



# **RECONNAISSANCE (RECON)**

***With great  
knowledge, comes  
successful attacks!***

# INTELLIGENCE GATHERING

- What is it
- Why do it
- What is it not

**Open source intelligence (OSINT)** is a form of intelligence collection management that involves finding, selecting, and acquiring information from publicly available sources and analyzing it to produce actionable intelligence.

# OPEN SOURCE INTELLIGENCE (OSINT)

Simply, it's locating, and analyzing publicly (open) available sources of information.

Intelligence gathering process has a goal of producing current and relevant information that is valuable to either an attacker or competitor.

- *OSINT is not only web searching!*

# OPEN SOURCE INTELLIGENCE (OSINT)

Takes three forms:

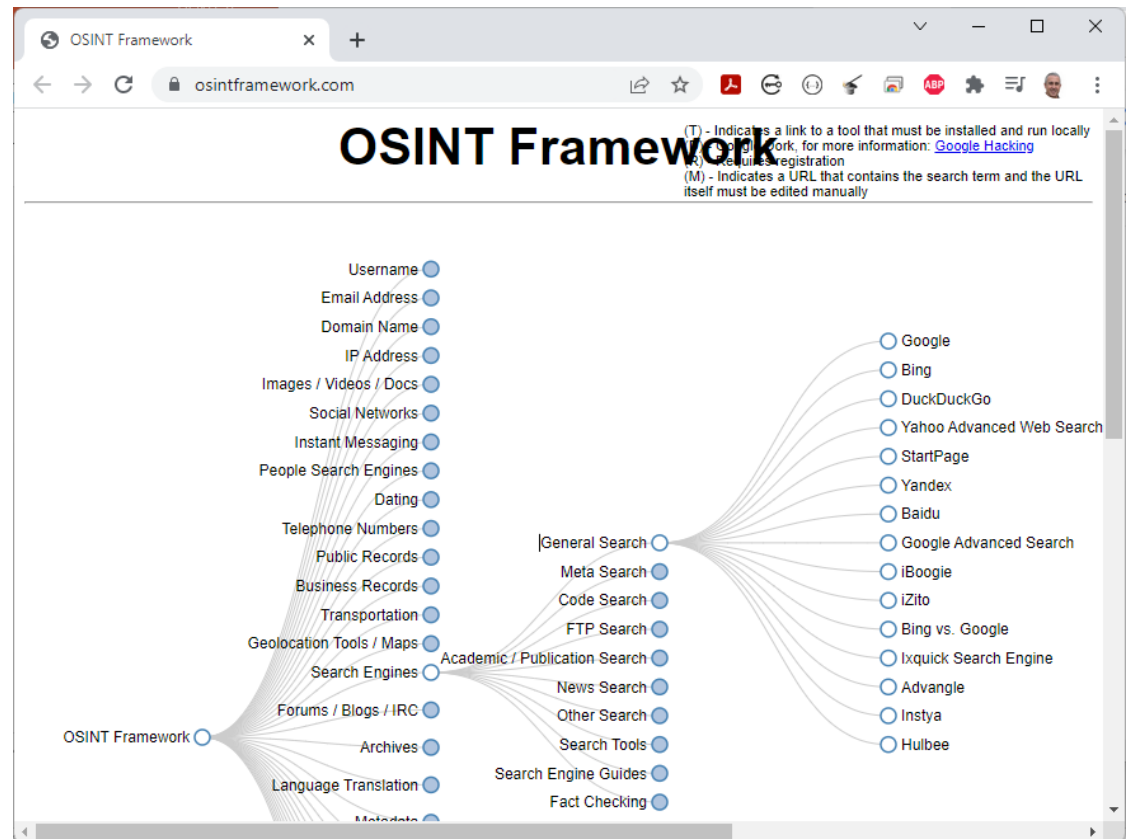
- Passive Information Gathering
- Semi-passive Information Gathering
- Active Information Gathering

Used for:

