

# INFORMATION GATHERING

LLC 500 Cybersecurity and Ethical Hacking  
Linux Learning Centre Ltd.

[www.linuxlearningcentre.co.ke](http://www.linuxlearningcentre.co.ke)



# AGENDA

- Introduction
- Active and Passive Recon
- Asset Discovery
- Content Discovery
- IpAddress Discovery
- Domain/Subdomain Discovery
- Email Discovery
- Network/Port Scanning
- Business Communication Infrastructure Discovery



# AGENDA

Source Code Aggregators

Cloud Infrastructure Discovery

Company Information And Associations

Internet Survey Data

Social media Profiling

Data Leaks

Internet Scan

# RECON

- Recon in general Information Gathering , Just collecting info about some organisation or personal we can say it as recon
- Recon in general Will be described as information collection/OSINT/Discovery Etc
- Two Types in general:
- Active and passive recon

# ACTIVE VS PASSIVE

- What is active reconnaissance? Active reconnaissance is a way of finding out information that does leave a footprint. (Think of a footprint like a digital signature, your thumb has a footprint, and so does your online activity although in a more abstract way) It involves an attempt to figure out things like the OS (Operating System) being used, any open ports, (a port being a pathway into a network basically. This is important because if you can find an open port, you can most likely find a way to get into the network) email addresses of the employees, etc.
- Gathering information without alerting the subject of the surveillance is passive reconnaissance. This is the natural start of any reconnaissance because, once alerted, a target will likely react by drastically increasing security in anticipation of an attack. This is like casing a place prior to robbing it.



# ASSET DISCOVERY

- Movable Assets
- Immovable assets



# CONTENT DISCOVERY

- Finding Sensitive Directories , Ofcourse Start with google , if u are ok start with familiar informations:

## Extensions

site:<http://example.com> filetype:php

site:<http://example.com> filetype:aspx

site:<http://example.com> filetype:swf

site:<http://example.com> filetype:wsdl

## Directory structure

site:<http://example.com> intext:"index of /"

## Juicy stuff

site: <http://target.com> filetype:txt

site: <http://target.com> inurl:.php.txt

site: <http://target.com> ext:txt

# CONTENT DISCOVERY

Discovering data may be in any form of data , Let me say some common content discovery methods

- News Discovery Apps and tools
- Social Search
- Twitter News Discovery
- Social news Discovery
- StartupAnd New tools discovery
- Video Content Discovery
- Content Trends
- RSS search Engines
- Alerts make use of google alerts
- Image Discovery
- Search Discovery



# SOME LINKS FOR CONTENT DISCOVERY

[https://content-discovery-  
tools.zeef.com/robin.good?utm\\_content=buffer3f3e3&utm\\_medium=social&ut  
m\\_source=twitter.com&utm\\_campaign=buffer](https://content-discovery-tools.zeef.com/robin.good?utm_content=buffer3f3e3&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer)

# IPADDRESS DISCOVERY

- [Mxtoolbox](#): Bulk Domain/IP lookup tool
- [Domaintoipconverter](#): Bulk domain to IP converter
- [Massdns](#): A DNS resolver utility for bulk lookups
- [GoogleappsDig](#): Online Dig tool by Google
- [DataSploit\(IP Address Modules\)](#): An OSINT Framework to perform various recon techniques
- [Domain Dossier](#): Investigate domains and IP addresses
- [Bgpview](#): Search ASN, IPv4/IPv6 or resource name
- [Hurricane Electric BGP Toolkit](#): Keyword to ASN lookup
- [Viewdns](#): Multiple domain/IP tools
- [Ultratoolsipv6Info](#): Multiple information related to IPv6 address
- [Whois](#): Command line utility usually used to find information about registered users/assignees of an Internet resource.
- [ICANN Whois](#): Whoisservice by Internet Corporation for Assigned Names and Numbers (ICANN)
- [NslookupLinux/Windows](#): Command line utility usually used for querying the DNS records
- [bgp](#): Internet Backbone and Colocation Provider ... Hurricane Electric IP Transit. Our Global Internet Backbone provides IP Transit with low latency, access to thousands of networks, and dual-stack

# JUST SMALL INFO ABOUT DNS RECORDS

Commonly used record types

- A(Host address)
- AAAA(IPv6 host address)
- ALIAS(Auto resolved alias)
- CNAME(Canonical name for an alias)
- MX(Mail eXchange)
- NS(Name Server)
- PTR(Pointer)
- SOA(Start Of Authority)
- SRV(location of service)
- TXT(Descriptive text)

