SEMINAR 1
TOOLS FOR INTERNET RESEARCH

Getting information about machines, people and…source code!

JOSÉ MANUEL REDONDO LÓPEZ *"S-81 ISAAC PERAL"* PROJECT V4.0

Source: Bing Chat AI

Departamento de Informática
Universidad de Oviedo

Universidad de Oviedo

Logo: @creative_vanesa (*Instagram*)

INTRODUCTION

- **You may wonder why all current focus on security these days…**
  - Even with no technical knowledge, you can have a lot of information about people / companies that can be exploited to obtain advantages
  - All is over the Internet for you to pick up…if you know where (and how) to look
  - This seminar will show you tools explaining how to do it
  - If you are serious about uses (and misuses) of the information over the Internet, you should enroll in the optional course Web Information Systems (SIW)
    - If you contact Daniel Gayo, he can give you access to the course materials and videos for you to check!
- **Let's do something that does not require a lot of technical skill ☺**
  - **OSINT (Open-Source INTelligence)**
  - Some of these applications will be used in laboratories
  - But, above all, we will exploit this branch more in the **Group Work** of the course
- **NOTE: If, using these tools, you find yourself exposing too many information on Internet ("egosurfing"), there are ways of recovering your privacy**
  - You can follow this guide: https://open.nytimes.com/how-to-dox-yourself-on-the-internet-d2892b4c5954?gi=57af4dc583d

