# Module 02: Footprinting and Reconnaissance

Footprinting is the first step in the hacking process. It involves gathering a blueprint of a target's network, security posture, and organizational structure. In CEH v13, this is largely referred to as **OSINT (Open Source Intelligence)**.

## 1. Footprinting Concepts

### Objectives of Footprinting

* **Knowledge of Security Posture:** Understanding firewalls, IDSs, and remote access.
* **Focus on Specific Ranges:** Finding IP addresses, subnets, and network blocks.
* **Identify Vulnerabilities:** Discovering outdated systems or misconfigurations.
* **Map the Network:** Creating a visual map of the target's infrastructure.

### Types of Footprinting

1. **Passive Footprinting:** Gathering information without direct interaction with the target's systems (e.g., searching social media, WHOIS, or Google).
2. **Active Footprinting:** Gathering information by interacting with the target's systems (e.g., using traceroute or querying their DNS server).

## 2. Google Hacking (Google Dorking)

Ethical hackers use advanced search operators to find sensitive information that is indexed but not meant to be public.

* **site:** Limits search to a specific domain (e.g., site:microsoft.com).
* **filetype:** Searches for specific files (e.g., filetype:pdf "confidential").
* **intitle:** Searches for terms in the webpage title (e.g., intitle:"index of").
* **inurl:** Searches for terms in the URL (e.g., inurl:admin).
* **cache:** Shows the version of the site Google has stored.

## 3. Website and Social Media Footprinting

### Website Analysis

* **Wappalyzer / BuiltWith:** Identify the CMS, backend languages (PHP, Python), and web server (Apache, Nginx).
* **HTTrack / Website Copier:** Used to mirror (clone) a website to browse it offline and search for hidden scripts or developer comments.
* **Burp Suite:** Intercepting requests to see hidden headers and parameters.

### Social Engineering Recon

* **LinkedIn:** The "Gold Mine" for finding employee names, roles, and software used in the company (via job postings).
* **theHarvester:** A tool to gather emails, subdomains, and hostnames from public sources like LinkedIn and Google.
* theHarvester -d example.com -b google

| **Option** | **Meaning** |
| --- | --- |
| -d | Target domain |
| -b | Data source (google, bing, shodan, etc.) |

## 4. Network and DNS Footprinting

### WHOIS Lookup

Provides details about domain ownership, registrar, contact info, and name servers.

### DNS Records (KNOW THESE FOR THE EXAM)

|  |  |
| --- | --- |
| **Record** | **Function** |
| **A** | Maps a hostname to an IPv4 address. |
| **AAAA** | Maps a hostname to an IPv6 address. |
| **MX** | Mail Exchange; identifies mail servers for the domain. |
| **NS** | Name Server; identifies the authoritative DNS for the domain. |
| **CNAME** | Canonical Name; an alias for another domain. |
| **SOA** | Start of Authority; contains zone management info. |
| **PTR** | Pointer Record; used for Reverse DNS (IP to Name). |
| **TXT** | Text record; often used for SPF/DKIM (email security). |

### DNS Zone Transfer (AXFR)

An attempt to get the entire DNS database for a domain. If a DNS server is misconfigured, an attacker can use dig axfr @nameserver domain.com to get a list of every host in the company.

### Network Path Analysis

* **Traceroute:** Uses the **ICMP TTL (Time to Live)** field to map the path between your computer and the target. Each "hop" is a router.
* **Visual Traceroute:** Tools like NeoTrace or VisualRoute.

## 5. Footprinting Tools to Memorize

* **Maltego:** A visual link analysis tool that shows relationships between people, companies, domains, and IP addresses.
* **Recon-ng:** A full-featured reconnaissance framework written in Python (looks like Metasploit).
* **FOCA:** Used to find metadata in documents (PDF, DOCX) exposed on a website.
* **Shodan:** The "Search Engine for IoT." Finds connected devices like webcams, industrial control systems, and servers.
* **Censys:** Similar to Shodan, used for searching internet-connected devices.
* **Sherlock:** Search for usernames across hundreds of social networks.
* **Photon:** A fast crawler used for extracting URLs, emails, and files from a target website.