# DNS RECORDS CONTINUED

Records types used for DNSSEC

- DNSKEY(DNSSEC public key)
- DS(Delegation Signer)
- NSEC(Next Secure)
- NSEC3(Next Secure v. 3)
- NSEC3PARAM(NSEC3 Parameters)
- RRSIG(RRsetSignature)

# DNS RECORDS CONTINUED

Less commonly used record types:

- AFSDB(AFS Data Base location)
- ATMA(Asynchronous Transfer Mode address)
- CAA(Certification Authority Authorization)
- CERT(Certificate / CRL)
- DHCID(DHCP Information)
- DNAME(Non-Terminal DNS Name Redirection)
- HINFO(Host information)
- ISDN(ISDN address)
- LOC(Location information)
- MB, MG, MINFO, MR(mailbox records)
- NAPTR(Naming Authority Pointer)
- NSAP(NSAP address)
- RP(Responsible person)
- RT(Route through)
- TLSA(Transport Layer Security Authentication)
- X25(X.25 PSDN address)

# BGP

- I recommend you to go through this before starting with BGP
- What is TCP
- What is peering
- What is Rib failure
- Why BGP is important

Some Tools or scripts :

<https://github.com/tdubs/bgp-recon>

Please go through about BGP Hijacking

# DOMAIN / SUBDOMAIN DISCOVERY

- [SubFinder](#): SubFinder is a subdomain discovery tool that discovers valid subdomains for websites. Designed as a passive framework to be useful for bug bounties and safe for penetration testing.
- [Amass](#): A subdomain enumeration utility
- [Sublist3r](#): Subdomains enumeration tool with multiple sources
- [Aiodnsbrute](#): Asynchronous DNS brute force utility
- [LDNS](#): A DNS library useful for DNS tool programming
- [Dns-nsec3-enum](#): NmapNSE Script for NSEC3 walking
- [Nsec3map](#): A tool to NSEC and NSEC3 walking
- [Crt.sh](#): Domain certificate Search
- [Ct-exposer](#): A tool to discover sub-domains by searching Certificate Transparency logs
- [Certgraph](#): A tool to crawl the graph of certificate Alternate Names
- [Appsecco-The art of subdomain enumeration](#): The supplement material for the book "The art of sub-domain enumeration"
- [SSLScrape](#): A scanning tool to scrape hostnames from SSL certificates
- [Wolframalpha](#): Computational knowledge engine
- [Project Sonar](#): Forward DNS Data
- [Project Sonar](#): Reverse DNS Data
- [GoBuster](#): Directory/File, DNS and VHostbusting tool written in
- [Go Bluto](#): Recon, Subdomain Bruteforce, Zone Transfers



# SOME IMPORTANT CONCEPT

- Check Forward DNS Data
- Check Reverse DNS Data
- <https://rapiddns.io/>

# EMAIL DISCOVERY

\* Through Google Dorks

- Hunter: Email search for a domain
- Skrapp: Browser addonto find emails on LinkedIn
- Email Extractor: Chrome extension to extract emails from web pages
- Convertcsv: Online tool to extract email addresses in text, web pages, data files etc.
- linkedin2username: OSINT Tool: Generate username lists for companies on LinkedIn
- Office365UserEnum: Enumerate valid usernames from Office 365 using ActiveSync.

# NETWORK PORT SCANNING

- Zmap: A fast network scanner designed for Internet-wide network surveys
- Masscan: An asynchronously TCP port scanner
- ZMapv6: A modified version of Zmap with IPv6 support.
- Nmap: A free and open source utility for network discovery. The most popular port scanner

# BUSINESS COMMUNICATION INFRASTRUCTURE DISCOVERY

- Mxtoolbox: Online tool to check mail exchanger (MX) records
- MicroBurst: PowerShell based Azure security assessment scripts
- Lyncsmash: Tools to enumerate and attack self-hosted Lync/Skype for Business
- Enumeration-as-a-Service: Script for SaaS offering enumeration through DNS queries
- ruler: A tool to abuse Exchange service

# SOURCE CODE AGGREGATORS / SEARCH -INFORMATION DISCOVERY

- Github: GithubAdvanced Search
- Bitbucket: BitbucketSearch using Google
- Gitrob: Reconnaissance tool for GitHub organizations
- Gitlab: Search Gitlabprojects
- Publicwww: Source Code Search Engine
- builtwith: Web technology information profiler tool. Find out what a website is built with.

