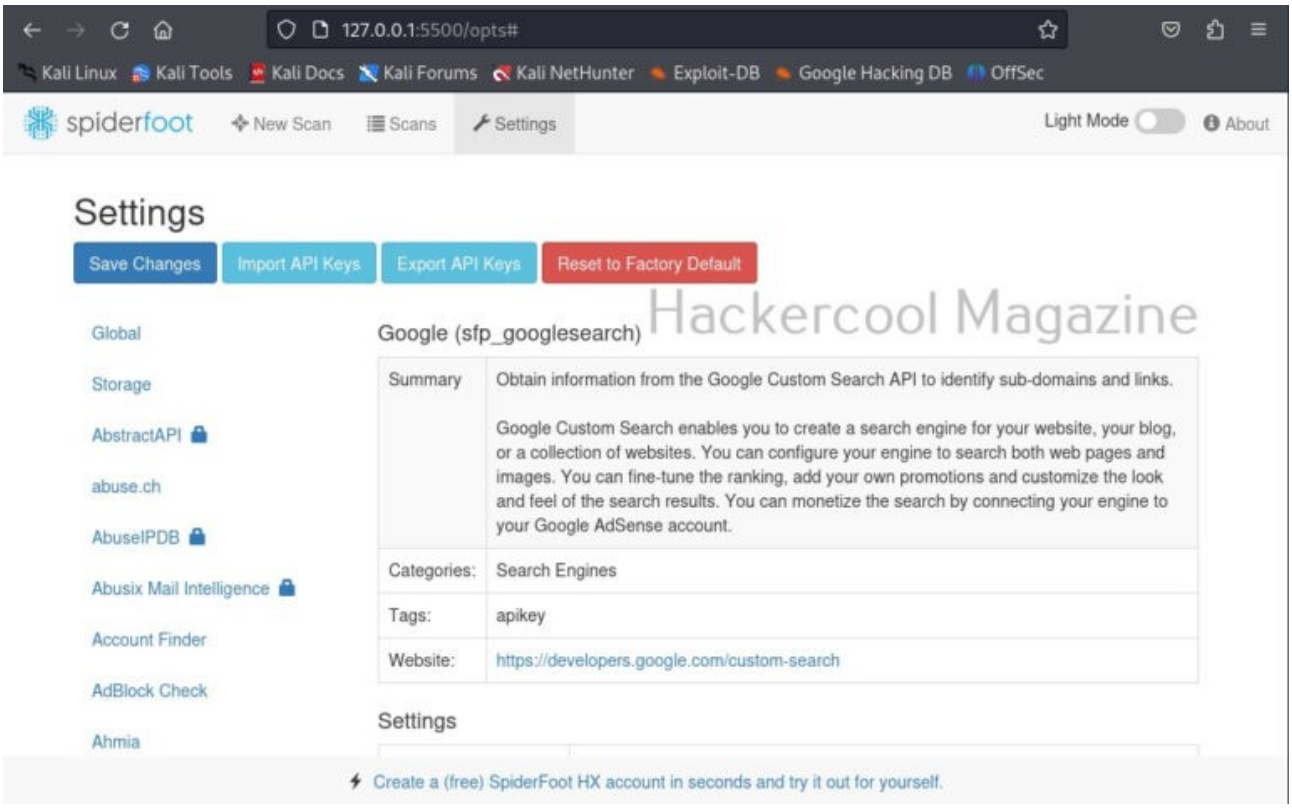


Another important tab here is the “settings” tab. It consists of settings for this tool. But just not that. Remember, I told you at the beginning of this article that Spiderfoot can collect information from almost all data sources. These data sources are listed here to the left in settings section.