INDEX

< Go to Index

OSINT AGAINST THE MACHINE

Investigating machines using only public information

Departamento de Informática
Universidad de Oviedo

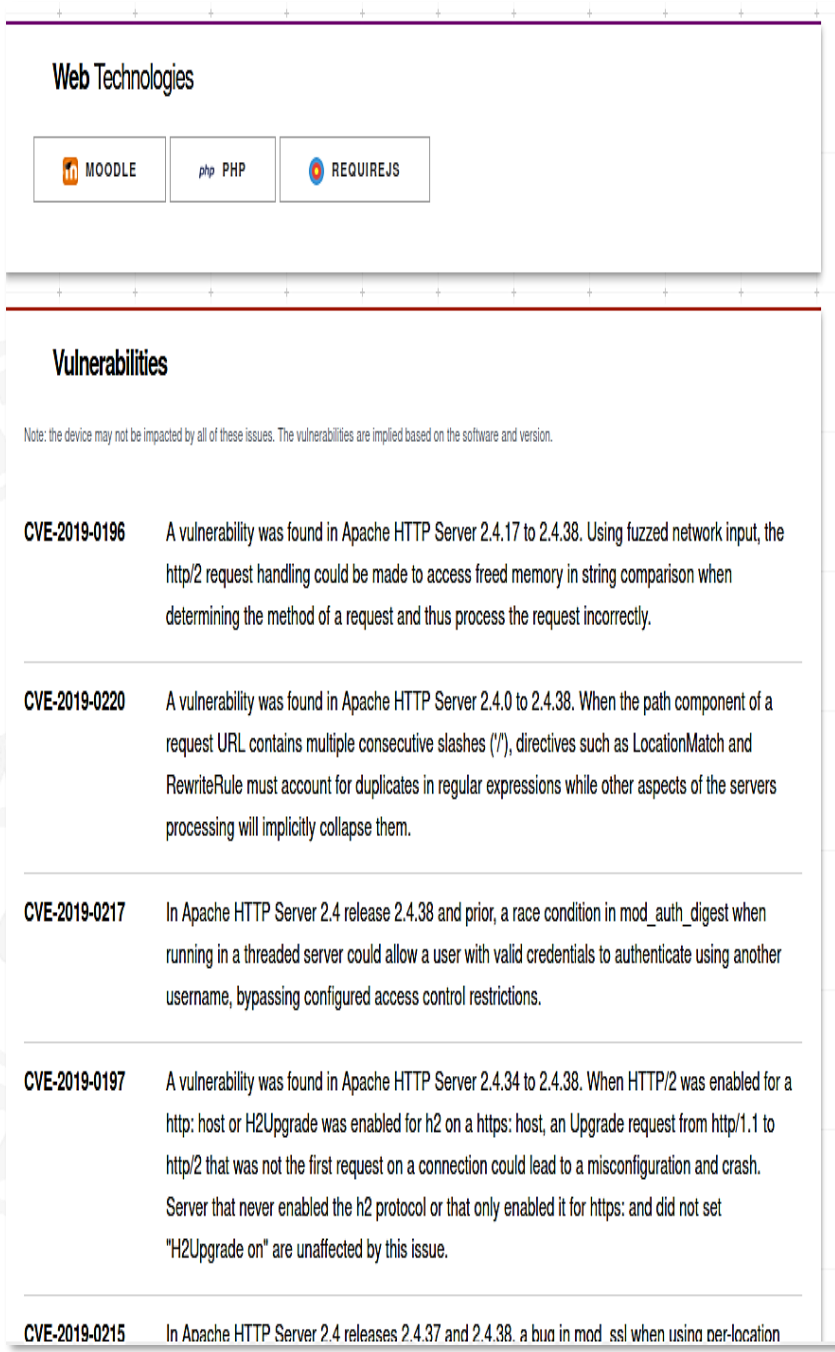
Universidad de Oviedo

SHODAN

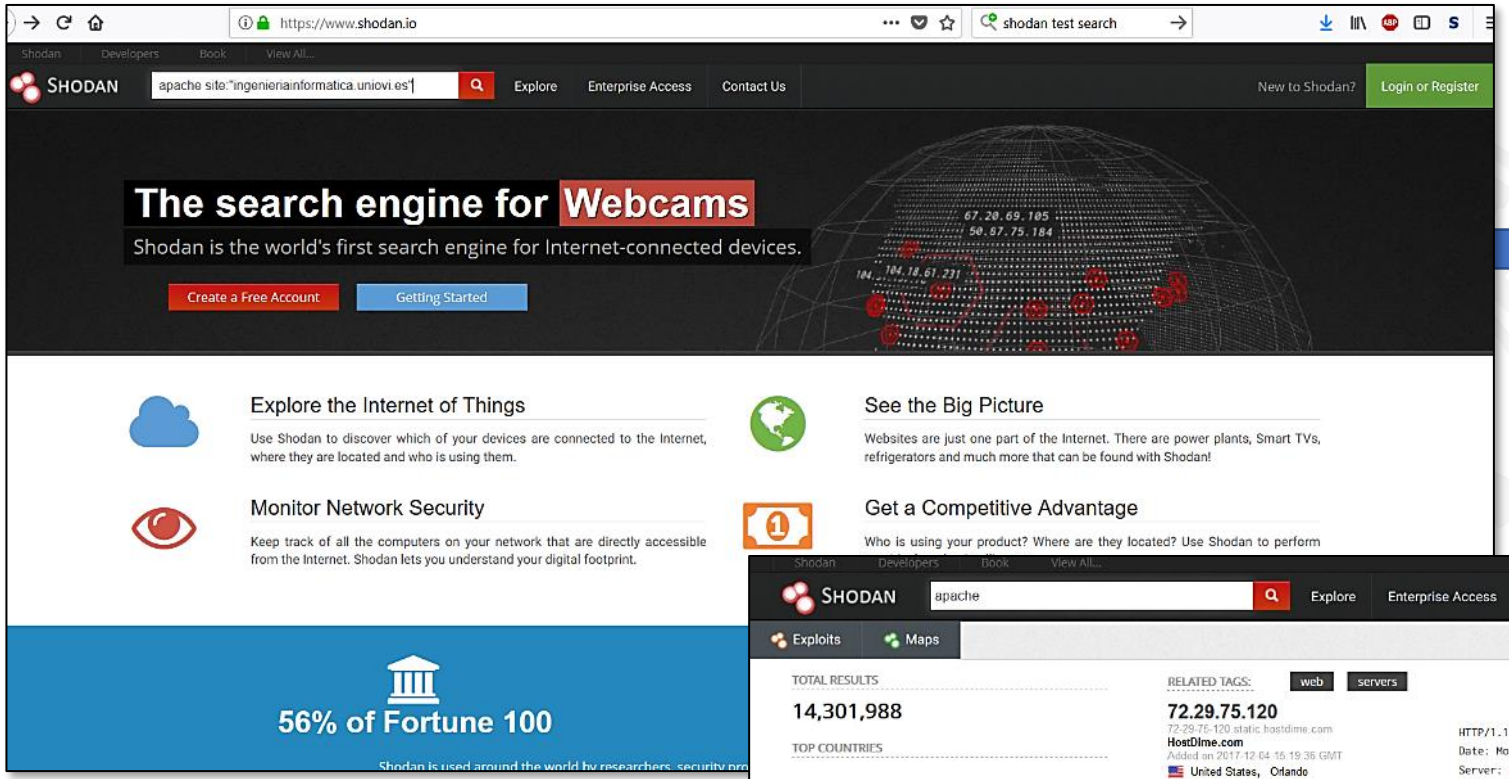
http://www.shodanhq.com/

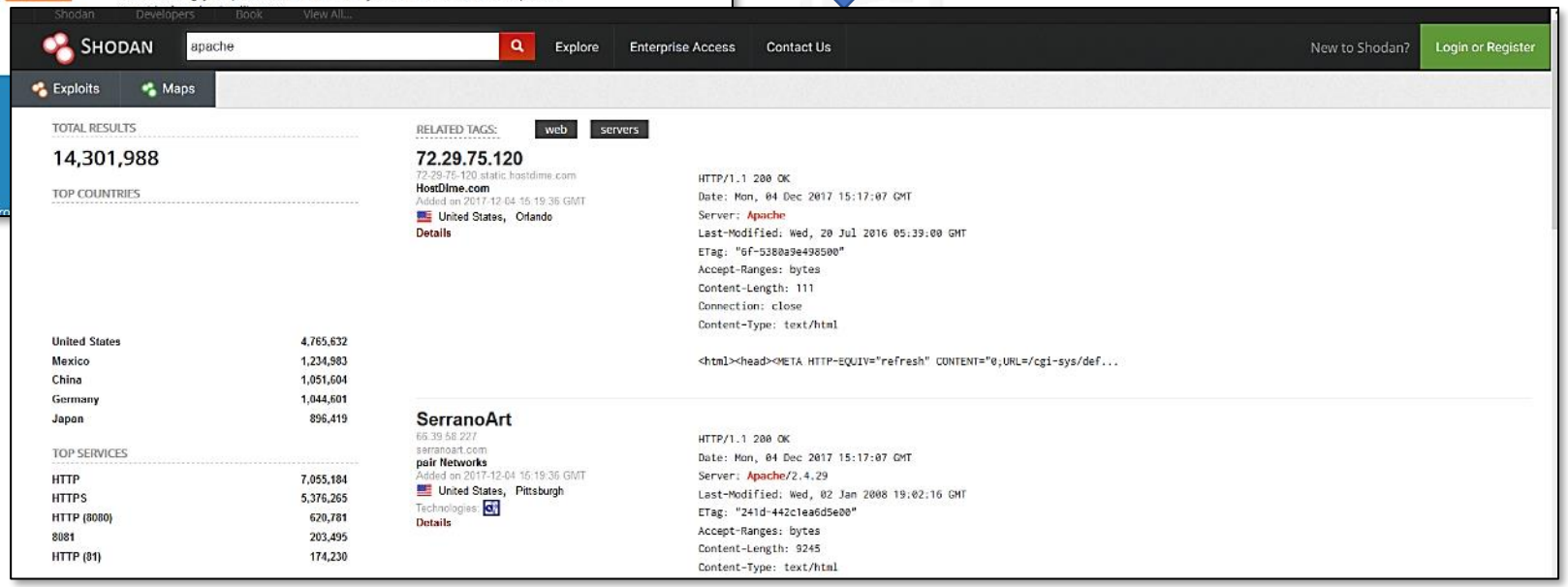
Computer Security

- **Machine (not web) search engine**
  - Routers, servers, cameras... (IoT)
  - Reports the type of service found, version, etc.
  - Finds SCADA (Supervisory Control And Data Acquisition)
    - Industrial engineering devices
- **Search for up to 10 devices**
  - Free registration allows more and using logical operators
  - Used to view devices exposed to the Internet and if they pose a danger
    - They have management software that can be accessed, and attacked
- **Tutorials**
  - https://danielmiessler.com/study/shodan/
  - https://hacking-etico.com/2016/02/12/4979/



SHODAN





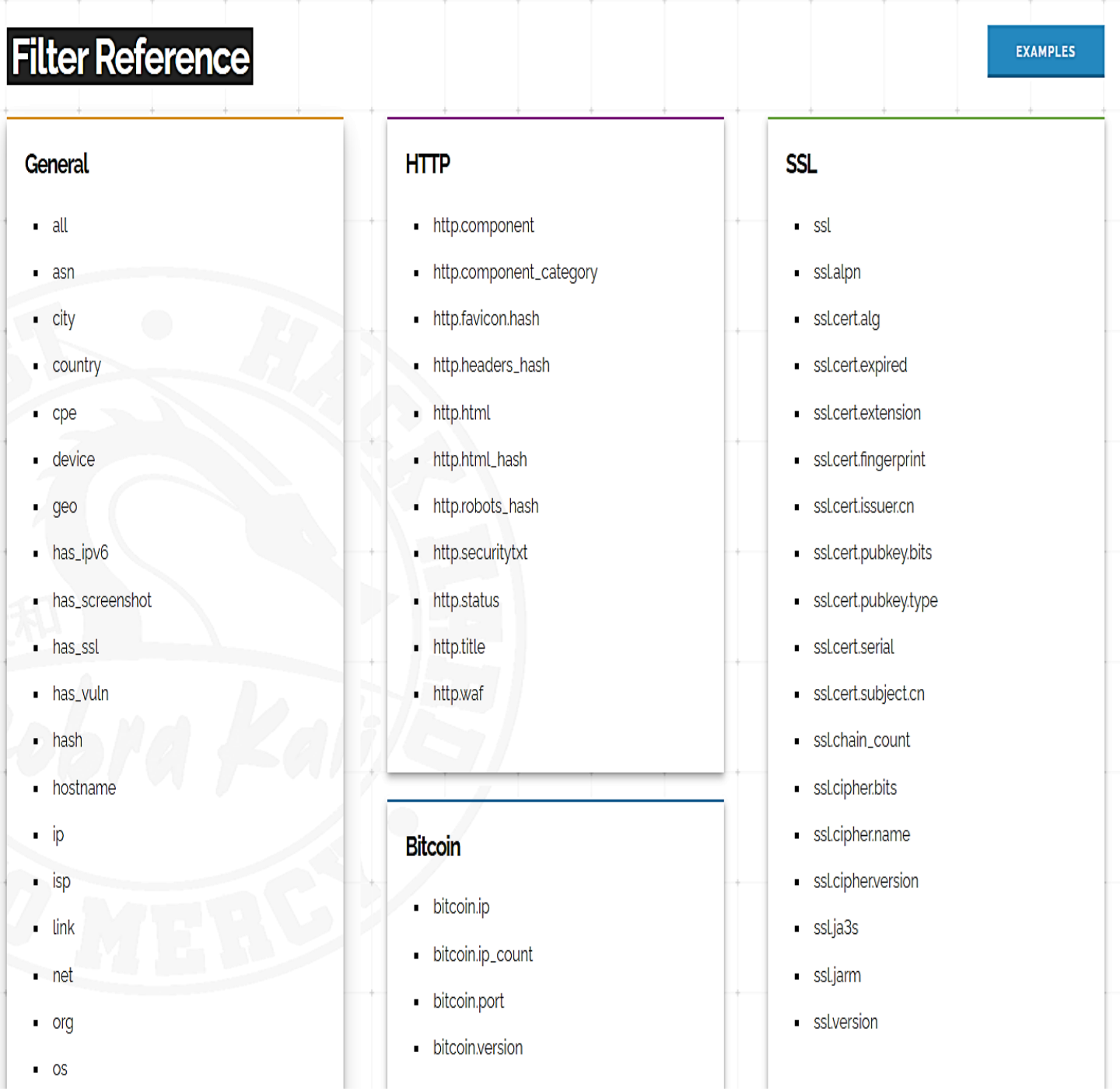
Search these service's banners

- HTTP/HTTPS
- FTP
- SSH
- Telnet
- SNMP
- SIP (IP telephony)
- RTSP (video *streaming* and webcams)
- …

SHODAN

Computer Security

- **The best part of having a Shodan account is the power of its search filter**
  - It is like a Google for machines!
- **You can search machines using an enormous amount of criteria**
- **Including advanced HTTP and SSL features**
  - That is, specially against web servers and remote accessible machines over Internet

