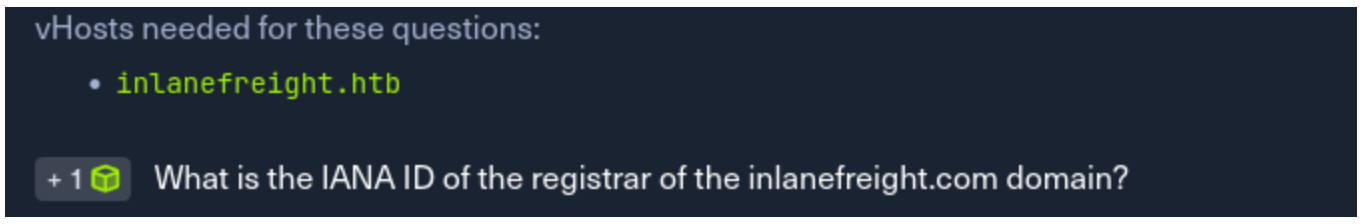
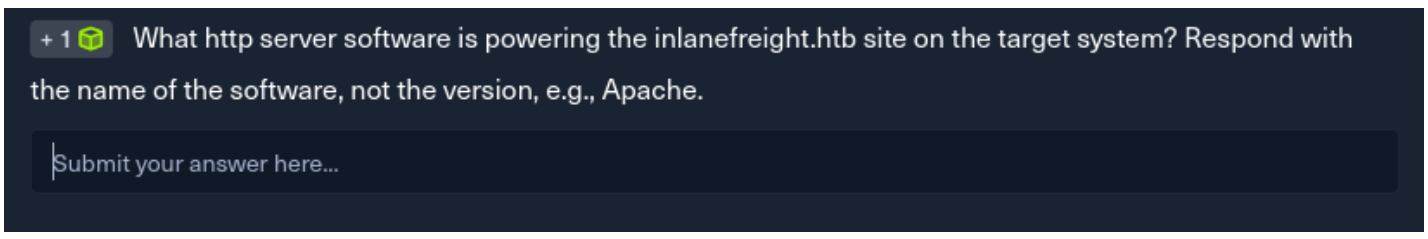
Web Information Gather HTB



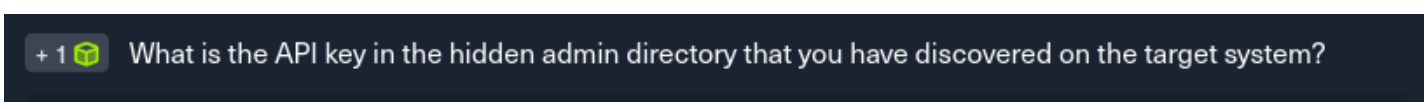
- to find this what we can do is run `whois`

```
whois inlanefreight.com | grep "IANA ID"
```



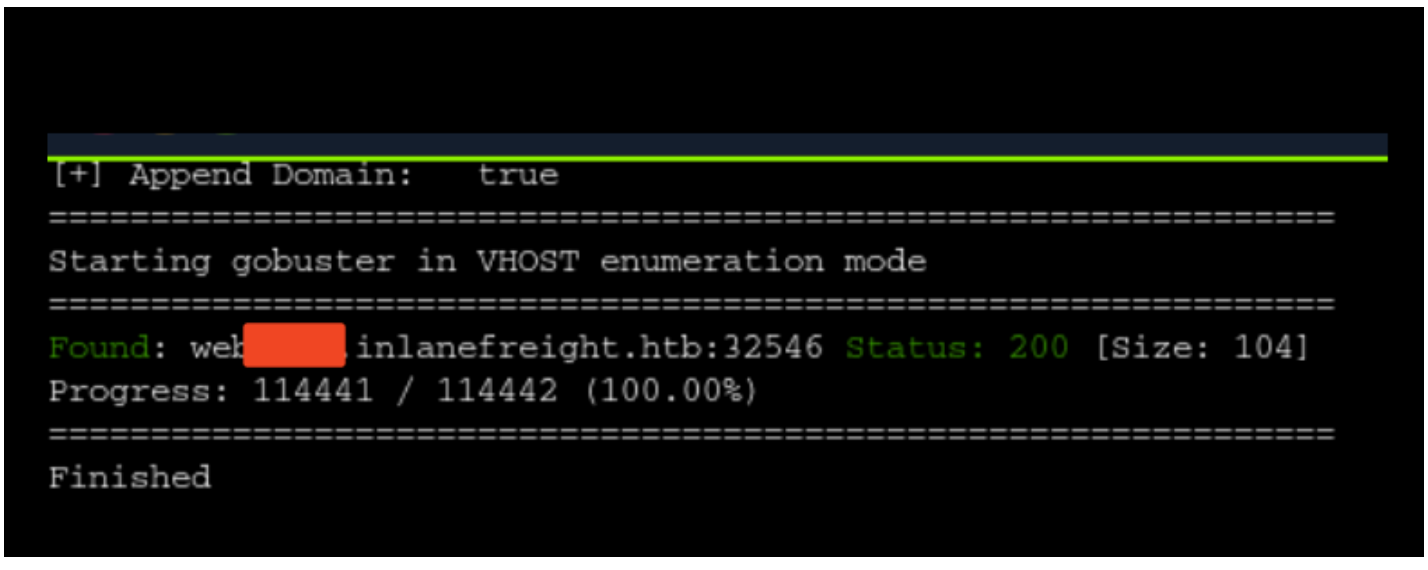
- to complete this we can use a tool like `whatweb`

```
┌──[us-academy-4]─[10.10.15.112]─[htb-ac-1326293@htb-emddjjl2bv]─[~]
└──╼ [★]$ whatweb http://inlanefreight.htb:32546/
```