- Corporate
- Individuals

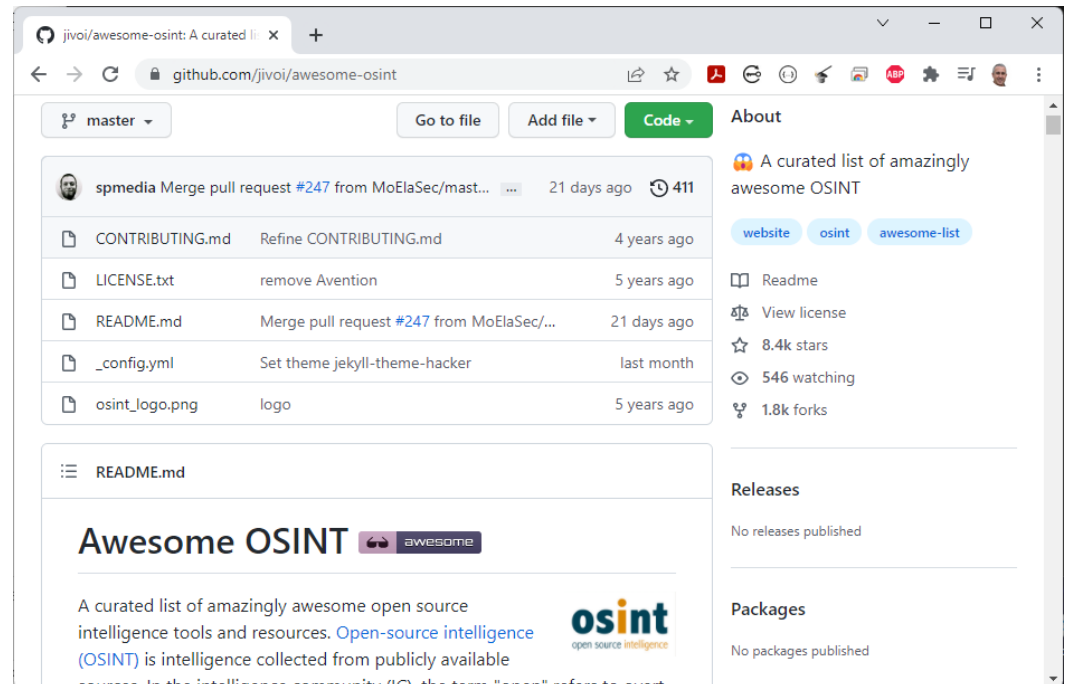
# OPEN SOURCE INTELLIGENCE (OSINT)

<https://osintframework.com>



# OPEN SOURCE INTELLIGENCE (OSINT)

<https://github.com/jivoi/awesome-osint>



# TARGET SELECTION

- Identification and Naming of Target
- Consider any Rules of Engagement limitations
- Consider **time length** for test
- Consider **end goal** of the test

# CORPORATE - PHYSICAL

## Locations

- Public sites can often be located by using search engines such as: Google, Yahoo, Bing, Ask.com, Baidu, etc.

## Relationships



# CORPORATE - LOGICAL

Business Partners

Business Clients

Competitors

Product line

Market Vertical

Marketing accounts

Meetings

Significant company dates

Job openings

Charity affiliations

Court records

Political donations

Professional licenses or registries

# JOB OPENINGS WEBSITES

- Monster, <https://www.monster.com>
- LinkedIn, <https://www.linkedin.com/mynetwork>,
- Viadeo (France), <https://fr.viadeo.com/fr>
- ...

