OSINT

- ✓ OSINT refers to all the information that is publicly available.
- ✓ There is no specific date on when the term OSINT was first proposed; however, a relative term has probably been used for hundreds of years to describe the act of gathering intelligence through exploiting publicly available resources.
- ✓ The U.S. Department of Defense (DoD) defines OSINT as follows:

"Open-source intelligence (OSINT) is an intelligence that is produced from publicly available information and is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement."

- ✓ OSINT sources are distinguished from other forms of intelligence because they must be legally accessible by the public without breaching any copyright or privacy laws. That's why they are considered "publicly available." This distinction makes the ability to gather OSINT sources applicable to more than just security services. For example, businesses can benefit from exploiting these resources to gain intelligence about their competitors.
- ✓ During the search for OSINT sources, classified information that is not protected properly can appear. This includes leaked documents, such as those published by WikiLeaks. This type of information is called NOSINT, as opposed to OSINT. Intelligence usually considers all sources regardless of their legal accessibility.

OPEN SOURCE INFORMATION CATEGORIES

There are four categories of open information and intelligence as per NATO literature on OSINT:-

- ✓ **Open source data (OSD):** This is generic data coming from a primary source. Examples include satellite images, telephone call data and metadata, datasets, survey data, photographs, and audio or video recordings that have recorded an event.
- ✓ **Open source information (OSINF):** This is generic data that has undergone some filtering first to meet a specific criterion or need; this data can also be called a secondary source. Examples include books about a specific subject, articles, dissertations, artworks, and interviews.
- ✓ **Open source intelligence (OSINT):** This includes all the information that has been discovered, filtered, and designated to meet a specific need or purpose. This information can be used directly in any intelligence context. OSINT can be defined in a nutshell as the output of open source material processing.
- ✓ **Validated OSINT (OSINT-V**): This is OSINT with a high degree of certainty; the data should be confirmed (verified) using a non-OSINT source or from a highly reputable OSINT source. This is essential, as some outside adversaries may spread inaccurate OSINT information with the intent to mislead OSINT analysis. A good example of this is when a TV station broadcasts live the arrival of a president to another country; such information is OSINT, but it has a large degree of certainty.

- ✓ The set of sources legally available to the public through specific channels is called gray literature.
- ✓ These sources include books, journals, dissertations, technical reports, and internal documents of commercial enterprises, commercial imagery, and any information that is controlled by its producer.
- ✓ Gray literature is a major element of OSINF and can be obtained legally by acquiring the permission of its copyright holder or by paying for it (for example, through subscriptions agencies, commercial bookstores, and so on.

- ✓ **OSD and OSINF** comprise the main sources (primary and secondary) of information that OSINT uses to drive its results.
- ✓ Another issue you need to understand within the OSINT context is the difference between data, information, and knowledge.
- ✓ The three terms are usually used interchangeably; however, each one has a different meaning, although the three do interact with each other.
  - ✓ **Data**: This is a set of facts describing something without further explanation or analysis. For example, “The price of gold per ounce is $1,212.”
  - ✓ **Information**: This is a kind of data that has been interpreted properly to give a useful meaning within a specific context. For example, “The price of gold per ounce has fallen from $1,212 to $1,196 within one week.”
  - ✓ **Knowledge**: This is a combination of information, experience, and insight that has been learned or inferred after some experimentation. Knowledge describes what your brain has recorded in the past, and these records can help you to make better decisions about the future when facing similar contexts. For example, “When the price of gold falls more than 5 percent, this means the price of oil will fall too.”