# CLOUD INFRASTRUCTURE DISCOVERY

- CloudScraper: A tool to spider websites for cloud resources (S3 Buckets, Azure Blobs, DigitalOceanStorage Space)
- InSp3ctor: AWS S3 Bucket/Object finder
- Buckets Grayhatware: Search for Open Amazon s3 Buckets and their contents
- Spaces-finder: A tool to hunt for publicly accessible DigitalOceanSpaces
- GCPBucketBrute: A Google Storage buckets enumeration script
- CloudStorageFinder: Tools to find public data in cloud storage systems

# COMPANY INFORMATION AND ASSOCIATIONS

- Mainly To Search Acquisitions, For example Facebook acquisition Whatsapp
- Crunchbase: Information about companies (funding, acquisition, merger etc.) and the people behind them
- Companieshouse: United Kingdom's registrar of companies
- OverSeasRegistries: List of company registries located around the world
- Opencorporates: Open database of companies in the world



# INTERNET SURVEYDATA

- [Project Sonar](#): Rapid7's internet-wide surveys data across different services and protocols
- [Scans.io](#): Internet-Wide Scan Data Repository, hosted by the ZMapTeam
- [Portradar](#): Free and open port scan data by packet.tel

# SOCIAL MEDIA \EMPLOYEE PROFILING

- LinkedInt: A LinkedIn scraper for reconnaissance
- Glassdoor: Company review and rating search
- SocialBlade: Track user statistics for different platforms including YouTube and Twitter
- Social-Searcher: Social Media Search Engine
- Checkuser: Social existence checker

# DATA LEAKS

- Dumpmon: A twitter bot which monitors multiple paste sites for password dumps and other sensitive information
- Pastebin\_scraper: Automated tool to monitor pastebin for interesting information
- Scavenger: Paste sites crawler (bot) looking for leaked credentials
- Pwnbin: Python based Pastebin crawler for keywords.
- PwnedOrNot: Tool to find passwords for compromised accounts
- Haveibeenpwned.com
- Ghostproject.fr

# SOME BREACH INFO SITES

- <https://cybernulled.com>
- <https://kleoz.net/>
- <https://cracking.org/>
- <https://demonforums.net/>
- <https://www.crackingpro.com/>
- <https://crackingking.com/>
- <https://www.leakzone.net/>
- <https://www.crackingdrift.to>
- <https://crackingitaly.co>

## Archived Information

- Cachedviews: Cached view of pages on the Internet from multiple sources
  - WaybackMachine: Internet Archive
  - Shodan: Search engine for Internet-connected devices
  - Censys: Another search engine for internet-connected devices
  - Zoomeye: Cyberspace Search Engine • <https://github.com/sdnewhop/grinder>
-

# INTELLIGENCE GATHERING

## Scenarios

- Mostly, there are only two scenarios either we are outside/ inside the organization.

### Outside -External

- If we are outside or doing an external pentest. We need to figure out the attack surface area first.
- This could be achieved by answering the following questions:
  - What are the
    - Domain/ subdomains present? (like example.com --domain; ftp.example.com --subdomain)
    - IP Addresses/ Network ranges/ ASN Number(s) assigned?
    - Different Services (open ports) running on those IP Addresses?
    - Email addresses or People working for the organization?
    - Different Operating Systems/ Software used in the organization?
  - Additionally it is also interesting to know if there have been any security breaches in the past.
  - We might be able to compromise user credential(s) or running vulnerable service(s) and get inside the internal network of the organization.

# INTELLIGENCE GATHERING

## **Inside -Internal**

- When we are inside the organization (let's say physically) there are two common situations:
- Potentially, posing as an employee (already having access to the internal network).
- External consultant (with no internal network access as of now).
- Let's first explore what options we have as external consultant sitting in a conference room.



# INTELLIGENCE GATHERING

## Wired LAN

- If there's a LAN cable laying around and we (obviously) plug it in our computer, the following situations can occur:
- DHCP (Dynamic Host Configuration Protocol) is enabled and your machine is provided with an IP Address.
- DHCP is disabled; however the LAN cable is working. In this case, we might be able to sniff the network and figure out the near-by IP Address, netmask and default gateway and configure our device to use a static IP.
- Network Access Control is enabled, then probably we would need to search for
- A device (such as printers) attached to network, clone it's MAC address and try again) or
- IP Phones or any Hub or
- Connect USB to LAN device to any already connected machine.
- LAN port is disabled (We can't do much here! Can we?).