# **CORPORATE – ORG. CHART**

Position identification

Transactions

Affiliates



# **CORPORATE – ELECTRONIC**

Document Metadata

Marketing Communications

# **CORPORATE – INFRASTRUCTURE ASSETS**

Network blocks owned

Email addresses

External infrastructure profile

Technologies used

Purchase agreements

Remote access

Application usage

Defense technologies

Human capability

# CORPORATE – FINANCIAL

Reporting

Market analysis

Trade capital

Value history

# INDIVIDUAL - HISTORY

Court Records

Political Donations

Professional licenses or registries

# INDIVIDUAL - SOCIAL NETWORK PROFILE(S)

Metadata Leakage

Location awareness

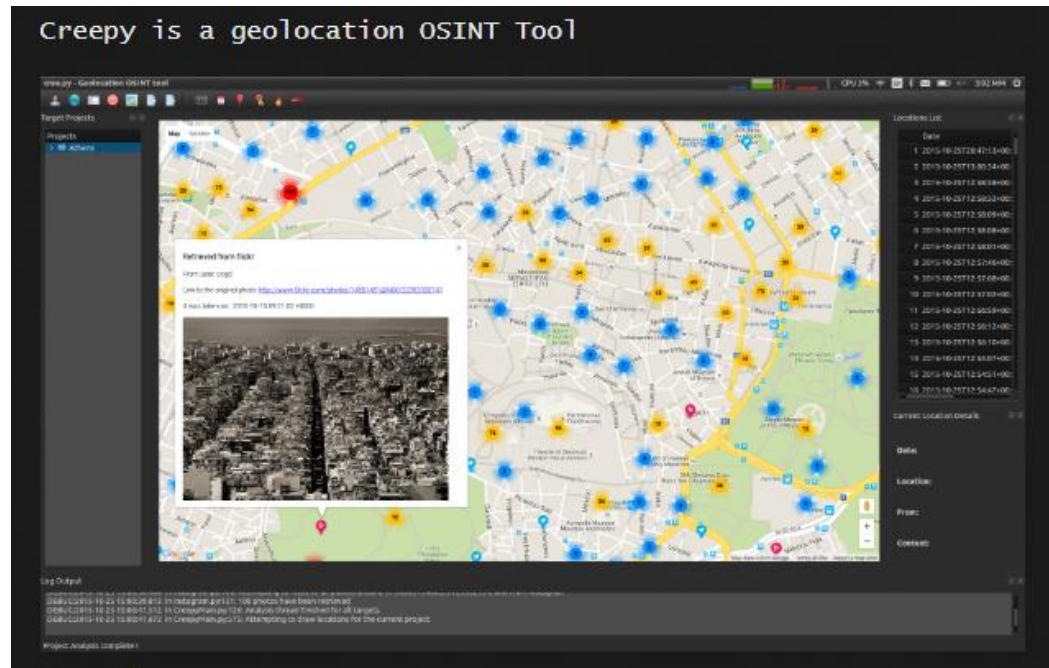
Social Media Presence



# LOCATION AWARENESS - CREE.PY

**Cree.py** (<https://www.geocreepy.com>) is an open source intelligence gathering application.

Can gather from Twitter and any geo-location data from several websites.



## ./ Creepy

A Geolocation OSINT Tool. Offers geolocation information gathering through social networking platforms.

### windows Downloads

Current version : v1.4.1

[Download 64bit Windows installer](#)[Download 32bit Windows installer](#)

### OSX Downloads

Current version : v1.4.1

[Download OSX installer](#)