OSINT TYPES

- ✓ OSINT includes all publicly accessible sources of information.
- ✓ This information can be found either online or offline, including in the following places:-
  - ✓ The Internet, which includes the following and more: forums, blogs, social networking sites, video-sharing sites like YouTube.com, wikis, Whois records of registered domain names, metadata and digital files, dark web resources, geolocation data, IP addresses, people search engines, and anything that can be found online.
  - ✓ Traditional mass media (e.g., television, radio, newspapers, books, magazines).
  - ✓ Specialized journals, academic publications, dissertations, conference proceedings, company profiles, annual reports, company news, employee profiles, and résumés.
  - ✓ Photos and videos including metadata.
  - ✓ Geospatial information (e.g., maps and commercial imagery products)

OSINT ENTITIES AND ORGANISATIONS

- ✓ The two government agencies that do OSINT globally are the Open Source Center in the United States and BBC Monitoring in Great Britain.
- ✓ **Open Source Center** (OSC); it is the largest OSINT organization and has vast resources to do its job. OSC works closely with other local intelligence agencies in the United States and offers its services to U.S. government intelligence agencies.
- ✓ **BBC Monitoring** (https://monitoring.bbc.co.uk/login) is a department within the British Broadcasting Corporation (BBC) that monitors foreign media worldwide.
- ✓ **Jane's Information Group** (http://www.janes.com) is a British company founded in 1898. Jane's is a leading provider that specializes in military, terrorism, state stability, serious and organized crime, proliferation and procurement intelligence, aerospace, and transportation subjects.
- ✓ **The Economist Intelligence Unit** (https://www.eiu.com/home.aspx) is the business intelligence, research, and analysis division of the British Economist Group. The main domain of the Economist Intelligence Unit is its business and financial forecasts; it offers a monthly report in addition to a country economic forecast for the coming five years with a comprehensive view about current trends on economic and political issues.

- ✓ **Oxford Analytica** (http://www.oxan.com) is a relatively small OSINT firm compared with the previous two. Oxford Analytica specializes in geopolitics and macroeconomics subjects. It has a global macro expert network to advise its clients on the best practices of strategy and performance when accessing complex markets.
- ✓ **Factiva** (http://new.dowjones.com/products/factiva) is a global news database with licensed content. It harvests data from more than 33,000 premium sources, and many of these sources (74 percent) are licensed and cannot be found freely online. Factiva collects sources in 28 languages in addition to its unique service of being able to provide access to resources that have not been published yet by their creators.
- ✓ **LexisNexis** (https://www.lexisnexis.com/en-us/gateway.page) is currently owned by RELX Group (formerly Reed Elsevier). It originally focused on providing high-quality legal and journalistic documents, but it has expanded its coverage to include more services such as media monitoring tools, supply management tools, sales intelligence solutions, market intelligence tools, and risk solutions that analyze public and industryspecific content to predict risk and improve decision-making.

PARTIES INTERESTED IN OSINT INFORMATION

- ✓ **Government** bodies, especially military departments, are considered the largest consumer of OSINT sources.
  - ✓ The huge technological developments and widespread use of the Internet worldwide have made governments a huge consumer for OSINT intelligence.
  - ✓ Governments need OSINT sources for different purposes such as national security, counterterrorism, cybertracking of terrorists, understanding domestic and foreign public views on different subjects, supplying policy makers with required information to influence their internal and external policy, and exploiting foreign media like TV to get instant translations of different events happening outside.
- ✓ International organizations like the **UN** use OSINT sources to support peacekeeping operations around the globe.
  - ✓ The UN balances superpowers' and emerging nationstates' concerns when creating its policy, which requires it to be as transparent as possible. To achieve this, the UN found that it is more convenient to exploit OSINT sources (including commercial satellite images) for intelligence needs instead of depending on reports from its member states, which may have conflicting policies.
  - ✓ Humanitarian organizations, like the International Red Cross, use OSINT sources to aid them in their relief efforts in a time of crisis or disaster. They use OSINT intelligence to protect their supply chain from terrorist groups by analyzing social media sites and Internet messaging applications to predict future terrorist actions.
  - ✓ NATO depends heavily on OSINT sources for intelligence purposes and for making plans for peacekeeping operations. It also benefits from commercial satellite imagery to plan operations because not all NATO member states have such facilities. NATO has published three standard references about how to exploit OSINT to the public.