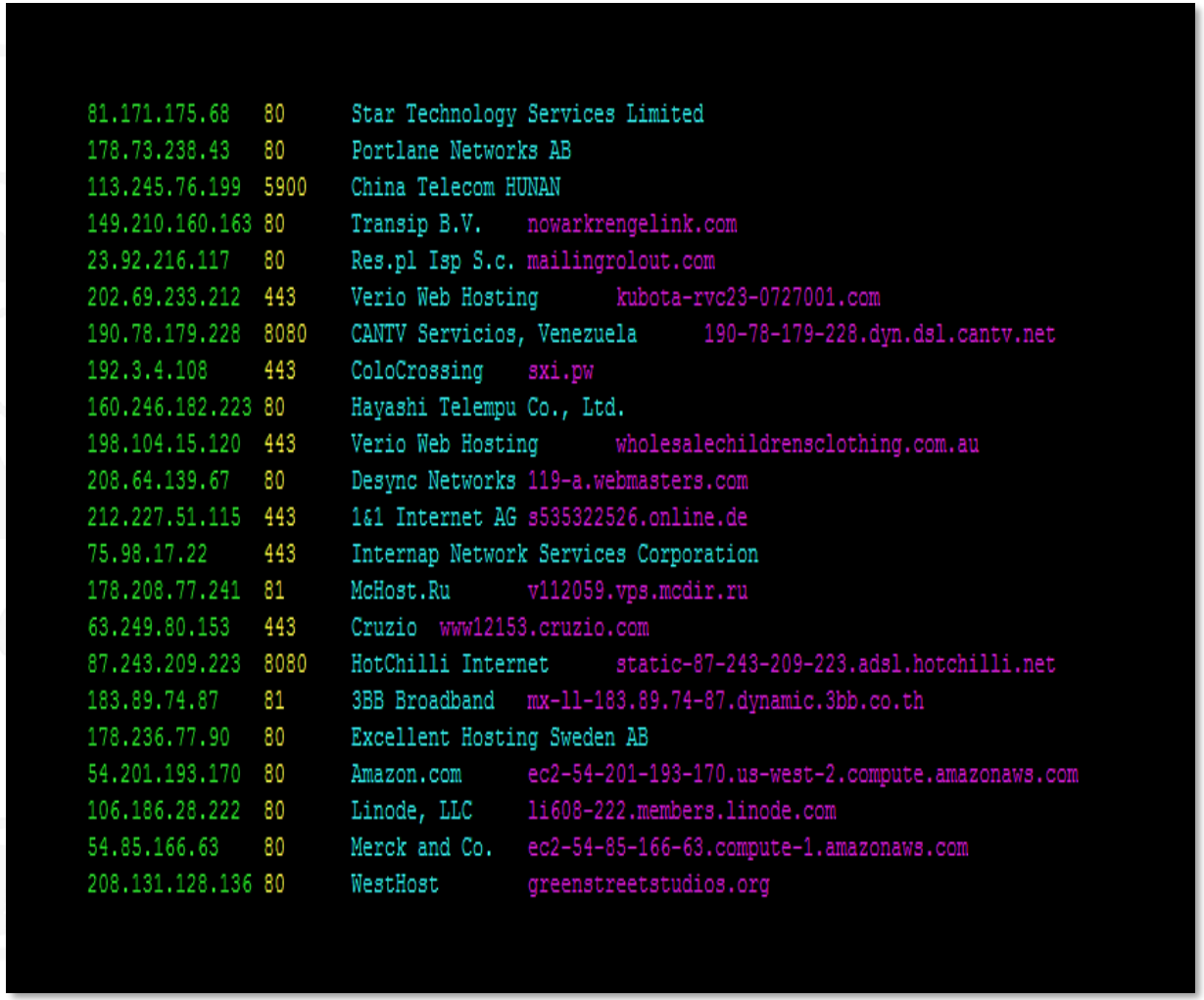


SHODAN CLI

https://cli.shodan.io/

Computer Security

- **Python library providing Shodan's capabilities from the command line**
  - Useful in GUI-less contexts
- **Allows to integrate Shodan's features into our own scripts or programs**
  - Or automate pentesting tasks including Shodan
  - The official website has usage examples
- **Its capabilities are equivalent to the web interface**
- **Set of useful searches**
  - https://github.com/jakejarvis/awesome-shodan-queries

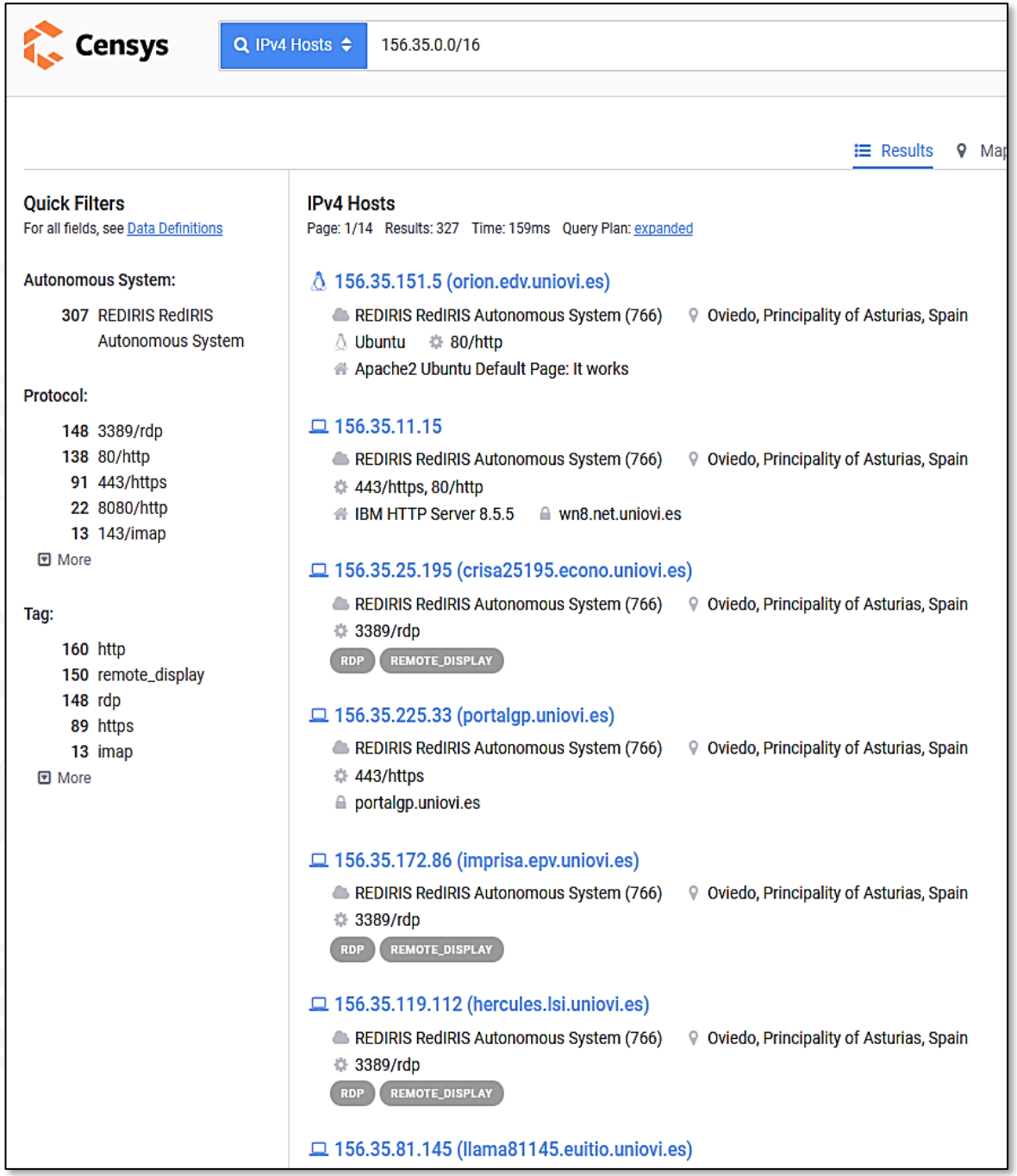


CENSYS

https://censys.io/

Computer Security

- **Like Shodan, but with structured info**
  - Searches server features: interesting running services (RDP), OS/web server banners…
  - Multiple search criteria
    - IP (v4 only) and networks: https://censys.io/ipv4/
    - Website domain names: https://censys.io/domain?q=
    - Certificates: https://censys.io/certificates?q=
- **Search tutorial**
  - https://developerinsider.co/censys-find-and-analyze-any-server-and-device-on-the-internet/
- **A similar tool is Onhype.io:**
  - https://fwhibbit.es/descubrimiento-y-recopilacion-de-activos-con-onyphe-io
  - Comparison with other tools (ZMap, Mr Looquer)
    - https://www.dragonjar.org/shodan-vs-scans-io-vs-censys-io-vs-zmap-vs-mr-looquer.xhtml