## 6. Hands-On Lab Sessions (CEH v13 Practice)

### Lab 1: Google Hacking Database (GHDB)

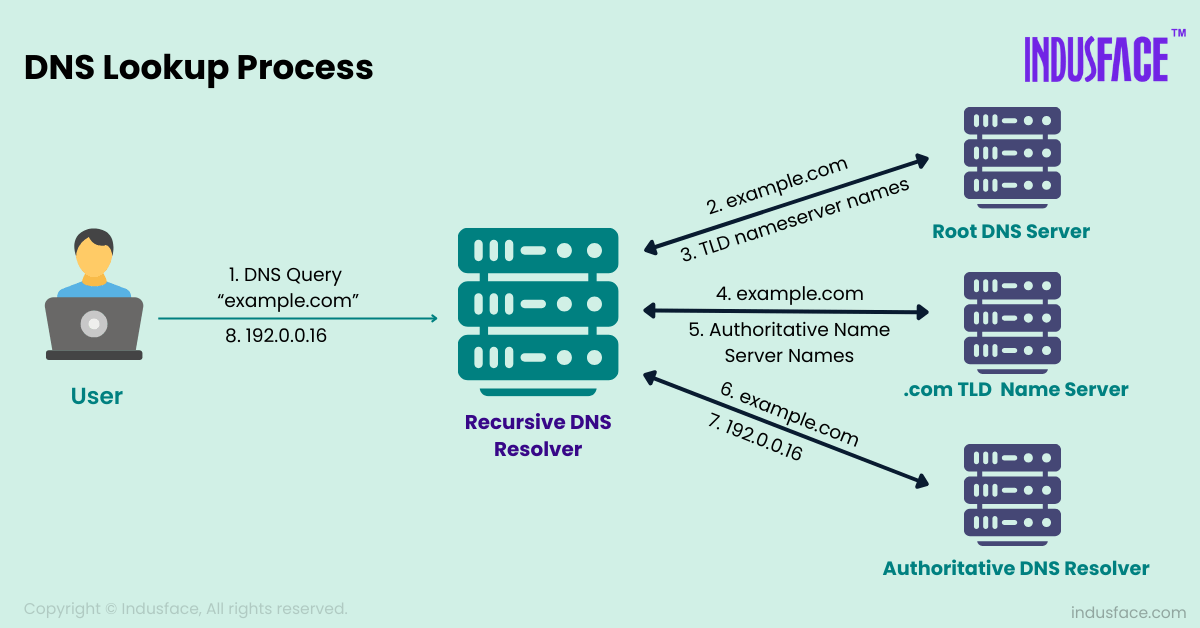
* **Goal:** Use GHDB to find vulnerable servers.
* **Action:** Navigate to exploit-db.com/google-hacking-database and use dorks to find login portals or exposed directories.
* **Key Command:** intitle:"index of" "parent directory"

### Lab 2: DNS Enumeration with nslookup and dig

* **Goal:** Identify mail servers and attempt a zone transfer.
* **Commands:**
  + nslookup kaushalya.tech
  + nslookup google.com

***Find mail servers***

nslookup -type=mx target.com



* + **dig** — Domain Information Groper
  + **dig target**.com
  + dig target.com any
  + Find IPv4 address (A record) - dig google.com
  + Find IPv6 address (AAAA) - dig google.com AAAA
  + Mail servers (MX) – dig google.com MX
  + AXFR = **Authoritative Zone Transfer**
  + dig axfr example.com @ns1.example.com
  + dig axfr @ns1.target.com target.com (To check for misconfigured zone transfers)

### Lab 3: Metadata Extraction with FOCA

* **Goal:** Extract hidden info from public documents.
* **Action:** Load FOCA, point it at a target domain, download PDFs/DOCXs, and click "Extract Metadata" to find usernames, software versions, and internal server paths.

### Lab 4: Social Media Recon with theHarvester

* **Goal:** Automate the gathering of emails and subdomains.
* **Command:** theHarvester -d target.com -b google,linkedin,bing -l 500

### Lab 5: Link Analysis with Maltego

* **Goal:** Visualize the relationship between a person and a company.
* **Action:** Open Maltego, drag a "Domain" entity onto the graph, and run "To DNS Name" and "To Email Address" transforms to map the infrastructure.

## 7. Footprinting Countermeasures

* **Hide WHOIS Info:** Use "Privacy Protection" services to mask contact details.
* **Disable DNS Zone Transfers:** Only allow zone transfers to trusted secondary DNS servers.
* **Monitor Search Engines:** Use Google Alerts to see what info is being indexed.
* **Social Media Policy:** Educate employees on what not to post (sensitive tech stacks, etc.).
* **Metadata Removal:** Ensure all public documents have metadata stripped.