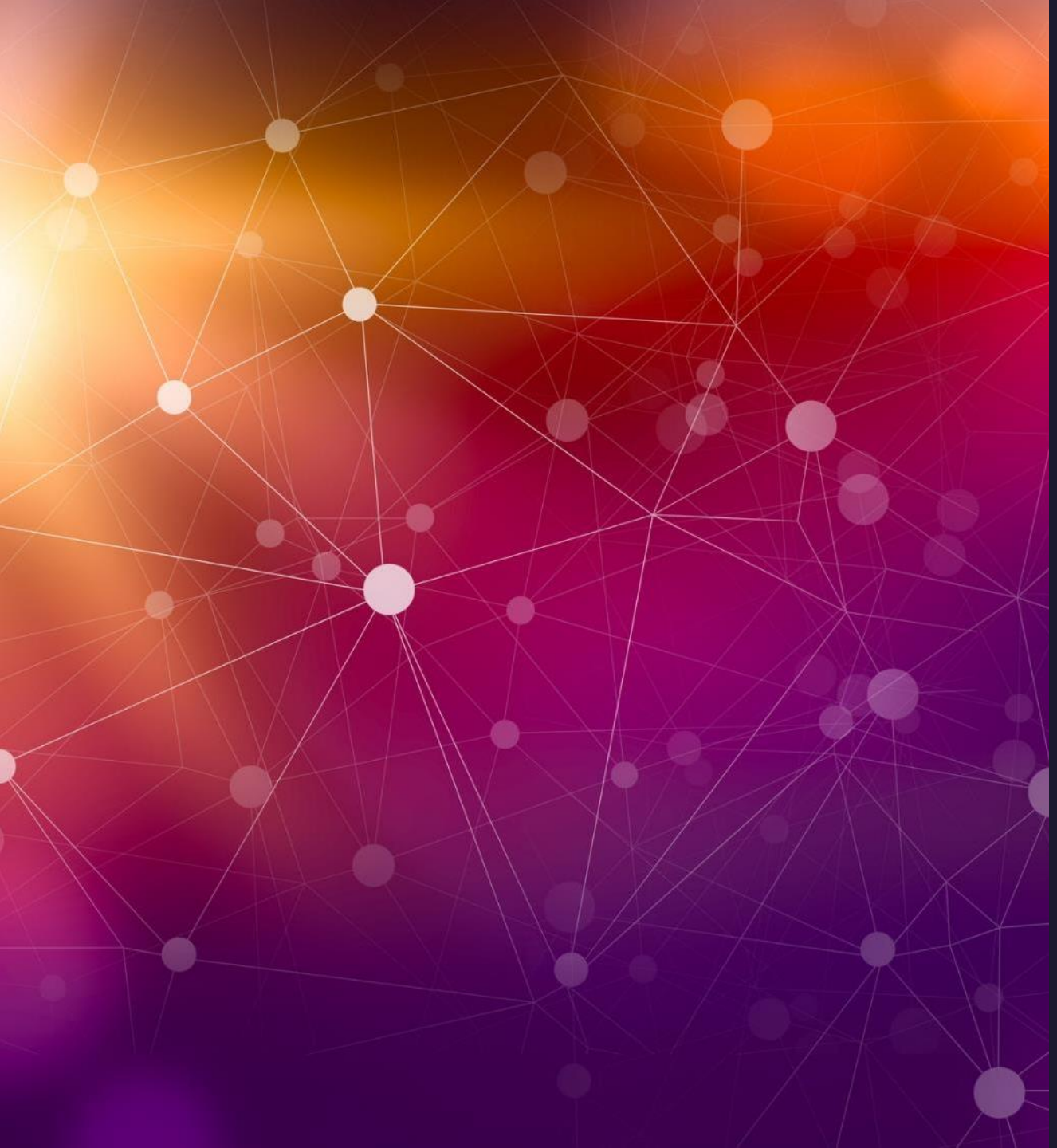


L08: OSINT

by Angel Vasques, Roland Cheremeh,
& Jai Ghude

What is Open-Source Intelligence?

- Open- Source Intelligence also known as OSINT can be defined as practice of collecting and analyzing information from publicly available sources to gain insights, knowledge, and understanding about a particular topic, individual, organization, or event.
- The information used in OSINT can be from many types of sources such as public records, news media, social media, websites, etc.
- The source of this information can come from public records, news media, social media, websites, etc.
- The use of OSINT can benefit law enforcement, research, marketing and journalism.
- OSINT can answer questions such as, what trends are currently popular, how many people use a certain website, what trends changed after a certain amount of time.

Significance of OSINT

- **Comprehensive collection of information :** OSINT tool provides a pool of open sources of information and therefore information analysts are able to obtain comprehensive insights (Smith, 2023)
- **Support decision making:** Comprehensive insights enables the stakeholders to make informed decisions based on the current age of technology
- **Note:** Large amount of information or content are shared online on a daily basis and therefore OSINT is important to gather and analyze the generated content as it informs decision making, threats and risk mitigations

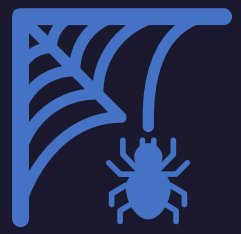
Sources of information

- Social Media: Platforms like Twitter, Facebook, Instagram, and LinkedIn provide real-time insights into public opinions, trends, and interactions.
- Websites and Blogs: Online articles, blogs, forums, and websites offer information on various subjects.
- News Outlets: News articles, press releases, and journalistic reports provide current events and information.
- Public Records: Government databases, court records, and regulatory filings contain official data.
- Academic Publications: Research papers and scholarly articles offer in-depth information on specific topics.