### Source Code Downloads

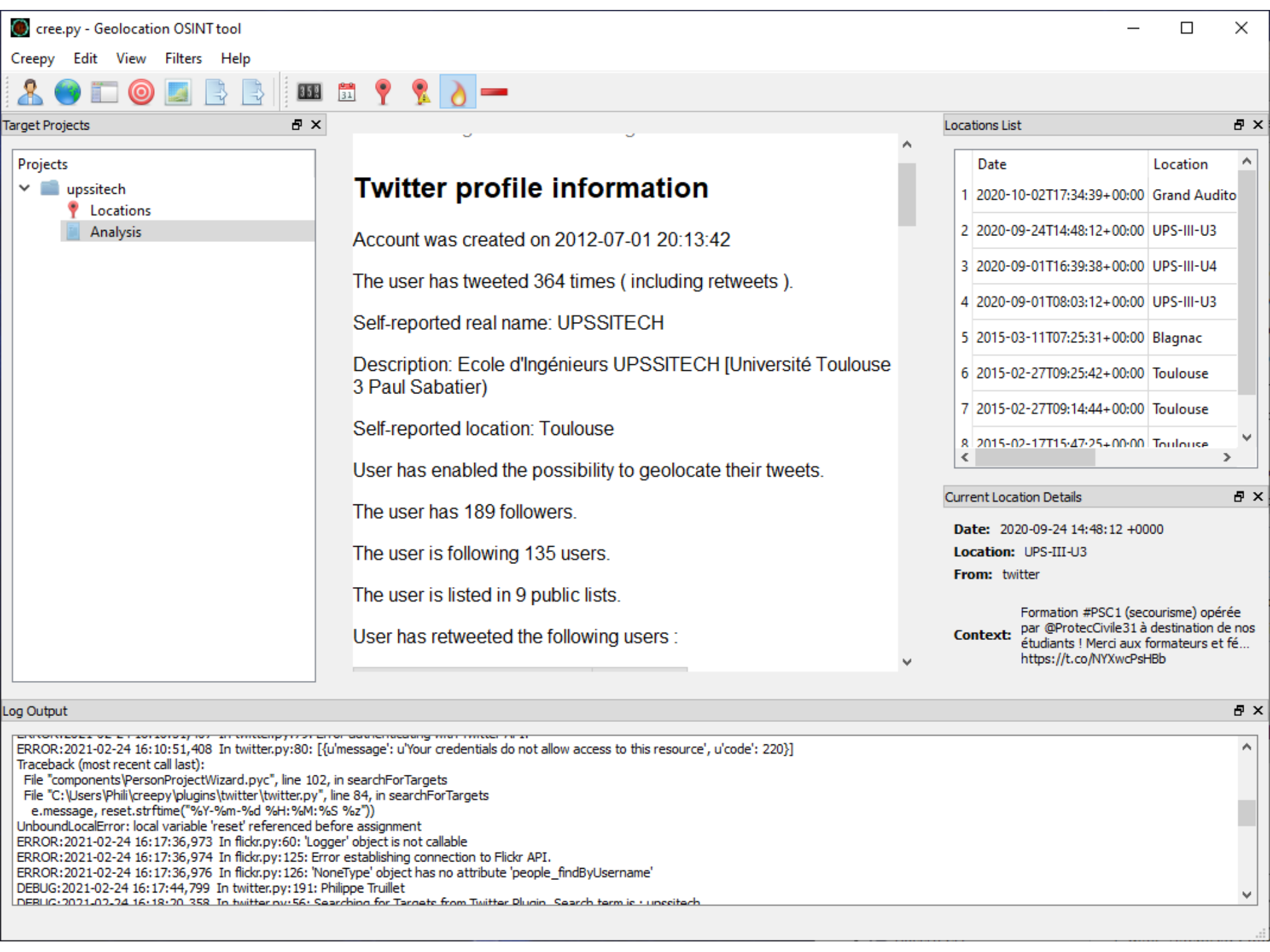
[Download as .zip](#)[Download as .tar.gz](#)[View on GitHub](#)

### Plugin Downloads

[Download as .zip](#)[Download as .tar.gz](#)

