## Lab 2: Information Gathering Using The Harvester

### Scenario

While working as a penetration tester in a cybersecurity company, you are assigned to perform the reconnaissance phase of a penetration test for a client. The goal is to gather as much Open-Source Intelligence (OSINT) as possible about the client’s public-facing infrastructure. Using the Parrot Security virtual machine, you decide to run theHarvester, a powerful OSINT tool, to collect information such as emails, subdomains, and IP addresses linked to the client’s domain. This information will help build a clear picture of the target’s digital footprint and potential attack surface.

### Solution

TheHarvester provides an effective way to perform footprinting by automating the collection of publicly available information about a target domain. Running on the Parrot Security virtual machine, it can query multiple search engines and data sources to gather emails, subdomains, IP addresses, and hostnames linked to the target. This data is crucial in the reconnaissance phase of penetration testing, as it helps build a clear picture of the organization’s external exposure. To make the results easier to analyze, theHarvester also allows exporting findings into a structured HTML report, which simplifies review and documentation.

|  |
| --- |
| 1. Turn on the Parrot Security virtual machine and log in with your credentials. Before starting this lab, integrate ShellGPT in the Parrot Security machine.  2. Open the terminal and write the command **theHarvester -v** to ensure Harvester is installed and working.    3. To launch an information gathering campaign on a target, type the command **theHarvester -d hackaday.com -l 300 -b bing,duckduckgo.** This will start theHarvester. It will begin searching bing and duckduckgo for the top 300 results related to hackaday.com. For this target, we found only one email using two search engines, and no additional information was available.    4. If we want to gather even more information about our target, we can specify the following commad **theHarvester -d hackaday.com -l 300 -b all.** The **-b all** tag will search all available search engines to the Harvester for information regarding hackaday.com. As you can see, it is an extremely useful tool for discovering email addresses, some other Urls, Hosts and IP addresses.        5. If we wanted to display this information in an easier-to-read format, we could add the **-f** tag at the end of the command **theHarvester -d hackaday.com -l 300 -b all -f hackaday.com.results.**    6. In Parrot OS, this saves the information gathered into **JSON** and **XML** files (e.g., **hackaday.com.results.json** and **hackaday.com.results.xml**). These files can then be opened with tools like **jq** (for JSON) or any XML viewer, allowing the results to be read in a structured and clearer way.    7. You can also use ShellGPT for gathering information using theHarvester. As you have already incorporated SHellGPT in your Parros OS machine, run **sgpt --chat footprint --shell “Use theHarvester to gather email accounts associated with 'microsoft.com', limiting results to 200, and leveraging 'baidu' as a data source”** command to a target organization for harvesting emails. You can write the script according to your requirements.      10. This concludes the demonstration of this lab. |