# INTELLIGENCE GATHERING

## **Wireless LAN**

- Check for Open/ Guest Wi-Fi -If you are connected somehow try to access the internal network range(s). Most probably, the organization would have segregated the network properly. However, sometimes DNS Names can be resolved.
- Check if any WEP/ WPA2 networks are present. If so, we can try to crack them.

# INTELLIGENCE GATHERING

Once, you are inside, probably the first thing would be to utilize **Responder** or **Inveigh** in **Analyze mode**.

Basically it's like this

- A user wants to access a file server named "NAS001" by \\NAS001, however, mistakenly types \\NAS01.
- The query goes to the DNS server to resolve the IP address of NAS01. However, as it's not a valid hostname, DNS Server responds to the user saying that it doesn't know that host.
- The user broadcasts on the local network and asks if anyone knows who is \\NAS01
- The attacker (if on the same network) seizing the opportunity says "I am NAS01 here is my IP Address"
- The user believes the attacker and sends its own username and NTLMv2 hash to the attacker.
- The attacker gathers all the hashes and cracks them (offline) to gain password.

# PASSIVE FINGERPRINTING

## Other Tools :

- [McafeeSite Digger](#) searches Google's cache to look for vulnerabilities, errors, configuration issues, proprietary information, and interesting security nuggets on web sites.
- [SearchDiggityv3](#) is Bishop Fox's MS Windows GUI application that serves as a front-end to the most recent versions of our Diggity tools: GoogleDiggity, BingDiggity, Bing, LinkFromDomainDiggity, CodeSearchDiggity, DLPDiggity, FlashDiggity, MalwareDiggity, PortScanDiggity, SHODANDiggity, BingBinaryMalwareSearch, and NotInMyBackYardDiggity.
- **Publicly available scans of IP Addresses**
- [Exfiltrated](#) provides the scans from the 2012 Internet Census. It would provide the IP address and the port number running at the time of scan in the year 2012.
- [Shodan](#): provides the same results may be with recent scans. You need to be logged-in. ShodanCLI is available at [ShodanCommand-Line Interface](#)

# PASSIVE FINGERPRINTING

Reverse IP Lookup by Domaintools: Domain name search tool that allows a wildcard search, monitoring of WHOIS record changes and history caching, as well as Reverse IP queries.

Passive Total: A threat-analysis platform created for analysts, by analysts.

- Server Sniff: A website providing IP Lookup, Reverse IP services.

- Robtex: Robtex is one of the world's largest network tools. At robtex.com, you will find everything you need to know about domains, DNS, IP, Routes, Autonomous Systems, etc. There's a nmap script http-robtex-reverse-ip which can be used to find the domain/ website hosted on that ip.

- Nmap Command

```
nmap--script http-robtex-reverse-ip--script-args http-robtex-reverse-ip.host='XX.XX.78.214'
```

# INFRA MAPPING

**Ping Gateway IP Addresses : `nmap -sn -v -PE 192.168.*.1`**

- **DNS Enumeration** :If you are connected to a internal dnsserver, you may query it with `dig -t any domainname`
- Reverse DNS Lookup : `nmap-sLip`
- **Identifying Alive IP Addresses : `nmap-sn-n ip`**

•

# RECON TOOLS

- **Aquatone: A tool for domain flyovers**
- The Datasploit tool performs various OSINT techniques, aggregates all the raw data, and returns the gathered data in multiple formats.
- SpiderFoot is an open source intelligence automation tool. Its goal is to automate the process of gathering intelligence about a given target, which may be an IP address, domain name, hostname or network subnet. SpiderFoot can be used offensively, i.e. as part of a black-box penetration test to gather information about the target or defensively to identify what information your organization is freely providing for attackers to use against you.
- Intrigue makes it easy to discover information about the attack surface connected to the Internet. Intrigue utilizes common OSINT sources via “tasks” to create “entities”. Each discovered entity can be used to discover more information, either automatically or manually.

Additional Read : <https://threat.tevora.com/apache-and-java-information-disclosures-lead-to-shells/>





THANK YOU