Social Media Monitoring

- This technique uses data from different social media platforms and uses this data to make inferences and answer questions.
- This technique can be used across a variety of social media networks.
- The purposes of social media monitoring include competitive analysis, crisis management, audience insight, etc.
- This technique is efficient due to its access to information from billions of daily internet users.

OSINT techniques



Web Scraping: This involves automated collection of data from websites.

Metadata Analysis: Extracting hidden information from files, images, and documents

Geolocation: Determining the geographic location of data sources.

The ethics of Social Media Monitoring

- One of the ethical problems with social media monitoring is the idea that users may not be aware the data being collected and how it is being used.
- The use of this information can be used to manipulate and persuade a group of people by using this information to attract more users and keep users coming back to the site.
- Social media monitoring can be biased and assume a user wants something based off of gender, race, etc.

Web Scraping

- Definition: Web scraping is the automated extraction of data from websites. It serves as a fundamental technique within Open Source Intelligence (OSINT) for gathering valuable information from online sources.
- Importance:Web scraping is crucial for OSINT practitioners as it enables them to collect vast amounts of data efficiently, which can then be analyzed to uncover insights and trends relevant to their objectives.

Process Overview

Steps in Web Scraping:

- Identify target websites: OSINT analysts begin by identifying the websites containing the desired information. These may include news websites, social media platforms, online forums, or any other publicly accessible online sources.
- Configure scraping tools: Once the target websites are identified, scraping tools are configured to extract specific data elements such as text, images, or links. Parameters are set to determine the scope and frequency of data collection.
- Extract desired information: The scraping tools systematically navigate through the website's structure, retrieving the designated data according to the configured parameters.
- Analyze collected data for insights: The extracted data is then processed and analyzed to uncover patterns, trends, or valuable information relevant to the OSINT objectives.

Use Cases

- Business and Competitive Intelligence: Companies use OSINT to monitor competitors, market trends, and consumer sentiments.
- Cybersecurity: Identifying potential threats and vulnerabilities by monitoring online activities.
- Investigative Journalism: Journalists use OSINT to gather facts and corroborate information for news stories.
- Law Enforcement: Supporting investigations, tracking criminals, and identifying potential threats.
- Academic Research: Collecting data for research purposes and gathering insights for studies.
- Digital Marketing: Understanding consumer behavior, sentiment, and trends to inform marketing strategies.

Considerations

Legal and Ethical Considerations:

- Compliance with website terms of service: OSINT practitioners must ensure that their web scraping activities comply with the terms of service and usage policies of the target websites.
- Avoidance of unauthorized access to restricted data: It is essential to refrain from accessing or scraping data that is protected by authentication mechanisms or intended for restricted use.
- While OSINT is a valuable tool, it also raises ethical concerns. Privacy, data protection, and the responsible use of information are crucial considerations. OSINT practitioners must respect legal and ethical boundaries, avoid invasion of privacy, and ensure that information is used for legitimate purposes.

Conclusion

Recap of Open Source Intelligence (OSINT):

- OSINT is a vital methodology for gathering and analyzing information from publicly available sources to gain insights and understanding about various topics, individuals, organizations, or events.
- It encompasses a wide range of sources, including social media, websites, news outlets, public records, geospatial data, and academic publications.
- OSINT techniques and tools, such as web scraping, social media analysis, and geolocation, enable practitioners to collect and analyze data effectively.
- Applications of OSINT span across intelligence, cybersecurity, journalism, law enforcement, business, and research domains, facilitating tasks such as competitive intelligence, threat detection, investigative journalism, and academic research.
- Ethical considerations, including privacy, data protection, and responsible information use, are paramount in OSINT practice.
- Continuous learning and adaptation to evolving technology and legal developments are essential for effective OSINT practice.
- OSINT is a versatile and powerful approach to gathering information from public sources. It plays a significant role in decision-making, problem-solving, and understanding the digital world, while also requiring responsible and ethical use of the collected data.

Sources

- https://www.sans.org/blog/what-is-open-source-intelligence/
- https://osintteam.blog/social-media-osint-a-comprehensive-guide-to-gathering-intelligence-from-social-media-platforms-b5dbb8d83f14
- https://www.puresquare.com/blog/the-ethics-of-social-media-tracking-is-it-time-for-change/
- https://medium.com/@investigator515/webscraping-osint-a-beginners-guide-bec9ef686dd
- https://www.opentext.com/what-is/open-source-intelligence-osint

Sources

- Smith, J. (2023). Leveraging Open Source Intelligence (OSINT) for Cybercrime Investigations: A Case Study. Journal of Cybersecurity Research, 7(2), 123-137.

- Johnson, A. & Lee, S. (2022). Proactive Cyber Threat Intelligence: Using OSINT Techniques to Identify Emerging Risks. International Journal of Information Security, 15(4), 321-335.

Thank you