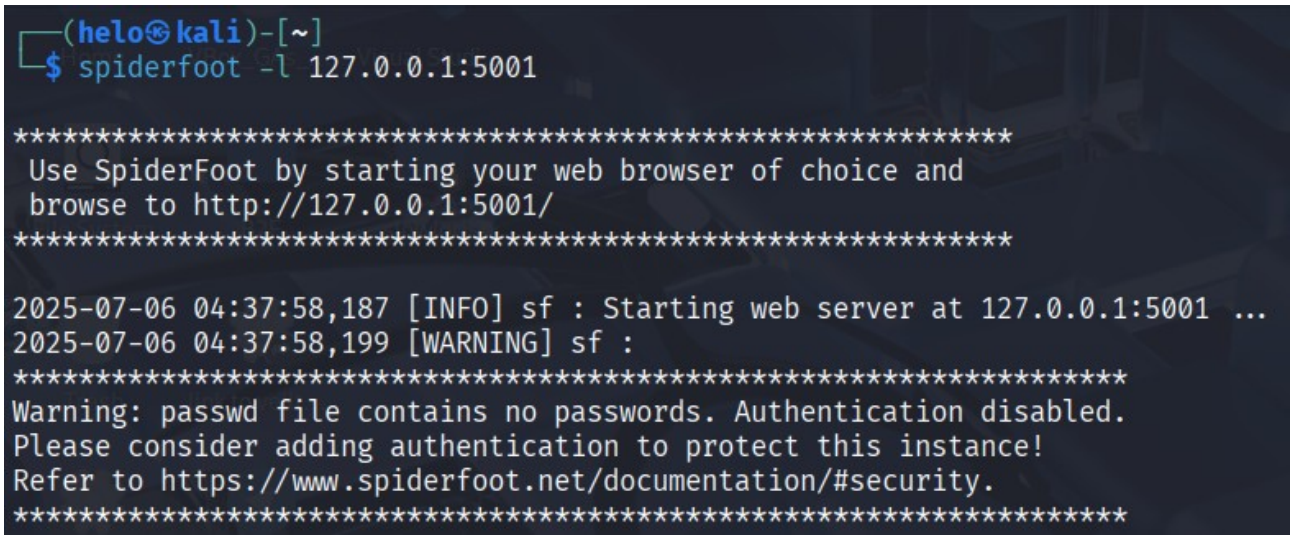
Almost all sources are free, but some need APIs belonging to that particular service (Did you see the lock sign next to some services?).



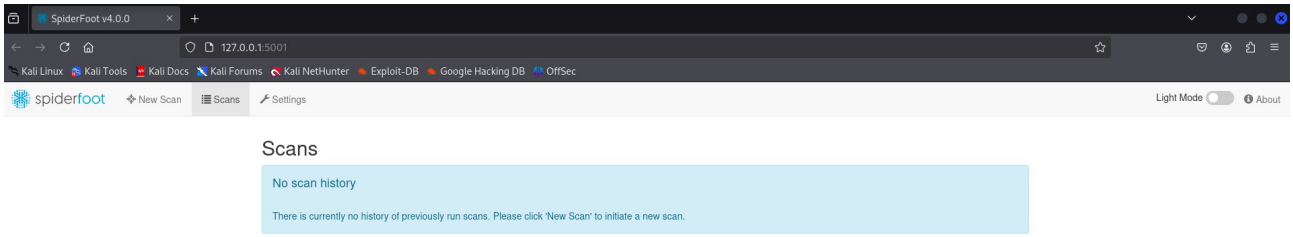
Setup the target by host 5001
Step 1: Setting up a Target

To start Spiderfoot's web interface, run:

- kali> spiderfoot -l 127.0.01:5001



By default, the web interface runs on **http://127.0.0.1:5001**. Open it in a browser to access the dashboard.



Step 2: Setting Up a Target

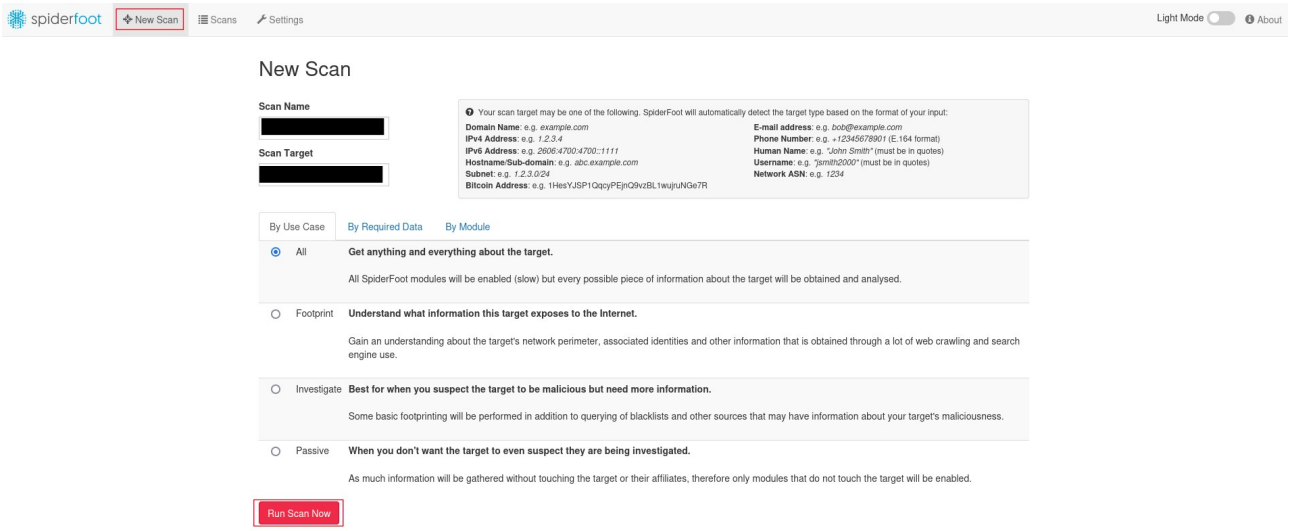
After launching Spiderfoot's web UI:

1. Navigate to **New Scan**.
2. Enter the target domain, IP, or organization name.

3. Choose a **scan profile**

- **All:** Runs all modules to gather every possible detail (slow but thorough).
- **Footprint:** Maps public-facing info about the target's network and identity.
- **Investigate:** Checks for malicious indicators along with basic footprinting.
- **Passive:** Gathers intel without directly interacting with the target.

4. Click **Run Scan** to initiate information gathering.



Step 3: Selecting Modules

Spiderfoot offers over 200 modules categorized under different data types. Key modules include:

1. Network and Infrastructure Intelligence

- **sfp_dnsresolve:** Resolves domain names to IP addresses.
- **sfp_whois:** Retrieves WHOIS registration data.
- **sfp_ports:** Identifies open ports.

2. Social Media & Public Profiles

- **sfp_twitter:** Extracts Twitter mentions.
- **sfp_facebook:** Finds Facebook pages related to the target.

3. Dark Web & Data Breaches

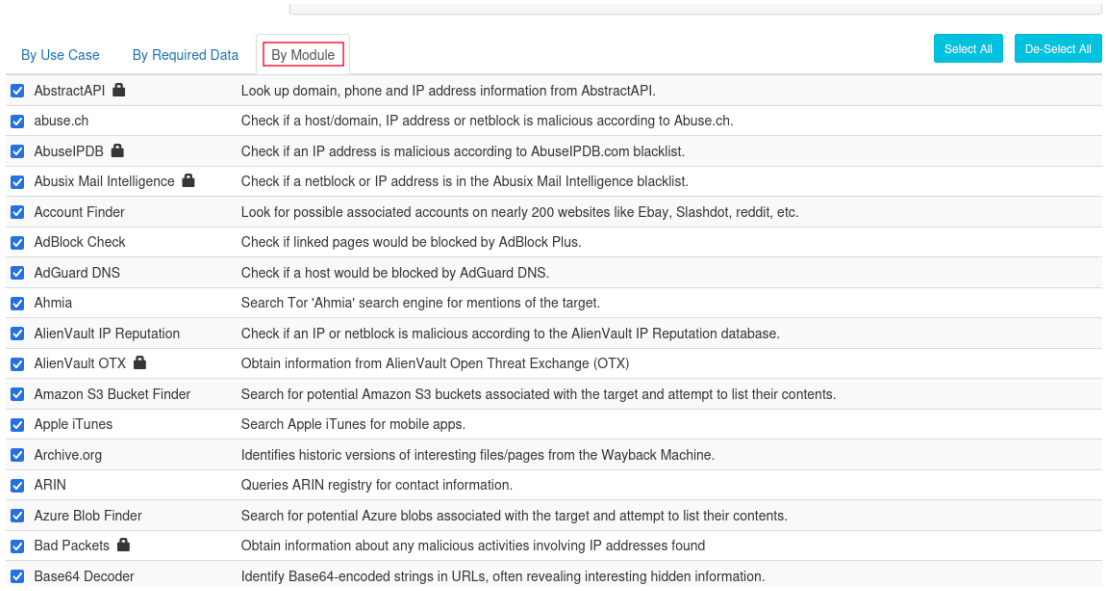
- **sfp_pwned:** Checks if the target appears in data breaches.
- **sfp_darkweb:** Searches for references to the target in dark web marketplaces and forums.