## Creepy

Creepy is a geolocation OSINT Tool



# INDIVIDUAL - INTERNET PRESENCE

Email Address

Personal Handles/Nicknames

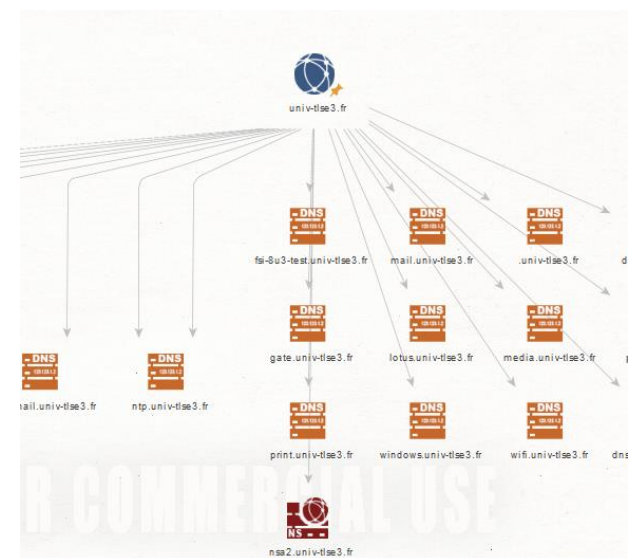
Personal Domain Names registered

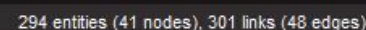
Assigned Static IPs/Netblocks

# MALTEGO

**Paterva Maltego** (<https://www.paterva.com>) is a data mining and information-gathering tool that maps the information gathered into a format that is easily understood and manipulated.

It saves you time by automating tasks such as email harvesting and mapping subdomains.

