

OSINT

```
Active measures by your target (Planting a false trail): Here we have a target that
                                                                                            has intentionally placed false data
                                                                                            Unintentionally incorrect data (OCR, Siri, typos): Humans enter data not accurate.
                                                                                            Old data that wasn't updated (whois): The whois system is an excellent example of
                                             Potential reasons for inaccurate data
                                                                                            this class of false data.
                                                                                            Similar target's data (Multiple people, same names): Your target's name is very
                                                                                            common for the region you are searching.
                                                                                            Human nature (Bragging, appearing to be someone they are not – think online
                                                                                            dating profiles): people who share/post data wrong in social network
     1. Data Analysis Challenges
                                                                                             Correlation: means that a relationship exists between two or more things.
                                             Correlation Does Not Imply Causality
                                                                                              Causality: is different. With causality, one thing makes something else happen.
                                             Sound Analysis
                                           Free PDF on Analytical Thinking — Psychology_of_Intelligence_Analysis.pdf &
                                                                                                   As proxy, Each report is stored on urlscan.io's server and is retrievable via its unique
                                                                                urlscan.io 🧔
                                                                                                   URL, to get web response and analysis web technologies
                                                                                spacex.com @ — More info about web app

    Using a proxy web application

                                                                                                    wget r l3 e robots=off nocheckcertificate A .pdf https://www.thirdpresbyterian.org/
                                                                           Download Content
                                                                                                              grep (
                                                                                                              cygwin 👌
                                       Using commandline tools
                                                                                                              GunWin d
                                                                                                              finds tr 🍼
                                                                           Search in downloaded content
                                                                                                                              for file in ls *.pdf; do pdftotext $file; done
                                                                                                                               converts PDFs to text files to investigate in it
                                                                                                              pdfgrep
                                                                                                              Mythicsoft — as grep
                                                                                                              Agent Ransack — GUI searching tool in files
                                                              OSINT Public Scripts that are for public consumption.
                                                              EyeWitness.py & — Take Screenshot from URL
                                       - Using a script
                                                               Browser Extension That Extracts Data [Instant Data Scraper]
                                                               viewdns.info P Reverse whoi
                                                             Search for victim web app API
                                        Using an API
                                                          API Reference &
2. Harvesting Web Data
                                                             We can use [Public API] Collection to collect data. 🔗
                                                                                       Privately funded site that archives web sites (text, images, JavaScript), Fewer
                                                                                       backups than archive.org.
                                                                      Archive.org web site or use CLI tool
                                                                                                                  https://webcache.googleusercontent.com/search?q=cache:q0ser3Y903YJ:https://

    Using cached content

                                                                                                                  www.sans.org/course/opensourceintelligencegathering
                                                                                                                  +&cd=1&hl=en&ct=clnkgl=us&strip=1
                                                                      Google domain search cached content
                                                                                                                  To prevent Google from pulling live resources, add & strip=1 to end of URL.
                                                                       Browser Extensions for Archives
                                                                                                                       Visit [Google's Mobile Friendly Test] 🔗
                                        How to bypass authentication to get e.g. LinkedIn person profile
                                                                                                                       Select the HTML tab
                                        details with out login into LinkedIn
                                                                                                                       Open web app to code beautify HTML and render html to web page
                                                                                                                        [codebeautify]
                                                                               Both Google products, AdSense and Analytics, use unique numbers for each
                                                                               customer account, This common connection helps us find other sites that
                                                                               may help our investigation
                                                                                                                                                             AdSense = pub12345678
                                                                               The unique codes are found in the source code of the pages and look like:
                                        Google AdSense and Analytics
                                                                                                                                                             Analytics = UA87654321
                                                                                                                                                              spyonweb 🔗
                                                                               Enter Google Analytics code in search field OR domain to get Analysis ID
                                                                                                                                                               sameid 🔗
                                                                                and results show other domains related with this Analysis ID
                                                                                                                                                               builtwith.com & — we can examine trackers across time
                                                                         Web sites using HTTPS present TLS or SSL certificates to browsers to
                                                                         create the encrypted connections.
                                                                         These certificates can contain sensitive or private data we can harvest

    Encryption Certificates

                                                                         Result in CA as Domains, IP addresses, Emails, Physical locations
                                                                         Searching for this site on censys.io: yielded the certificate contents on the
                                                                         right, another websites: entrust.com, transparencyreport.google.com, crt.sh
                                                           CLI or GUI
                                                                               exiftool c %+.7f GPSPosition *.jpg
   3. File Metadata Analysis
                                          thexi fer  Web APP WARNING: Sharing files
                                          verexif: Web APP @ — WARNING:this web app Sharing files
                                          FOCA (Fingerprinting Organizations with Collected Archives) & — FOCA is a tool that has been well known for metadata analysis of files.
                                                      OSINT Source Code &
                                        Technisette 🔗
                                       OSINT Team Links 🔗
                                       Onstrat website 🔗
                                        Reuser web site &
     4. OSINT Frameworks
                                        (OSINT) Recherchetraining &
                                        Terrorism & Radicalisation Research Dashboard & — Looking for international terrorists and the systems and tools they use?
                                       OSINT analysts that need to venture into the world of dating and sx sites &
                                       Bellingcat & Bellingcat maintains a Google document with well categorized links for international OSINT
                                     D3 by Observable & To Build you own Visualization
```