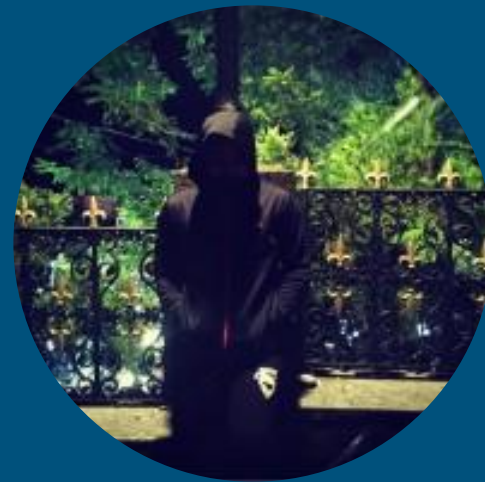


# Web Application Penetration Testing Series



---

## **Scraping Urls From The Wayback Machine With WayBackUrls**

# Who Am I

---



**Sumit Jain**  
(Ethical Hacker & Cyber Security Expert)  
Synack & Pentabug Red Team Member

# Fetch Urls With Waybackurls

- *github : <https://github.com/tomnomnom/waybackurls>*
- *Installation* :
  - i. Run this in your terminal

“ `sudo go install github.com/tomnomnom/waybackurls@latest && sudo cp /root/go/bin/waybackurls /usr/bin` ”
  - ii. This will install in your Root Directory
  - iii. Run with `waybackurls -h` ( help command )

# Waybackurls : Useful Commands

---

- `waybackurls <target>` : This command retrieves all the URLs of the Wayback Machine archive for the specified domain or target
- `waybackurls <target> -json` : This command retrieves all the URLs of the Wayback Machine archive for the specified domain or target in JSON format

# Waybackurls : Useful Commands

---

- `waybackurls <target> | grep <keyword>` : This command retrieves all the URLs of the Wayback Machine archive for the specified domain or target that contain the specified keyword
- `waybackurls <target> | httpprobe` : This command retrieves all the URLs of the Wayback Machine archive for the specified domain or target and tests them for HTTP/HTTPS connectivity

# Waybackurls : Useful Commands

---

- `waybackurls <target> -exclude <exclude-file> :`  
This command retrieves all the URLs of the Wayback Machine archive for the specified domain or target, but excludes the URLs listed in the specified file.
- `waybackurls <target> -filter "status_code:200" | sort -u :` This command retrieves all the URLs of the Wayback Machine archive for the specified domain or target that return a 200 status code.



**Sumit Jain**

**Follow me on**

**Twitter :**     **@sumit\_cfe**

**Linkedin :**   **@sumitthehacker**