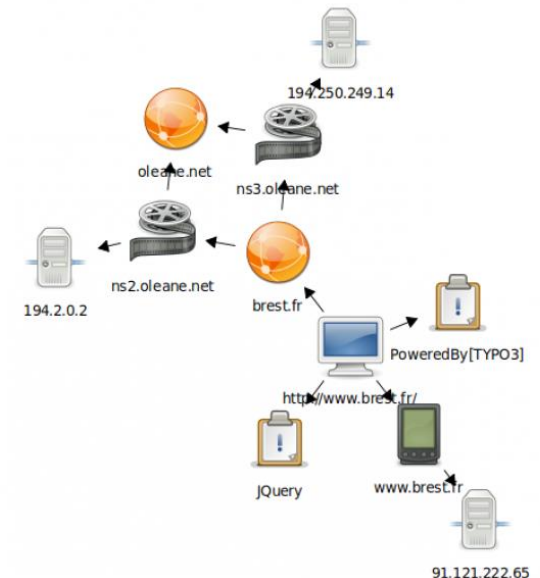


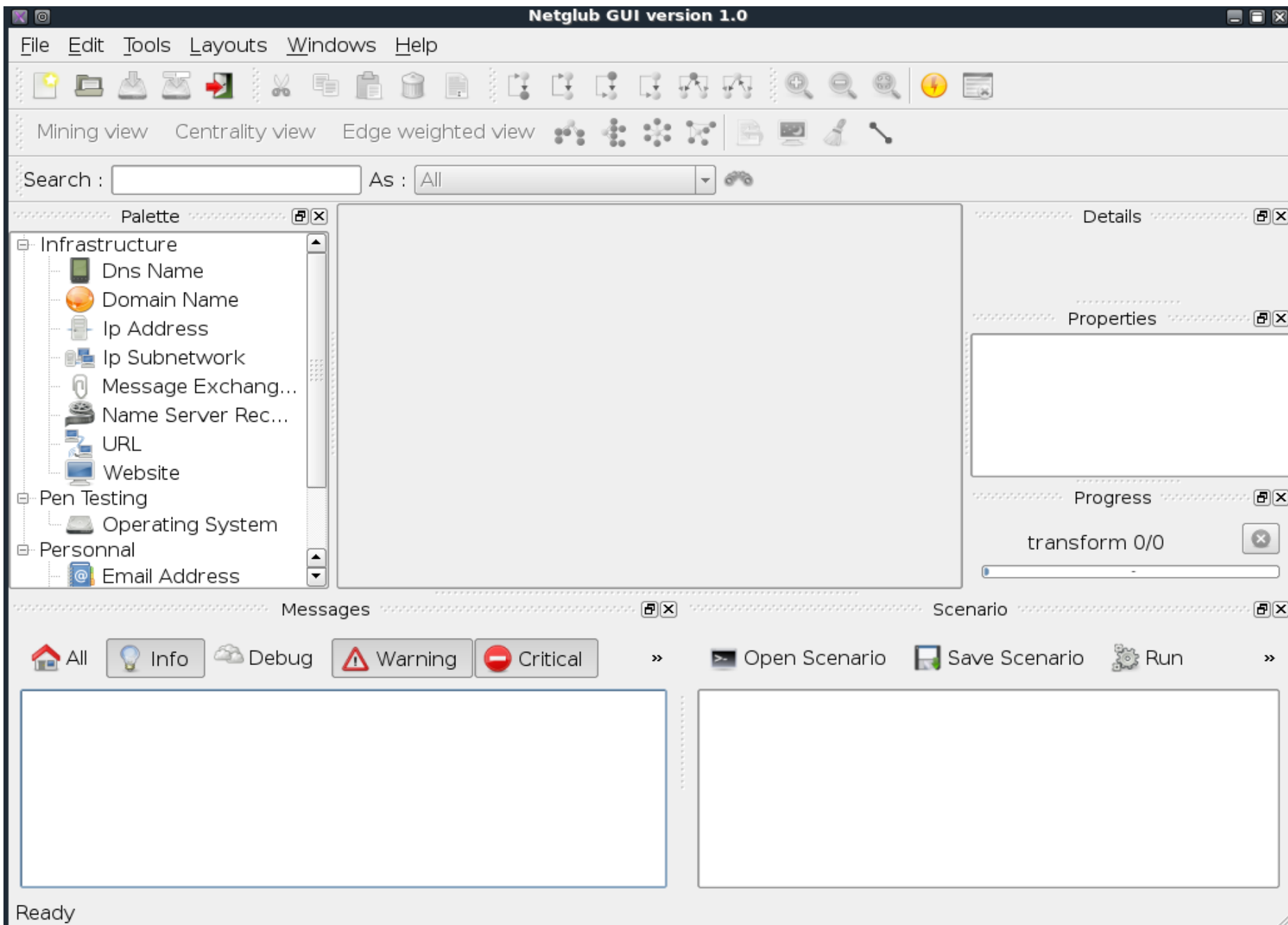


# NETGLUB

**NetGlub** (<http://www.netglub.org>) is an open source data mining and information-gathering tool that presents the information gathered in a format that is easily understood, (Similar to Maltego).

Consists of: Master, Slave, and GUI







Netglub GUI version 1.0

File Edit Tools Layouts Windows Help

Mining view

Centrality view

Edge weighted view

As : All

Palette

Infrastructure

Dns Name

**Domain Name**

Ip Address

Ip Subnetwork

Message Exchang...

Name Server Rec...

URL

Website

Pen Testing

Operating System

Personal

Email Address

Default Graph 1.ng

Details

Domain Name

google.com

Properties

Name	value
Entity Informations	
Domain name	goo...
Graph Informations	
Nb In Edges	0

Progress

transform 3/10

Messages

All

Info

Debug

Warning

Critical

15:46:58 : \*\*\* Transform from google.com To MX [Dig] finish  
15:46:58 : \*\*\* Transform from google.com To Domain [Top Le  
15:46:58 : \*\*\* Transform from google.com To Location [Who  
15:46:58 : \*\*\* Transform from google.com To Website [www  
15:46:58 : \*\*\* Transform from google.com To NS [Dig] finishe  
15:46:58 : \*\*\* Transform from google.com To Email [Whois]  
15:46:58 : \*\*\* Transform from google.com To Dns Name [SE  
15:46:53 : \*\*\* Transform from "google.com" To Domain [Top

Open Scenario

Save Scenario

Run

# TheHarvester

(<https://github.com/laramies/theHarvester>) is a tool, written by Christian Martorella, that can be used to gather e-mail accounts and subdomain names from different public sources (search engines, pgp key servers).

