**PRACTICAL 3:**

**OSINT for Tracing Images, Videos, and Documents**

**Overview**

OpenSource Intelligence (OSINT) enables investigators and attackers alike to gather useful information from publicly accessible resources. Media files such as images, videos, and documents often contain valuable information in their content, metadata, and context, which can leave a digital footprint. This document explores how attackers gather useful information from these sources and how such footprints can be traced back to their origin.

**How Users Leave Digital Footprints through Images, Videos, and Documents**

1. **Images:**
   * **Metadata (EXIF Data):**
     + Embedded details like GPS coordinates, device type, camera settings, and timestamps.
     + Example: A photo uploaded to social media may reveal the location where it was taken.
   * **Visual Content:**
     + Background details like landmarks, text, or identifiable objects can provide contextual clues.
   * **Online Distribution:**
     + Sharing the same image on multiple platforms links them to a user’s online activity.
2. **Videos:**
   * **Metadata:**
     + Creation dates, editing software details, and sometimes geotags.
   * **Audio-Visual Content:**
     + Audio, background objects, or people can help identify locations or events.
     + Key frames can be extracted for further analysis.
   * **Platform-Specific Data:**
     + Videos uploaded to platforms like YouTube often include uploader details and timestamps.
3. **Documents:**
   * **Metadata:**
     + Author names, creation/editing history, and software information.
   * **Embedded Content:**
     + Hyperlinks, images, or media files within documents can point to external sources.
   * **Textual Content:**
     + Keywords, names, or contact information can be used to track associated online activity.

**OSINT Approach: Tracing Digital Footprints**

Let’s consider a scenario where an attacker has a random dog image and aims to determine its source and gather additional information.

**Step-by-Step OSINT Workflow**

**Step 1: Searching for Similar Images**

* **Objective:** Find the source or other instances of the image online.
* **Method:**
  + Use reverse image search tools like Google Images, Bing Visual Search, or Yandex Images.
  + Tools like [Creative Commons Search](https://search.creativecommons.org/) can compare results across various search engines.
* **Outcome:** Identify platforms, websites, or social media accounts where the image has been uploaded.

**Step 2: Extracting Metadata**

* **Objective:** Gather additional details embedded in the image file.
* **Tools:**
  + Use metadata extraction tools like ExifTool, Jeffrey’s Image Metadata Viewer, or Metapicz.
* **Insights Obtained:**
  + Location (GPS coordinates) and time the image was taken.
  + Device details, including camera make, model, and settings.
* **Further Action:** Use location data for additional reverse searches or cross-referencing with public databases.

**Step 3: Performing Forensic Analysis**

* **Objective:** Uncover tampered or hidden data and refine the search.
* **Tools:**
  + Analyze the image with forensic tools like FotoForensics or Amped Authenticate.
  + Identify signs of editing or manipulation, such as cloning, cropping, or splicing.
* **Outcome:** Obtain precise information about the image’s authenticity and origin.

**Step 4: Contextual Analysis**

* **Objective:** Use contextual clues to trace activity or connections.
* **Method:**
  + Analyze visible text using OCR tools like Tesseract to extract keywords or names.
  + Identify recognizable landmarks, objects, or brands in the image.
  + Cross-reference with social media profiles or news sources to establish connections.

**Scenario: Tracing the Source of an Image**

**Scenario Description**

An investigator finds a random image of a dog shared online and aims to trace its origin and gather as much information as possible using OSINT techniques.

**Step 1: Searching for Similar Images**

* **Action Taken:**
  + The investigator uses Google Reverse Image Search by uploading the dog image.
  + Yandex and TinEye are also used to find visually similar images or websites hosting the same file.
* **Result:**
  + Google Images identifies a match on a public Instagram account.
  + TinEye shows the image uploaded to a pet adoption website three months earlier.

**Step 2: Extracting Metadata**

* **Action Taken:**
  + The investigator downloads the image and uses **ExifTool** to extract metadata.
* **Metadata Details Obtained:**
  + **Timestamp:** Image taken on January 15, 2024, at 3:15 PM.
  + **Location:** GPS coordinates embedded in the image lead to a park in Austin, Texas.
  + **Device:** Photo captured using an iPhone 13 Pro.
* **Follow-Up:**
  + The GPS coordinates are cross-referenced with Google Maps to pinpoint the exact park.

**Step 3: Performing Forensic Analysis**

* **Action Taken:**
  + **FotoForensics** is used to analyze the image.
  + **Findings:**
    - No signs of tampering detected.
    - A watermark is visible in the bottom corner, suggesting the image belongs to "PawsLife Photography."
* **Follow-Up:**
  + A search for "PawsLife Photography" leads to their official website, where a similar dog image is featured in their portfolio.

**Step 4: Contextual Analysis**

* **Action Taken:**
  + - The investigator extracts text from the Instagram caption accompanying the image using **OCR (Tesseract):**
    - Caption: "Adopt me! Meet Max, a playful retriever looking for a loving home. Taken at Zilker Park, Austin."
  + A search for "Max retriever Zilker Park adoption" leads to a local pet shelter's website.
* **Result:**
  + The shelter confirms the image was taken during an adoption event at Zilker Park.

**Conclusion**

Using OSINT techniques, the investigator successfully traced the image's origin to a pet adoption website and identified contextual details such as the image's timestamp, location, and owner. This highlights how digital footprints in images can lead to a trail of connected information.

**Using Videos and Documents for Tracing Digital Footprints**

1. **Videos:**
   * Extract metadata using tools like FFmpeg or ExifTool.
   * Analyze audio for voices, background sounds, or accents.
   * Extract keyframes for reverse image searches.
2. **Documents:**
   * Use tools like FOCA or pdfinfo to extract metadata.
   * Analyze embedded links or media for related sources.
   * Search for specific phrases, file hashes, or unique content to find other instances online.

**Conclusion**

Images, videos, and documents provide a wealth of information that attackers can exploit for tracing a target’s digital footprint. By leveraging OSINT tools and techniques, it is possible to uncover the origins, context, and connections of such media. This underscores the importance of being cautious with the content shared online and understanding the potential risks associated with metadata and contextual clues.