- Run VHOST Scan:

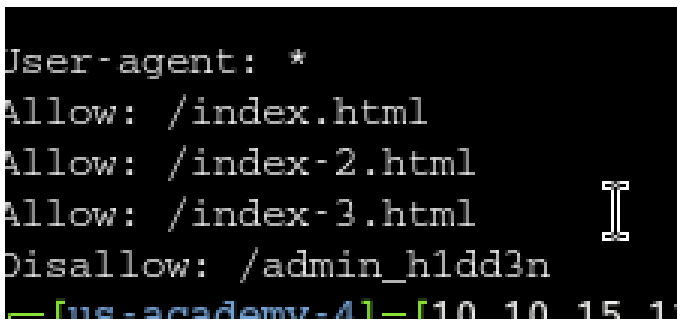
```
gobuster vhost -u http://inlanefreight.htb:32546 -w subdomains-top1million-110000.txt --append-domain
```



USE Curl to Navigate and Check The Site:

```
─[us-academy-4]─[10.10.15.112]─[htb-ac-1326293@htb-8mfc7imjzh]─[/root]
└──╼ [★]$ curl -i http://web1337.inlanefreight.htb:32546/robots.txt
```

- Check if Robots.txt has anything.



+4 After crawling the inlanefreight.htb domain on the target system, what is the email address you have found? Respond with the full email, e.g., mail@inlanefreight.htb.

- we will need to use the python webcrawler `scrapy`

```
─[us-academy-4]─[10.10.15.112]─[htb-ac-1326293@htb-8mfc7imjzh]─[~]
└──╼ [★]$ pip3 install scrapy

─[us-academy-4]─[10.10.15.112]─[htb-ac-1326293@htb-8mfc7imjzh]─[~]
└──╼ [★]$ wget -O ReconSpider.zip https://academy.hackthebox.com/storage/modules/144/ReconSpider.v1.2.zip
--2025-05-07 14:38:50--  https://academy.hackthebox.com/storage/modules/144/ReconSpider.v1.2.zip
Resolving academy.hackthebox.com (academy.hackthebox.com)... 109.176.239.69, 109.176.239.70
Connecting to academy.hackthebox.com (academy.hackthebox.com)|109.176.239.69|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 1706 (1.7K) [application/zip]
Saving to: 'ReconSpider.zip'

ReconSpider.zip                     100%[===================================================================>]   1.67K  --.-KB/s    in 0s

2025-05-07 14:38:51 (17.4 MB/s) - 'ReconSpider.zip' saved [1706/1706]

┌─[us-academy-4]─[10.10.15.112]─[htb-ac-1326293@htb-8mfc7imjzh]─[~]
└──╼ [★]$ unzip ReconSpider.zip
Archive:  ReconSpider.zip
  inflating: ReconSpider.py
```

```
─[us-academy-4]─[10.10.15.112]─[htb-ac-1326293@htb-8mfc7imjzh]─[/usr/share/wordlists/seclists/Discovery/DNS]
└──╼ [★]$ gobuster vhost -u http://web1337.inlanefreight.htb:32546 -w subdomains-top1million-110000.txt --append-domain
===============================================================
Gobuster v3.6
by OJ Reeves (@TheColonial) & Christian Mehlmauer (@firefart)
===============================================================
[+] Url:            http://web1337.inlanefreight.htb:32546
[+] Method:         GET
[+] Threads:        10
```

```
[+] Wordlist:       subdomains-top1million-110000.txt
[+] User Agent:     gobuster/3.6
[+] Timeout:        10s
[+] Append Domain:  true
===============================================================
Starting gobuster in VHOST enumeration mode
===============================================================
Found: XXX.webXXXX.inlanefreight.htb:32546 Status: 200 [Size: 123]
Progress: 488 / 114442 (0.43%)^C
[!] Keyboard interrupt detected, terminating.
Progress: 525 / 114442 (0.46%)
===============================================================
Finished
===============================
```

- using this result we can again curl the site and vist it.

```
tb-ac-1326293@htb-8mfc7imjzh]—[/usr/share/wordlists/seclists/Discovery/DNS]
└──╼ [★]$ curl -i http://dev.web1337.inlanefreight.htb:32546/robots.txt
```

- noticed a werid hyperlink:

- lets visit that endpoint
- it was like never ending loop so lets let the recon spider do it’s thing.
  - from this i could solve last to lab with it.