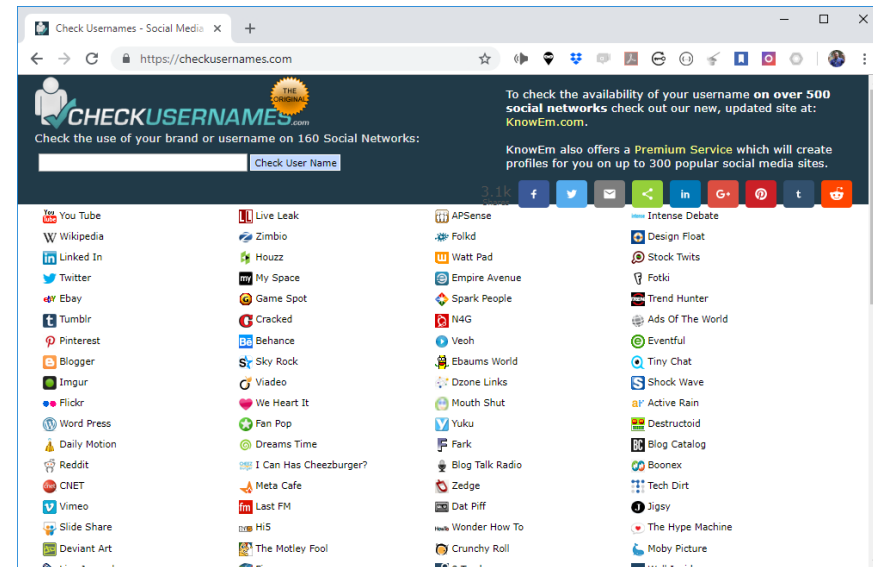
→ <http://www.edge-security.com>

[illegible]

# SOCIAL NETWORKS

Check Usernames - Useful for checking the existence of a given username across 160 Social Networks.

<http://checkusernames.com>



## Mail Lists

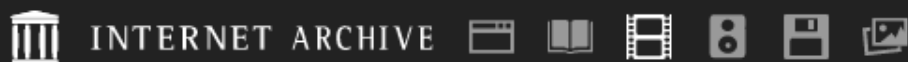
The Mail Archive - <https://www.mail-archive.com>

# ARCHIVED INFORMATION

There are times when we will be unable to access web site information due to the fact that the content may no longer be available from the original source.

Being able to access archived copies of this information allows access to past information.

1. Perform Google searches using specially targeted search strings: `cache:<site.com>`
2. Use the archived information from the Wayback Machine (<https://www.archive.org/web>).

[ABOUT](#) [BLOG](#) [PROJECTS](#) [HELP](#) [DONATE](#) ❤️ [CONTACT](#) [JOBS](#) [VOLUNTEER](#) [PEOPLE](#)

Explore more than 538 billion [web pages](#) saved over time

[BROWSE HISTORY](#)

Find the Wayback Machine useful?

[DONATE](#)

## Tools

[Wayback Machine Availability API](#)

Build your own tools.

[WordPress Broken Link Checker](#)

Banish broken links from your blog.

[404 Handler for Webmasters](#)

Help users get where they were

## Subscription Service

Archive-It enables you to capture, manage and search collections of digital content without any technical expertise or hosting facilities. [Visit Archive-It to build and browse the collections.](#)

## Save Page Now

[SAVE PAGE](#)

Capture a web page as it appears now for use as a trusted citation in the future.

Only available for sites that allow crawlers.

# METADATA LEAKAGE

The goal is to identify data that is relevant to the target corporation.

It may be possible to identify locations, hardware, software and other relevant data from Social Networking posts.

## Examples:

- ixquick - <https://ixquick.com>
- MetaCrawler - <https://metacrawler.com>
- Dogpile - <https://www.dogpile.com>
- Search.com - <https://www.search.com>
- Jeffery's Exif Viewer <https://exif.regex.info/exif.cgi>

# METADATA LEAKAGE - FOCA

FOCA (<https://github.com/ElevenPaths/FOCA>) is a tool that reads metadata from a wide range of document and media formats.