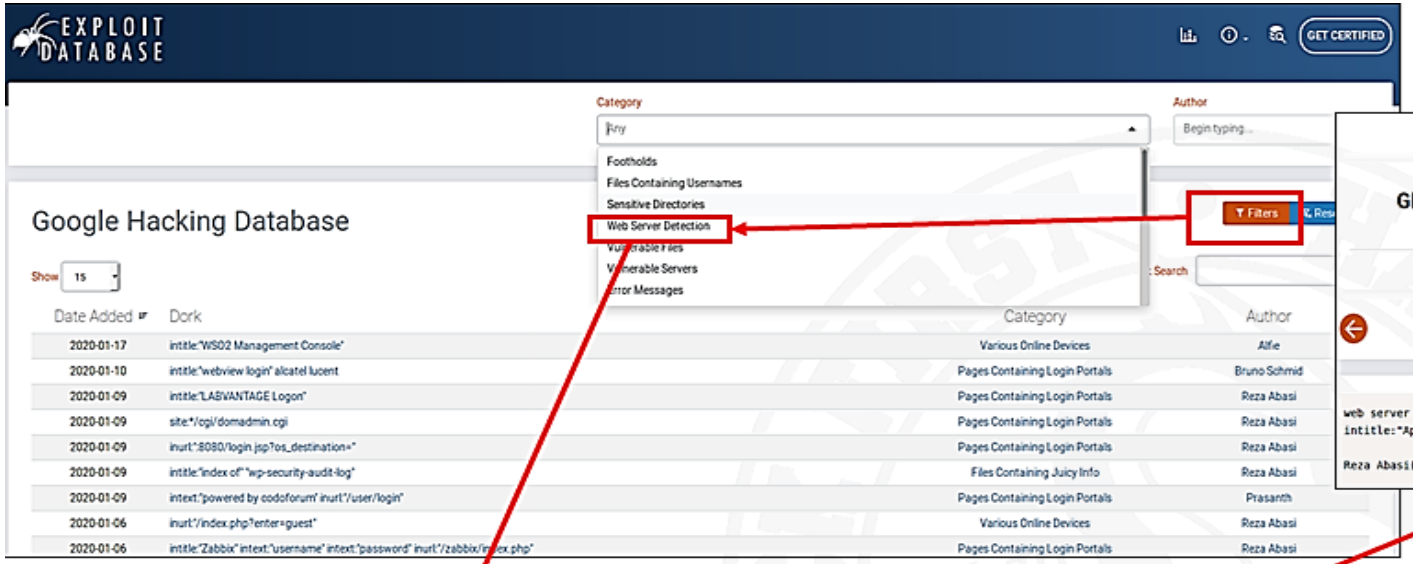


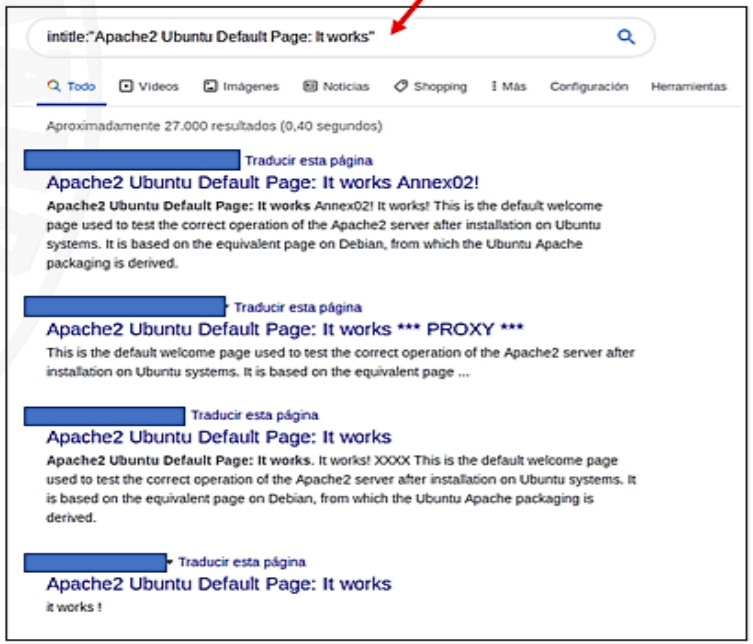
Using Shodan to find information

Google hacking or "Google dorking"

- **Ability of doing certain searches that reveal compromising information**
  - Using standard search engines (normally targeting a concrete domain with the `site:` operator)
  - These queries find information leading to vulnerabilities, configuration issues, or facilitate an attack
  - It can be done in other search engines too (although Google is the most used one)
- **There is a database of organized prebuilt searches: Google Hacking Database**
  - These are called "Google Dorks" : https://www.exploit-db.com/google-hacking-database
  - Choose the right search category: each one represents a possible vulnerability!
  - Search the target with the chosen query (`site:`)
- **"Dorking" can also be used with other products**
  - https://github.com/cipher387/Dorks-collections-list
- **Examples**
  - https://wifibit.com/google-hacking/
  - https://ciberpatrulla.com/osint-con-google/
  - https://ciberpatrulla.com/buscar-google/
  - https://medium.com/@logicbomb_1/one-misconfig-jira-to-leak-them-all-including-nasa-and-hundreds-of-fortune-500-companies-a70957ef03c7

GOOGLE HACKING: PROCESS





Final Recon

https://github.com/thewhiteh4t/FinalRecon

- Python script for web reconnaissance
- Implements a lot of features to create a full profile of a web page server and its contents
  - DNS and WHOIS information
  - Full crawler
  - Traceroute
  - Directory search
  - Port scanning the server
- Tutorial:
  - https://hakin9.org/final-recon-osint-tool-for-all-in-one-web- reconnaissance/

```
python3 finalrecon.py -h

usage: finalrecon.py [-h] [--headers] [--sslinfo] [--whois] [--crawl] [--dns] [--
sub] [--trace] [--dir] [--ps] [--full] [-t T] [-T T] [-w W] [-r] [-s] [-d D] [-e E]
[-m M] [-p P] [-tt TT] [-o O] url

FinalRecon - The Last Recon Tool You Will Need | v1.0.7

positional arguments:
  url         Target URL

optional arguments:
  -h, --help  show this help message and exit
  --headers   Header Information
  --sslinfo   SSL Certificate Information
  --whois     Whois Lookup
  --crawl     Crawl Target
  --dns       DNS Enumeration
  --sub       Sub-Domain Enumeration
  --trace     Traceroute
  --dir       Directory Search
  --ps        Fast Port Scan
  --full      Full Recon

Extra Options:
  -t T        Number of Threads [ Default : 30 ]
  -T T        Request Timeout [ Default : 30.0 ]
  -w W        Path to Wordlist [ Default : wordlists/dirb_common.txt ]
  -r          Allow Redirect [ Default : False ]
  -s          Toggle SSL Verification [ Default : True ]
  -d D        Custom DNS Servers [ Default : 1.1.1.1 ]
  -e E        File Extensions [ Example : txt, xml, php ]
  -m M        Traceroute Mode [ Default : UDP ] [ Available : TCP, ICMP ]
  -p P        Port for Traceroute [ Default : 80 / 33434 ]
  -tt TT      Traceroute Timeout [ Default : 1.0 ]
  -o O        Export Output [ Default : txt ] [ Available : xml, csv ]
```

GooFuzz

https://github.com/m3n0sd0n4ld/GooFuzz

- **Bash script that uses Google Hacking to obtain various types of sensitive information**
  - Without making queries to the concrete web server
  - Files or directories, subdomains…
- **Tutorial**
  - https://www.youtube.com/watch?v=DY8k-WEravY

```
> ./GooFuzz -t nasa.gov -e pdf,bak,old -d 10
************************************************************
* GooFuzz 1.2.2 - The Power of Google Dorks                 *
************************************************************

Target: nasa.gov

================================================================
Extension: pdf
================================================================

https://history.nasa.gov/alsj/a11/A11_PressKit.pdf
https://history.nasa.gov/alsj/a13/A13_PressKit.pdf
https://history.nasa.gov/alsj/a14/A14_PressKit.pdf
https://history.nasa.gov/alsj/a15/A15_PressKit.pdf
https://history.nasa.gov/alsj/a410/A07_PressKit.pdf
https://history.nasa.gov/alsj/a410/A09_PressKit.pdf
https://history.nasa.gov/monograph15.pdf
https://www.hq.nasa.gov/alsj/LLRV_Monograph.pdf

================================================================
Extension: bak
================================================================

https://bocachica.arc.nasa.gov/ATTREX_2013/rh/rh_omega_oct3.html.bak
https://bocachica.arc.nasa.gov/ATTREX_2013/rh/rh_omega_sep30.html.bak
https://bocachica.arc.nasa.gov/ATTREX_2014/attrex2014_satmap.html.bak
https://bocachica.arc.nasa.gov/ATTREX_2014/schom/schomfigures_20140126.html.bak
https://bocachica.arc.nasa.gov/ATTREX_2014/schom/schomfigures_20140226.html.bak
https://bocachica.arc.nasa.gov/ATTREX_2014/trajfigures/trajfigures.html.bak
https://bocachica.arc.nasa.gov/POSIDON/posidon.html.bak
https://hesperia.gsfc.nasa.gov/hessi/solar_install/installation.bak
https://ndeaa.jpl.nasa.gov/nasa-nde/outreach/outreach.html.bak
https://ndeaa.jpl.nasa.gov/nasa-nde/yosi/yosi.htm.bak

================================================================
Extension: old
================================================================

https://bocachica.arc.nasa.gov/HAVE/nmc_rh_omega_plots.html.old
https://echo.jpl.nasa.gov/asteroids/1988TA/1988TA_planning.html.old
https://echo.jpl.nasa.gov/asteroids/1998QE2/1998QE2_planning.html.old
https://echo.jpl.nasa.gov/asteroids/2014JO25/2014JO25_planning.html.old
https://echo.jpl.nasa.gov/asteroids/goldstone_asteroid_schedule.html.old
https://fits.gsfc.nasa.gov/users_guide/users_guide.ps.old
https://image.msfc.nasa.gov/ChrisDocs/UDFInstall/uLibInstall.old
https://seawifs.gsfc.nasa.gov/OCEAN_PLANET/HTML/titanic.html.old
https://umbra.nascom.nasa.gov/sobo/bga_history.html.old
```

