TRAINING DAY 16 REPORT:

- ## Installation of Tor

  Tor (The Onion Router) is a browser and a privacy network that anonymizes internet traffic.

  ### To install Tor on Kali Linux:

  ```
  sudo apt update
  sudo apt install tor -y
  ```

  Once installed, start the Tor service:

  ```
  sudo systemctl start tor
  sudo systemctl enable tor
  ```

  For browser version (GUI):

  ```
  sudo apt install torbrowser-launcher
  torbrowser-launcher
  ```

- ## Tor Browser Settings

  Tor Browser can be customized for:

  1. Security Level: Standard / Safer / Safest

  2. NoScript settings: Block JavaScript on untrusted websites

  3. Bridge connections: Bypass censorship

  4. Privacy & Security: Disable history, cookies, and fingerprinting

  Access via:
  (menu) > **Settings** > **Privacy & Security**

- ## Tor in Kali Linux

  Once installed, you can use **Tor in terminal** as a **SOCKS5 proxy**:

  Start Tor service:

  ```
  sudo systemctl start tor
  ```

  Route traffic through Tor using `proxychains`:

```
proxychains firefox
```

Or configure tools like `curl` or `nmap`:

```
proxychains curl https://check.torproject.org
```

Make sure `/etc/proxychains.conf` has:

```
socks5 127.0.0.1 9050
```

- ## Fix the Error in Tor Browser

  Common errors:

  1. Tor Browser not opening

  2. Cannot connect to Tor network

  3. Signature verification failed

  Fixes:

  1. Make sure system time is correct.

  2. Reinstall using:

```
sudo apt purge torbrowser-launcher
sudo apt install torbrowser-launcher
```

  Try running with:

```
torbrowser-launcher --settings
```

  Or download latest Tor manually from:
  https://www.torproject.org/download

- ## Introduction to Footprinting / Reconnaissance

  Footprinting (or **Reconnaissance**) is the **first phase of ethical hacking**.

  Goal: Gather as much information as possible about a **target system, network, or organization**.

  ### Types:

1. Passive Footprinting: Collecting data without directly interacting with the target (e.g., via search engines, WHOIS, DNS records).

2. Active Footprinting: Direct interaction with the target (e.g., ping, traceroute, port scanning).

Example:

Finding an organization's IP range, subdomains, employee emails, or exposed technologies before attempting any attacks.

- ## Footprinting Through Search Engines

Search engines like **Google**, **Bing**, and **DuckDuckGo** can be powerful tools for gathering target information.

Information you can gather:

1. Cached pages and hidden directories

2. Employee details, emails, office locations

3. Past security issues or data leaks

4. File types using `filetype:` (e.g., `.pdf`, `.docx`)

Google Dorking Example.

- ## Introduction to OSINT (Open Source Intelligence)

OSINT is the collection and analysis of **publicly available information**.

Sources of OSINT:

1. Social media (LinkedIn, Twitter, Facebook, Instagram)

2. Public records, job portals

3. News articles, forums, GitHub

4. Shodan (for IoT devices)

Example: Finding a company’s internal tools on GitHub, or exposed credentials on Pastebin.

- ## Email Footprinting

  This involves gathering information about email IDs related to the target.

  ### What you can learn:

  1. Email patterns (e.g., firstname.lastname@company.com)
  2. Validity of email addresses
  3. Employee details via email
  4. Possible phishing targets

  ### Tools & Techniques:

  1. theHarvester
  2. Hunter.io
  3. Email verification tools (e.g., verify-email.org)
  4. Social engineering possibilities

- ## Website Footprinting

  Gathering all possible information about a target **website/domain**.

  Includes:

  1. Technologies used (CMS, server, frameworks)
  2. Subdomains (e.g., `dev.target.com`)
  3. Robots.txt file
  4. WHOIS info
  5. DNS records (A, MX, TXT)
  6. File paths exposed

  ### Tools:

  1. Netcraft
  2. BuiltWith
  3. Wappalyzer

4. DNSdumpster

5. Nikto (vulnerability scanner)

Footprinting Using Google

Google Hacking or **Google Dorking** uses advanced search queries to extract sensitive data.

You can find:

1. Exposed credentials
2. Login portals
3. Database files
4. Error logs
5. Admin panels

Examples:

```
intitle:"index of" site:target.com
filetype:log inurl:"/logs/"
```

Competitive Intelligence

This involves collecting and analyzing info about **business competitors** through legal and ethical means.

Techniques:

1. Analyzing competitor websites, press releases
2. Tracking job postings (to know what tech they use)
3. Studying reviews, investor reports
4. Monitoring patents, social media, blogs

Useful for:

1. Business strategy
2. Marketing

3. Understanding vulnerabilities or market gaps

- ## Internet Archive

  The **Internet Archive** is a **free, non-profit digital library** that preserves and provides access to historical versions of websites, books, audio, videos, and software.

  ### Key Tool: Wayback Machine

  - Lets you **view old versions of websites** by date.

  - Example use: Investigate changes in a company’s site, recover deleted pages, or analyze historical web content.

  - Website: web.archive.org

  ### Other Features:

1. **Books Library** – Millions of digitized books, including rare and historical texts.
2. **Software Archive** – Old operating systems, games, and tools for testing or emulation.
3. **Audio/Video Library** – Public domain and user-uploaded media (lectures, music, documentaries).
4. **TV News Archive** – Search and watch news broadcasts for media research.

- ## What is a Web Crawler?

  A **Web Crawler** (also called a **spider** or **bot**) is a software program that **automatically browses the internet** and **indexes web pages** for search engines or data collection.

  ### What It Does:

  1. Starts with a list of URLs (called seeds).

  2. Visits each URL, reads the page content.

  3. Extracts hyperlinks from the page.

  4. Adds new links to the crawl queue and repeats the process.

Uses:

Purpose	**Description**
Search Engine Indexing	Google, Bing use crawlers to index websites
OSINT & Footprinting	Hackers & analysts use tools like **HTTrack** or **Maltego** to map target sites
Market Intelligence	Companies track competitors' prices or content
Security Testing	Used to find exposed or vulnerable pages

Ethical/Legal Note:

1. Web crawling should respect **robots.txt** rules.
2. Unauthorized or aggressive crawling can result in **IP bans** or **legal action**.