

# Gathering Electronic OSINT

---



**Jerod Brennen**

CISSP, INFOSEC GEEK

@slandail [www.slandail.net](http://www.slandail.net)



# Overview



**Marketing communications**

**Document metadata**

**Automation**



# Electronic OSINT

Information extracted from public documents and document metadata



“We leave traces of ourselves  
wherever we go, on whatever we  
touch.”

**Lewis Thomas, Artist & Scientist**



# Marketing Communications (L1/L2)

---



# Past and Present





## The Wayback Machine

- <https://archive.org/web/>

Enter target site > Browse History  
Cached URL's

# Phishing

The act of sending a fraudulent email to an unsuspecting recipient with the intention of deceiving that recipient





**Newsroom**  
**PR Newswire**  
**Social Media Profiles**  
**Mailing Lists**

## Marketing Communication Resources



`site:company_domain.com "company_name" logo`

## Search Engine Queries

**Usage: Locating current brand resources for phishing campaigns**



# Contextual Data

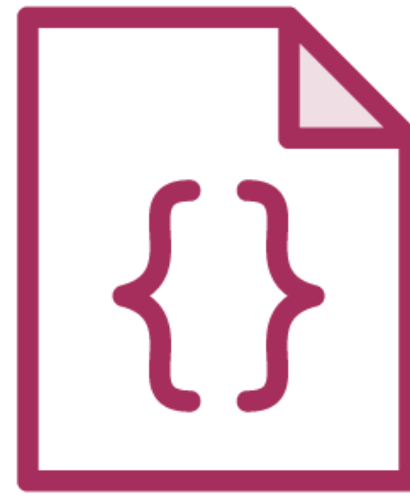
---



# Contextual Data vs. Metadata



Contextual Data



Metadata



## Google Hacking Database

- <https://www.exploit-db.com/google-hacking-database/>

## Johnny Long

- <http://johnny.ihackstuff.com/>

# GHDB Categories

Files containing usernames

Files containing passwords

Files containing juicy info

Sensitive directories



`site:company_domain.com filetype:pdf`

`site:company_domain.com filetype:docx`

`site:company_domain.com filetype:xlsx`

`site:company_domain.com filetype:pptx`

## Search Engine Queries

**Usage:** Identify publicly available documents on a target organization's website





## HTTrack

- <https://www.httrack.com/>

Website copier

Finds and downloads files

Spiders hyperlinks



# Document Metadata (L1/L2)

---



# Extracted from Metadata

**Author/Creator  
Name**

**Date and Time  
Created**

**Endpoint  
Technologies**

**Network  
Information**

**Geolocation**



# Platform/Language Comfort Level



Windows  
vs.  
Linux



Perl



Python



Java/XML





## FOCA

- <https://www.elevenpaths.com/labstools/foca/index.html>

**Fingerprinting Organizations with  
Collected Archives**

**Windows / GUI**



## ExifTool

- <http://www.sno.phy.queensu.ca/~phil/exiftool/>

## Perl library

## Platform Independent

## Useful One-Line Commands

- <https://sourceforge.net/projects/exiftool1line/files/>



## Metagoofil

- <http://www.edge-security.com/metagoofil.php>
- <https://github.com/laramies/metagoofil>

## Python

Similar to FOCA

Automated process



## Meta-Extractor

- <http://meta-extractor.sourceforge.net/>

## Now ManualDeposit

- <https://github.com/DIA-NZ/ManualDeposit>

## Java/XML

Documents, images, and sound files

# One Two Punch



**Step 1:**  
Download, extract, and  
analyze metadata  
(e.g., FOCA, Metagoofil)



**Step 2:**  
Perform in-depth analysis  
of specific files  
(e.g., ExifTool)





# Demo



## Tesla Motors

- <https://bugcrowd.com/tesla>



# Summary



**Finding publicly available documents**

**Reviewing contextual data**

**Extracting and analyzing metadata**



Onto the next module:  
Identifying Infrastructure  
Assets