< Go to Index

Source: Bing Chat AI

Social Networks >

Data Integration >

OSINT AGAINST THE PEOPLE

Investigating people using public information only

Departamento de Informática
Universidad de Oviedo

Universidad de Oviedo

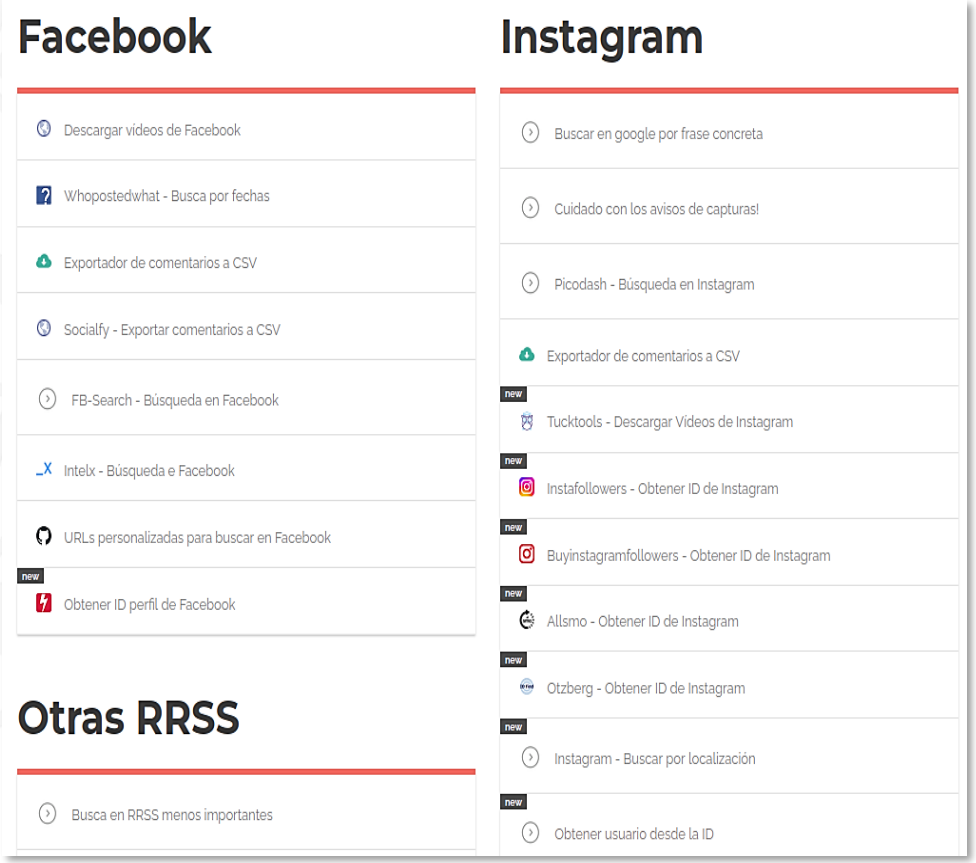
Social networks

Know what the people do...and post ☺

THE PROBLEM OF TOOLS THAT WORK WITH SOCIAL MEDIA

- **Social networks don't like tools that extract information from their profiles** ☺
  - And that gets us into a cat-and-mouse race
  - Forcing us to look for tools that keep working
- **You can find a lot of them in Ciberpatrulla**
- **Twitter (now X) is a perfect example of this**
  - Access to its API became non-free, which killed many research tools
  - This left (some) of the scraping-based ones
    - That is, **processing the content that is displayed on the screen**, acting as if they were normal users
  - Some are (but don't rule out that they stop working)
    - https://github.com/tweepy/tweepy
    - https://github.com/dataquestio/twitter-scrape

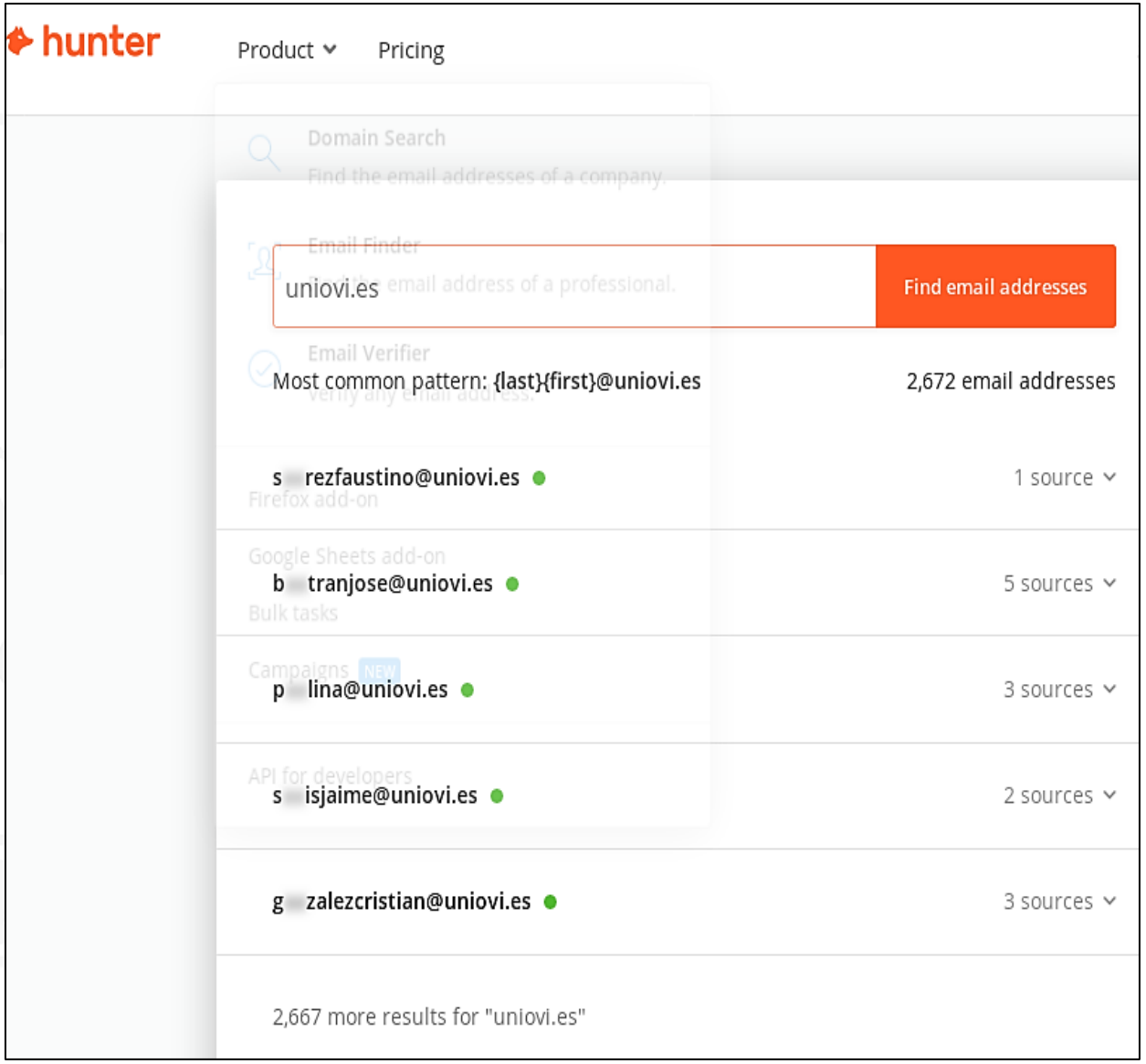
Ciberpatrulla is a huge repository of OSINT tools, among which there are many that work on social networks: https://ciberpatrulla.com/links/



HUNTER.IO

https://hunter.io/

- **Website that allows you to discover email addresses related to domains / companies**
- **With them, OSINT attacks, phishing attempts or APTs (Advanced Persistent Threats) can be started**
- **It is not a free service**
  - Free registration allows limited usage of its features
  - Complete results
  - CSV downloads
  - Up to 50 searches/month



Project Sherlock

https://github.com/sherlock-project/sherlock

- **Finds a username on lots of different social networks and webpages**
  - More than 320!
- **It is very common to find users with the same name on different social networks**
  - We can use it to collect information from different sources about the same person
  - This may reveal data about people with closed profiles only in a network, but an open profile in another one
- **Basic usage is simple**
  - `python3 pherlock.py <username>`

```
┌──(vagrant㉿kali)-[~]
└─$ sherlock IbaiLlanos
[*] Checking username IbaiLlanos on:
[+] Academia.edu: https://independent.academia.edu/IbaiLlanos
[+] CapFriendly: https://www.capfriendly.com/users/IbaiLlanos
[+] ChaturBate: https://chaturbate.com/IbaiLlanos
[+] Coil: https://coil.com/u/IbaiLlanos
[+] DeviantART: https://IbaiLlanos.deviantart.com
[+] Duolingo: https://www.duolingo.com/profile/IbaiLlanos
[+] F3.cool: https://f3.cool/IbaiLlanos/
[+] Facebook: https://www.facebook.com/IbaiLlanos
[+] Fiverr: https://www.fiverr.com/IbaiLlanos
[+] FortniteTracker: https://fortnitetracker.com/profile/all/IbaiLlanos
[+] Giphy: https://giphy.com/IbaiLlanos
[+] GitHub: https://www.github.com/IbaiLlanos
[+] Gumroad: https://www.gumroad.com/IbaiLlanos
[+] Instagram: https://www.instagram.com/IbaiLlanos
[+] LeetCode: https://leetcode.com/IbaiLlanos
[+] Lichess: https://lichess.org/@/IbaiLlanos
[+] Linktree: https://linktr.ee/IbaiLlanos
[+] Minecraft: https://api.mojang.com/users/profiles/minecraft/IbaiLlanos
[+] PSNProfiles.com: https://psnprofiles.com/IbaiLlanos
[+] Pinterest: https://www.pinterest.com/IbaiLlanos/
[+] Pokemon Showdown: https://pokemonshowdown.com/users/IbaiLlanos
[+] Pornhub: https://pornhub.com/users/IbaiLlanos
[+] Quizlet: https://quizlet.com/IbaiLlanos
[+] Reddit: https://www.reddit.com/user/IbaiLlanos
[+] Roblox: https://www.roblox.com/user.aspx?username=IbaiLlanos
[+] RuneScape: https://apps.runescape.com/runemetrics/profile/profile?user=Ib
aiLlanos
[+] Scratch: https://scratch.mit.edu/users/IbaiLlanos
[+] SlideShare: https://slideshare.net/IbaiLlanos
[+] Smule: https://www.smule.com/IbaiLlanos
[+] Spotify: https://open.spotify.com/user/IbaiLlanos
[+] Telegram: https://t.me/IbaiLlanos
[+] Tenor: https://tenor.com/users/IbaiLlanos
[+] TikTok: https://tiktok.com/@IbaiLlanos
[+] TradingView: https://www.tradingview.com/u/IbaiLlanos/
[+] Trello: https://trello.com/IbaiLlanos
[+] Twitch: https://www.twitch.tv/IbaiLlanos
[+] Twitter: https://twitter.com/IbaiLlanos
[+] Venmo: https://venmo.com/u/IbaiLlanos
[+] Wattpad: https://www.wattpad.com/user/IbaiLlanos
[+] WordPress: https://IbaiLlanos.wordpress.com/
[+] Xbox Gamertag: https://xboxgamertag.com/search/IbaiLlanos
[+] YouNow: https://www.younow.com/IbaiLlanos/
[+] mercadolivre: https://www.mercadolivre.com.br/perfil/IbaiLlanos
[+] osu!: https://osu.ppy.sh/users/IbaiLlanos
[+] xHamster: https://xhamster.com/users/IbaiLlanos
```