4. Email and Contact Information

- **sfp_email:** Finds email addresses linked to the domain.
- **sfp_phone:** Extracts phone numbers if available.

5. Threat Intelligence

- **sfp_virustotal:** Checks if the target appears in VirusTotal reports.
- **sfp_shodan:** Gathers exposed services and vulnerabilities from Shodan.

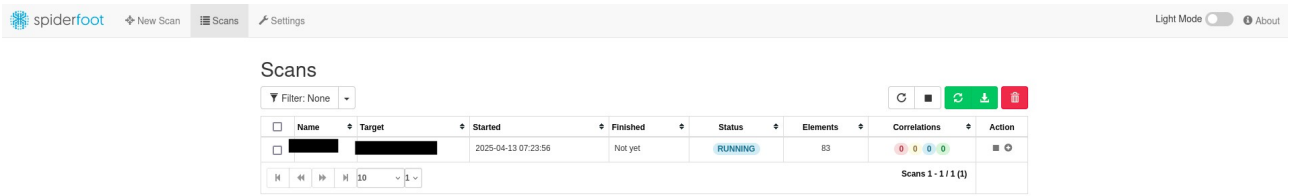


You can enable or disable modules based on your reconnaissance scope.

Step 4: Running the Scan

Once modules are selected, Spiderfoot starts scanning in the background. The progress can be monitored in **Scan Status**.

- Active modules show real-time data collection.
- The system automatically correlates results from multiple sources.
- Depending on the scope, scans can take minutes to hours.

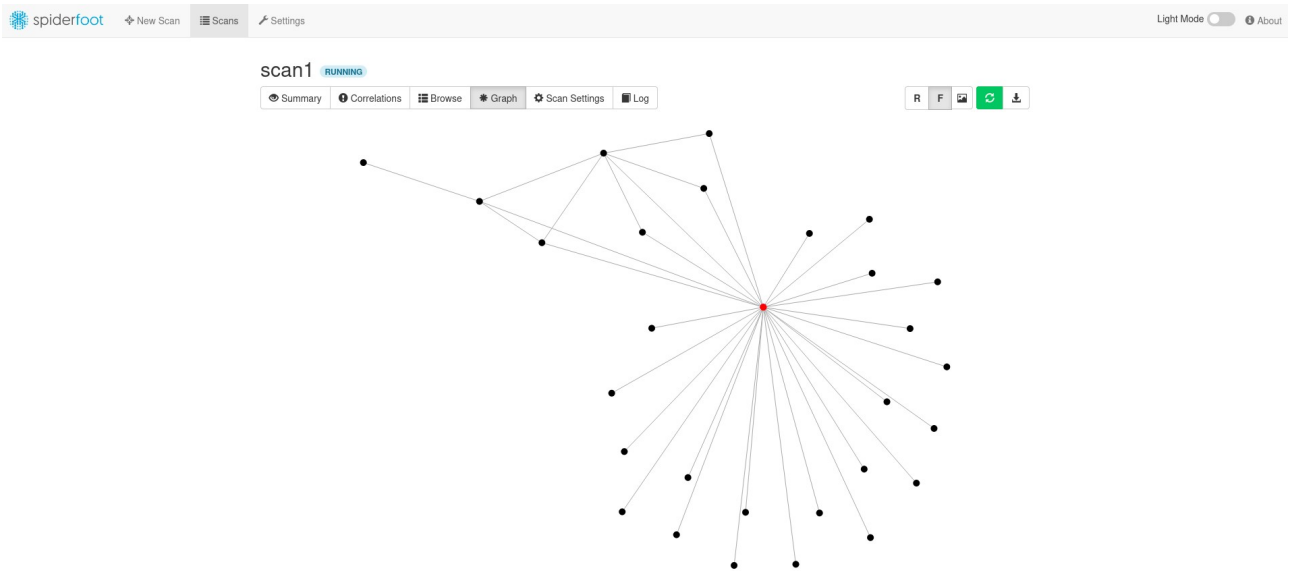


Step 5: Analyzing the Results

Viewing Data in the Web UI

Spiderfoot provides structured visualization:

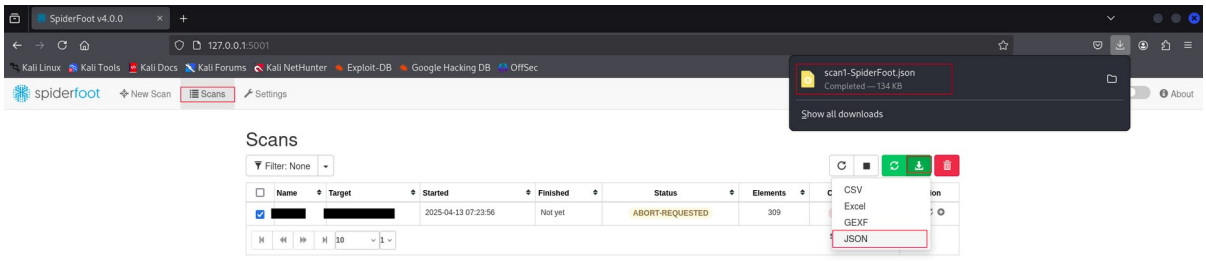
- **Graph View:** Shows relationships between entities like domains, emails, and IPs.
- **Raw Data:** Displays detailed logs from each module.
- **Dashboard:** Summarizes key findings.



Exporting Results

Spiderfoot allows exporting scan results in multiple formats:

1. After the scan completes, go to the **Scans** tab
2. Select the scan
3. On the top right of Scans, click “**Export**” and select “**JSON**” format



Alternatively, use the web UI to download results as:

- CSV for spreadsheets
- GEXF – For graph analysis (e.g. in Gephi).

Advanced Usage of Spiderfoot

1. API Integration

Spiderfoot can integrate with external APIs to enhance OSINT collection. Supported APIs include Shodan, VirusTotal, and Have I Been Pwned.

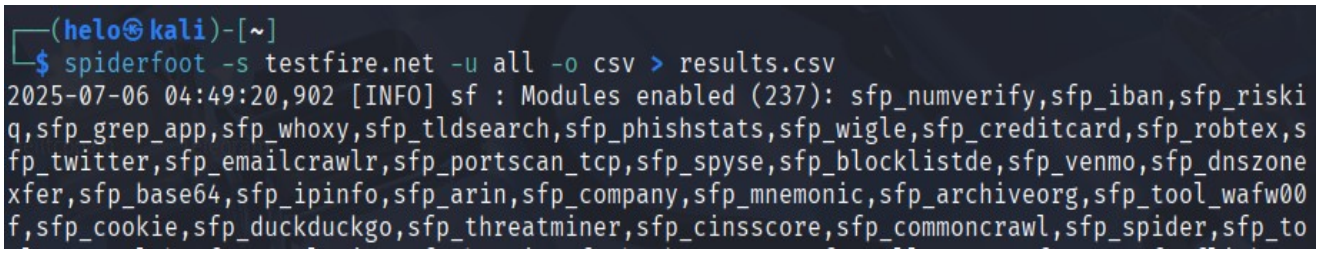
To configure an API:

- Navigate to **Settings > API Keys**.
- Enter the API key from the respective service.
- Save changes and rerun scans for enriched data.

2. CLI Mode for Automation

For those preferring command-line usage, Spiderfoot can be run in CLI mode:

- kali> spiderfoot -s testfire.net -u all -o csv > results.csv



This command scans testfire.net using all modules and saves results in results.csv.

Mitigation Steps

- Secure exposed services (e.g., disable unnecessary open ports).
- Change compromised credentials found in data leaks.
- Implement monitoring for threats mentioned on the dark web.

Best Practices for Using Spiderfoot

- **Define Clear Objectives:** Avoid unnecessary modules to keep scans efficient
- **Respect Legal Boundaries:** Use Spiderfoot responsibly and follow ethical guidelines
- **Leverage API Keys:** Unlock extended data sources for comprehensive results
- **Automate Scans:** Utilize CLI mode for scheduled reconnaissance tasks

REFERENCE

Hackercool: https://www.hackercoolmagazine.com/beginners-guide-to-spiderfoot/

Geekforgeeks: https://www.geeksforgeeks.org/linux-unix/spiderfoot-a-automate-osint-framework-in-kali-linux/

Infosectrain: https://www.infosectrain.com/blog/information-gathering-using-spiderfoot-a-practical-walkthrough/

Contact:

Gmail: ninkanong200620@gmail.com