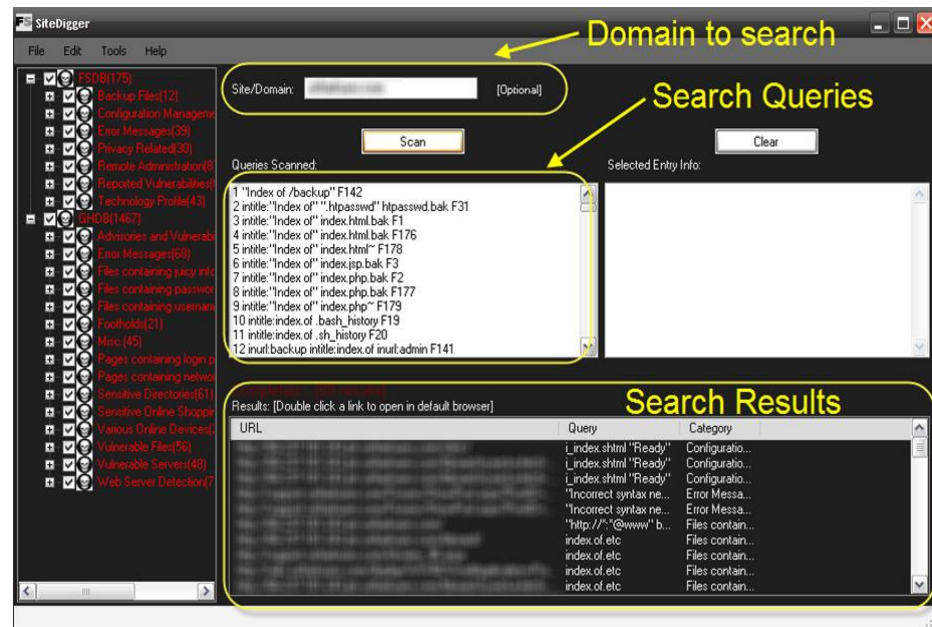
FOCA pulls the relevant usernames, paths, software versions, printer details, and email addresses.



# METADATA LEAKAGE - FOUNDSTONE SITEDIGGER

## SiteDigger

(<https://sitedigger.software.informer.com/3.0/>) allows us to search a domain using specially strings from both the Google Hacking Database (GHDB) and Foundstone Database (FSDB).





# METADATA LEAKAGE - METAGOOFIL

**Metagoofil** (<https://github.com/laramies/metagoofil>) is a Linux based information gathering tool designed for extracting metadata of public documents (.pdf, .docx, .xlsx, .pptx, .odp, .ods) available on the client's websites.

Metagoofil generates an html result page with the results of the metadata extracted, plus a list of potential usernames that could prove useful for brute force attacks. It also extracts paths and MAC address information from the metadata.

# INDIVIDUAL

**Physical Location**

**Mobile footprint**

Phone #

Device type

Installed applications

# COVERT GATHERING - CORPORATE

## **On-Location Gathering**

Physical security inspections

Wireless scanning / RF frequency scanning

Employee behavior training inspection

Accessible/adjacent facilities (shared spaces)

Types of equipment in use

## **Offsite Gathering**

Data center locations

Network provisioning/provider

# OTHER GATHERING FORMS

## Human Intelligence (**HUMINT**)

Methodology always involves direct interaction - whether physical, or verbal.

Gathering should be done under an assumed identity (*remember pretexting?*).

- Key Employees
- Partners/Suppliers

# OTHER GATHERING FORMS

Signals Intelligence (**SIGINT**):

Intelligence gathered through the use of interception or listening technologies.

Example:

- Wired/Wireless Sniffer
- TAP devices

# OTHER GATHERING FORMS

## Imagery Intelligence (IMINT):

Intelligence gathered through recorded imagery, i.e. photography.

IMINT can also refer to satellite intelligence, (cross over between IMINT and OSINT if it extends to Google Earth and its equivalents).

- <https://www.geoportail.gouv.fr>
- <https://remonterletemps.ign.fr>
- <https://earth.google.com/web>