User reconnaissance integration / from multiple sources

Aggregating user data

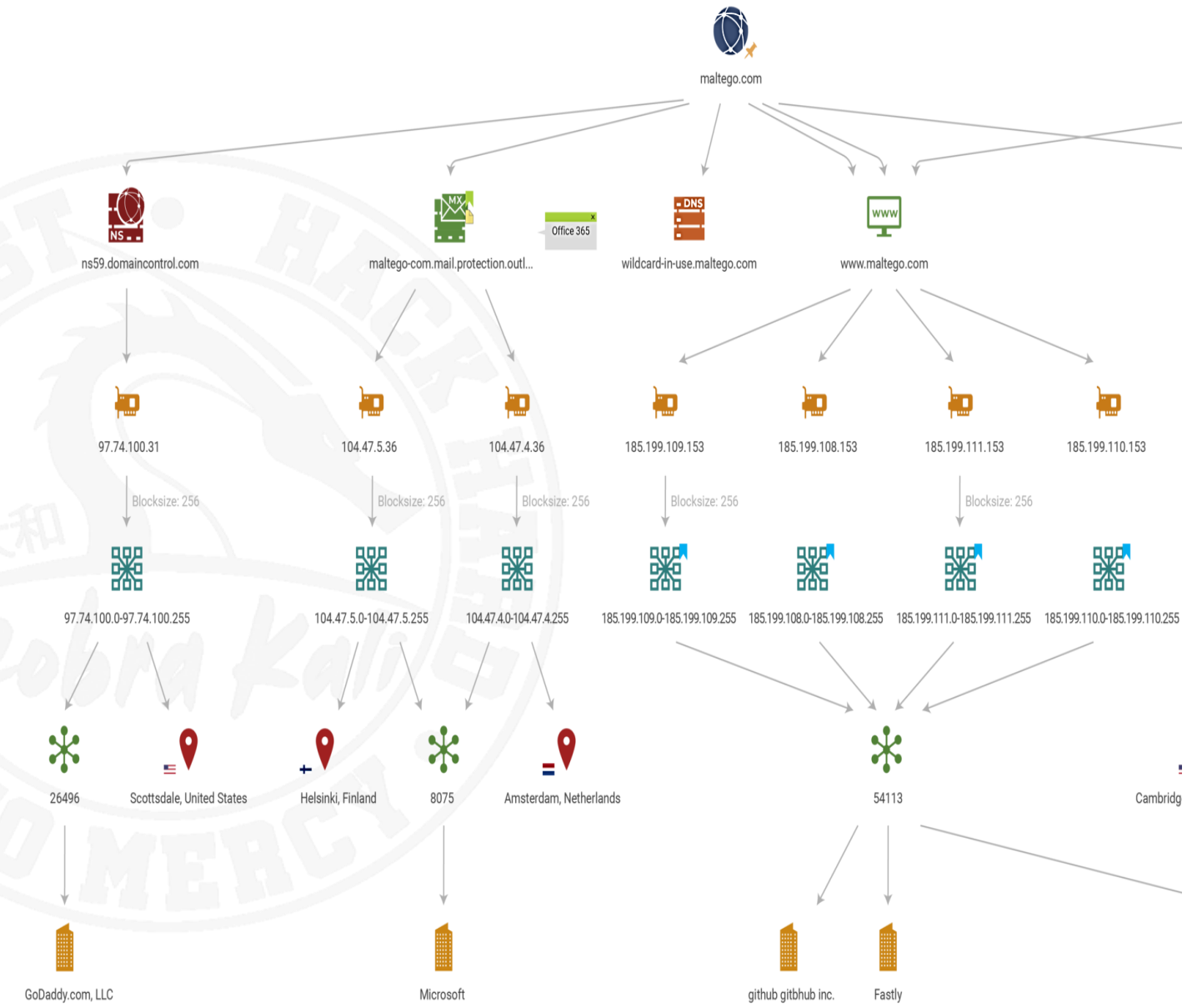
MALTEGO

https://www.paterva.com/buy/maltego-clients/maltego-ce.php

Computer Security

Source: https://en.wikipedia.org/wiki/Maltego

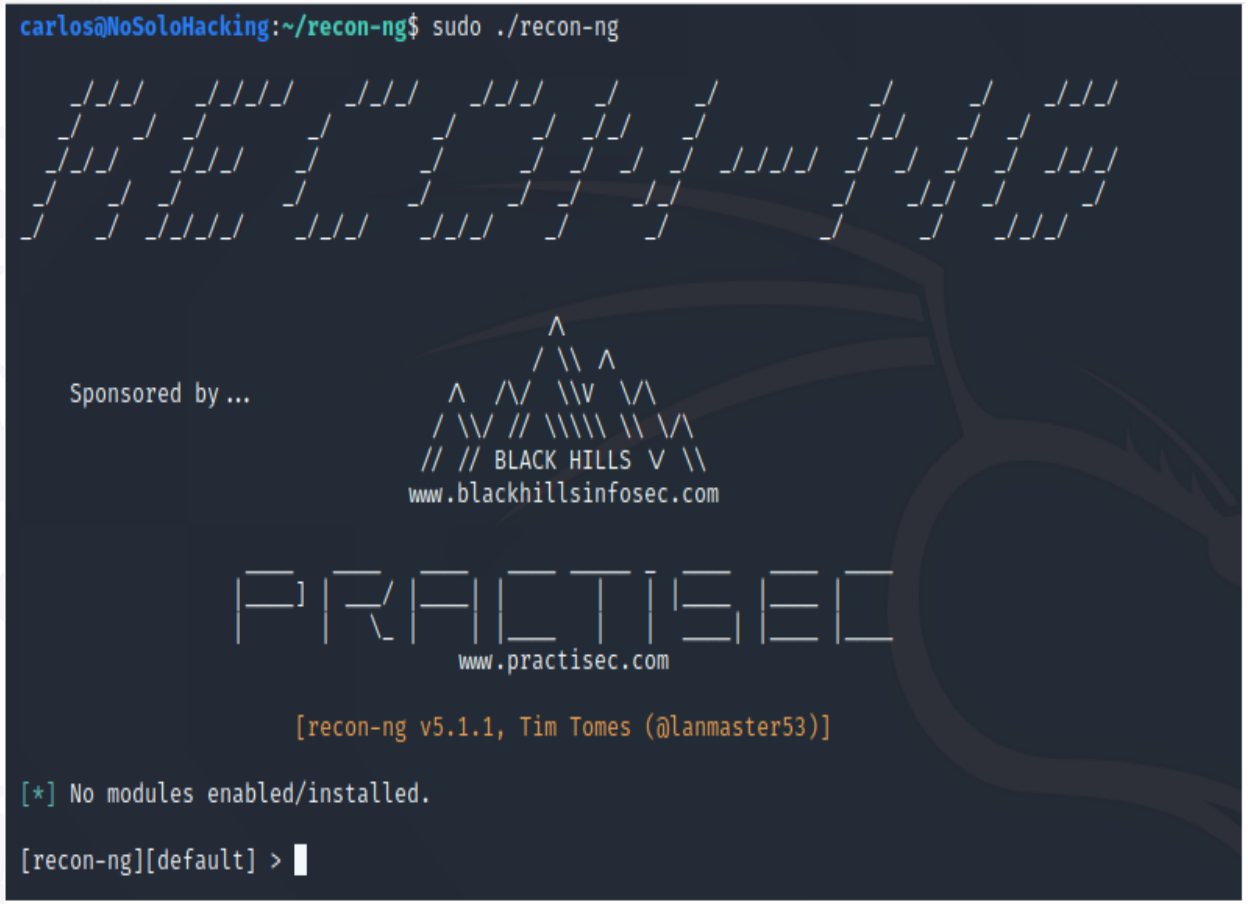
- **Collects information from Internet entities**
  - Creates custom entities that represent any kind of information
  - Specializes in finding relationships between them
    - Social networks, computer networks, domains...
  - Uses DNS and `whois` records, search engines, known social networks, APIs... as sources
- **Tutorials**
  - https://www.fwhibbit.es/introduccion-a-maltego
  - https://ciberpatrulla.com/maltego/
  - www.elladodelmal.com/2010/08/mineria-de-datos-con-maltego-1-de-2.html
  - We will see more about it on the **Group Works**



RECON-NG

https://github.com/lanmaster53/recon-ng

- Used for recollecting information about users from different sources
- It has modules to search on different sources of information
  - Search engines, APIs...
- Copies **Metasploit** UI to make it easier for people who are already familiar with it
- **Tutorials**
  - https://hackertarget.com/recon-ng-tutorial/
  - https://noticiasseguridad.com/tutoriales/recon-ng-herramienta-para-recoleccion-de-informacion/

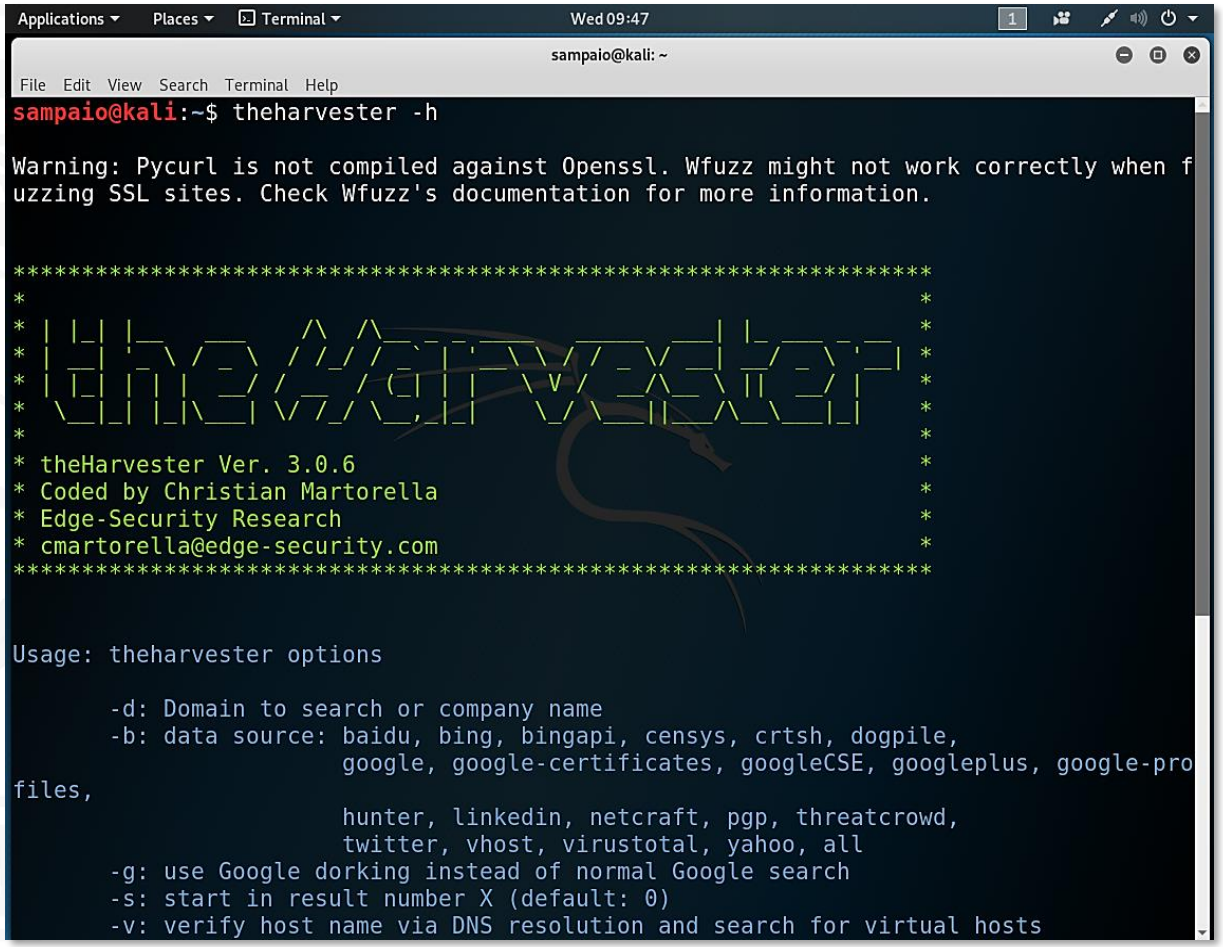


Source: https://www.nosolohacking.info/recon-ng-instalacion/

The Harvester

https://github.com/laramies/theHarvester

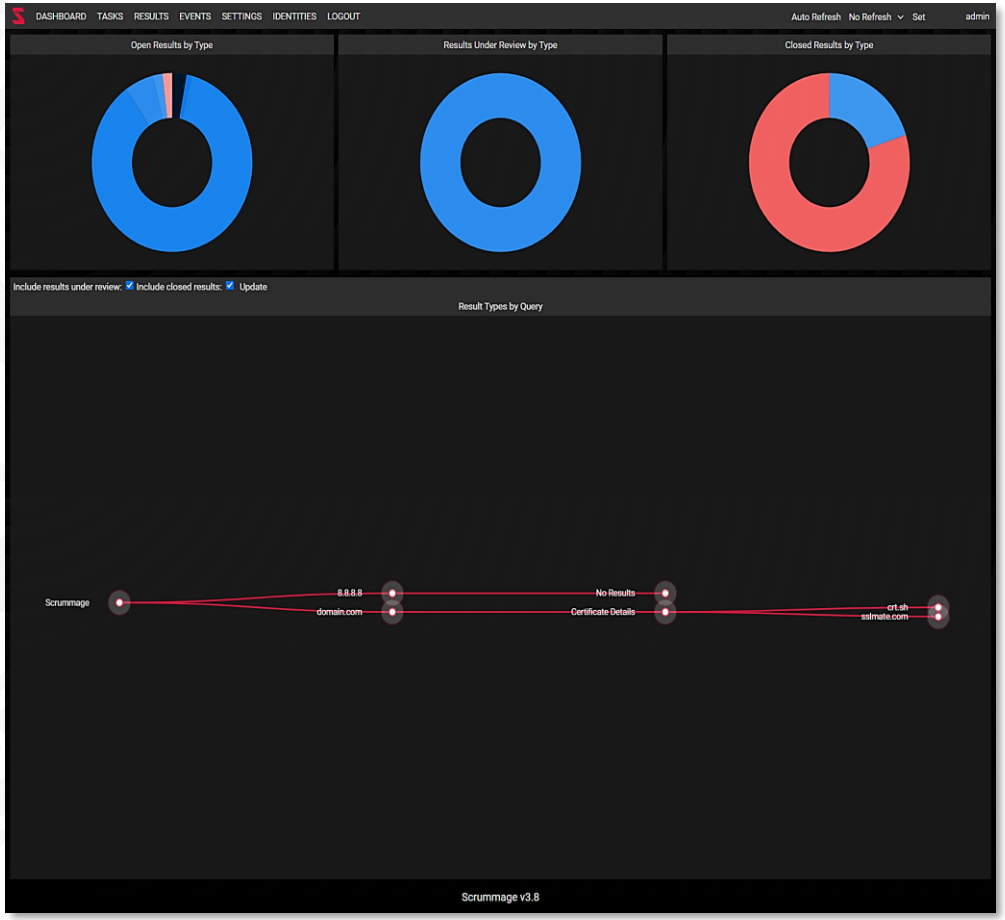
- **Obtains data of a certain target using many different sites as sources**
  - You get a lot of information from public sources: emails, names, subdomains, IPs, URLs...
  - Uses known search engines for different kinds of information
    - baidu, bing, bingapi, censys, Comodo Certificate search, dnsdumpster, dogpile, duckduckgo, github-code, Google (optional Google dorking), google-certificates, hunter, intelx, Linkedin, Netcraft Data Mining, securityTrails (DNS history information), Shodan, threatcrowd, trello, Twitter, Bing virtual hosts search, Virustotal, and Yahoo search engine
- **Also, information from DNS**
  - Brute force, reverse lookup, TLD expansion…
- **Tutorial**
  - https://www.welivesecurity.com/la-es/2015/04/08/the-harvester-riesgo-nformacion-publica/



SCRUMMAGE

https://github.com/matamorphosis/Scrummage

- **Centralizes OSINT scans**
  - Provides us a high-level view of sites with information about a target
- **Implements plugins to customize the type of scans and optimize resource usage**
- **Such plugins provide diverse capabilities**
  - Blockchain search
  - Domain fuzzing
  - Twitter analytics
  - Instagram search
  - Search on Have I Been Pwned? (https://haveibeenpwned.com/)
    - In case any account identities have been leaked …
- **Similar tools**
  - OSINT Framework: https://github.com/lockfale/osint-framework
  - Scumblr: https://portunreachable.com/automated-osint-with-scumblr-a4d81a048e54?gi=ae6ebac43a92

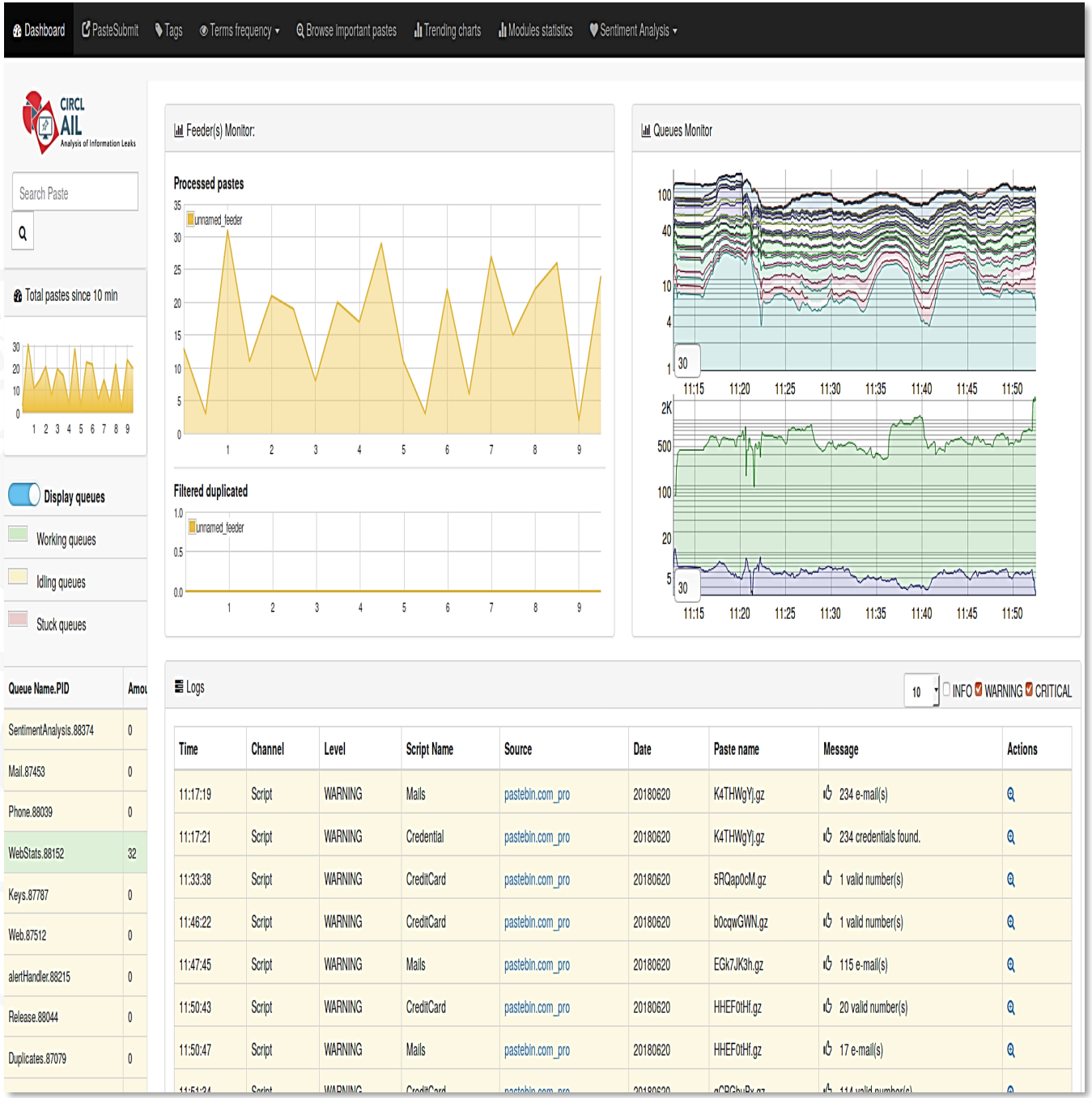


AIL Framework

https://github.com/CIRCL/AIL-framework

Computer Security

- **Highly modular framework that analyzes filtered information**
  - From unstructured sources
  - A popular source is **Pastebin**
- **Its highly modular so it could be extended**
  - And used to mine and process private information of different kinds
  - And therefore, detect and prevent information leakage
- **To find out more**
  - https://www.kitploit.com/2019/08/ail-framework-framework-for-analysis-of.html



More OSINT tools...

Computer Security

- **OSINT's techniques are very popular and powerful due to the information they can find**
  - There are a huge number of tools that get information from different media and social networks on Internet
  - It's very difficult to list them all
- **Remember: There's a classified list of tools available: https://ciberpatrulla.com/links/**
  - This web is probably the most complete repository on OSINT tools
  - It contains hundreds of tools classified for its basic usage in the OSINT world
  - Social networks, vehicles, phones, documents...

DO YOU HAVE EXAMPLES?

- With such a number of tools available, it is very easy to get lost when starting an OSINT research
- Therefore, it is best to see examples of how to make one
- You can follow this four-part tutorial
  - https://navisec.io/a-pentesters-guide-part-1-osint-passive-recon-and-discovery-of-assets/
  - https://navisec.io/a-pentesters-guide-part-2-osint-linkedin-is-not-just-for-jobs/
  - https://navisec.io/a-pentesters-guide-part-3-osint-breach-dumps-password-spraying/
  - https://navisec.io/a-pentesters-guide-part-4-grabbing-hashes-and-forging-external-footholds/
  - https://navisec.io/a-pentesters-guide-part-5-unmasking-wafs-and-finding-the-source/
- Or this OSINT guide
  - https://www.hackers-arise.com/osint
- We can find more tools in this web page
  - Metadata extraction, geolocation, subdomain extraction…
  - https://derechodelared.com/herramientas-osint-recopilatorio/
- Or in the material that we are going to give you for the Group Works ☺

Using Google Hacking to find Uniovi-bound resources

< Go to Index

Source: Bing Chat AI

OSINT AGAINST THE SOURCE

Source code can be an OSINT threat too

Departamento de Informática
Universidad de Oviedo

Universidad de Oviedo

GitGot

https://github.com/BishopFox/GitGot

Computer Security

- **Finding private information is, disgracefully, common in source code**
  - Connection strings, usernames, keys, API Keys (with which to request services on behalf of the application), DB server addresses, unprotected Amazon buckets (cloud storage)...
  - When code is uploaded to GitHub, this information is an OSINT target!
- **GitGot allows searching for sensitive data in code hosted on GitHub**
  - General searches, search for a particular purpose, or only for certain types of sensitive information
  - A similar tool, but specialized in finding API Keys, is GitHound: https://www.kitploit.com/2019/07/git-hound-find-exposed-keys-across.html?m=1

GitGraber

https://github.com/hisxo/gitGraber

Computer Security

- Tool to find sensitive data in GitHub repositories
- Includes data from various sources that may have been left in repository files
- Among other information sources it includes
  - Google
  - Amazon (AWS)
  - Paypal
  - Github
  - Facebook
  - Twitter
  - …

```
root@bugbounty# python3 gitGraber.py -k wordlists/keywords.txt -q "yahoo" -s

[i] Github query : https://api.github.com/search/code?q=yahoo access_key&sort=indexed&o=desc
[i] Status code : 200
[!] POSSIBLE AWS TOKEN FOUND (keyword used:yahoo)
[+] Commit date : 2019-09-10T13:40:08Z by [redacted]
[+] RAW URL : https://raw.githubusercontent.com/[redacted]/cmd_bash.md
[+] Token : [redacted]
[+] Repository URL : https://github.com/[redacted]
[i] Github query : https://api.github.com/search/code?q=yahoo access_token&sort=indexed&o=desc
[i] Status code : 200
[!] POSSIBLE GOOGLE_FIREBASE_OR_MAPS TOKEN FOUND (keyword used:yahoo)
[+] Commit date : 2019-09-11T20:12:28Z by [redacted]
[+] RAW URL : https://raw.githubusercontent.com/[redacted]/connectApp/[redacted]/app.js
[+] Token : [redacted]
[+] Repository URL : https://github.com/[redacted]
[!] POSSIBLE TWILIO TOKEN FOUND (keyword used:yahoo)
[+] Commit date : 2019-07-30T06:28:32Z by [redacted]
[+] RAW URL : https://raw.githubusercontent.com/[redacted]/project.html
[+] Token : [redacted]
[+] Repository URL : https://github.com/[redacted]
```

GitHub itself!

Computer Security

- The GitHub search feature can also be used to find private information on public repos
- It is a powerful search function that allows searching by multiple terms
  - Like Google Hacking (Google Dorks), **but for code**
  - In fact, many people calls this **GitHub Dorks**!
- The image shows several popular search examples

Self-evaluation achievement list: Seminar 1

?

- Understand the OSINT concept
- Understand what types of devices we can find on the internet
- Know what type of information you should not put into source code
- Know how to use Shodan and Censys OSINT tools to find machines of a concrete target
- Know how to use Google Hacking to find information of a concrete target
- Know multiple OSINT tools that work with personal information of individuals / companies
- Know tools to study source code in public repositories
- Know how to study a target using Google Hacking
- Know how to study a person on social media and the problem with the tools that do this type of activity

SEMINAR 1
TOOLS FOR INTERNET RESEARCH

Departamento de Informática
Universidad de Oviedo

Universidad de Oviedo