- ✓ **LEAs**
  - ✓ Police uses OSINT sources to protect citizens from abuse, sexual violence, identity theft, and other crimes.
    - ✓ This can be done by monitoring social media channels for interesting keywords and pictures to help prevent crimes before they escalate.
    - ✓ Law enforcement uses OSINT to monitor and track a criminal's networks across different countries.
- ✓ **Business Corporations**
  - ✓ Information is power, and corporations use OSINT sources to investigate new markets, monitor competitors' activities, plan marketing activities, and predict anything that can affect their current operations and future growth.
  - ✓ In the past, exploiting OSINT sources was limited to big businesses with good intelligence budgets. Nowadays, with the widespread use of the Internet, small companies with limited budgets can exploit OSINT sources effectively and merge acquired information into their business plans.
- ✓ **Penetration Testers** and Black Hat Hackers/Criminal Organizations
  - ✓ OSINT is used extensively by hackers and penetration testers to gather intelligence about a specific target online. It is also considered a valuable tool to assist in conducting social engineering attacks.
  - ✓ The first phase of any penetration testing methodology begins with reconnaissance (in other words, with OSINT).

(source: http://www.DarknessGate.com)

- ✓ Privacy-Conscious People
  - ✓ These are ordinary people who might want to check how outsiders can break into their computing devices and what their ISP knows about them.
  - ✓ They also need to know their online exposure level to close any security gap and delete any private data that may have been published inadvertently.
  - ✓ OSINT is a great tool to see how your digital identity appears to the outside world, allowing you to maintain your privacy.
- ✓ Terrorist Organizations
  - ✓ Terrorists use OSINT sources to plan attacks, collect information about targets before attacking them (like when using satellite images such as Google Maps to investigate the target location), procure more fighters by analyzing social media sites, acquire military information revealed accidentally by governments (like how to construct bombs), and spread their propaganda across the world using different media channels.

Information Gathering Types

- ✓ OSINT sources can be collected using three main methods:-
  - ✓ Passive
  - ✓ Semipassive
  - ✓ Active
- ✓ **Passive Collection**
  - ✓ This is the most used type when collecting OSINT intelligence.
  - ✓ The main aim of OSINT gathering is to collect information about the target via publicly available resources only.
  - ✓ In this type, your target knows nothing about your intelligence-collecting activities.
  - ✓ This kind of search is highly anonymous and should be done secretly.
  - ✓ From a technical perspective, this type of gathering reveals limited information about the target because you do not send any traffic (packets) to the target server—either directly or indirectly— and the main resources that you can gather are limited to archive information (mainly outdated information), unprotected files left on target servers, and content present on the target website.

- ✓ **Semi-passive**
  - ✓ From a technical view, this type of gathering sends limited traffic to target servers to acquire general information about them.
  - ✓ This traffic tries to resemble typical Internet traffic to avoid drawing any attention to your reconnaissance activities.
  - ✓ In this way, you are not implementing in-depth investigation of the target's online resources, but only investigating lightly without launching any alarm on the target's side.
  - ✓ Although this type of gathering is considered somehow anonymous, the target can know that there is reconnaissance happening if they investigate the issue (by checking the server or networking device logs).
- ✓ **Active Collection**
  - ✓ In this type, you interact directly with the system to gather intelligence about it.
  - ✓ The target can become aware of the reconnaissance process since the person/entity collecting information will use advanced techniques to harvest technical data about the target IT infrastructure such as accessing open ports, scanning vulnerabilities (unpatched Windows systems), scanning web server applications, and more.
  - ✓ This traffic will look like suspicious or malicious behavior and will leave traces on the target's intrusion detection system (IDS) or intrusion prevention system (IPS).
  - ✓ Conducting social engineering attacks on the target is also considered a type of active information gathering.

BENEFITS OF OSINT

- ✓ In today's information age, no one can underestimate the vital role that OSINT plays in the different intelligence arenas.
- ✓ The benefits of OSINT span many areas in today's world. The following are the main ones:-
  - ✓ **Less risky**:
    - ✓ Using publicly available information to collect intelligence has no risk compared with other forms of intelligence such as using spying satellites or using human sources on the ground to collect information, especially in hostile countries.
  - ✓ **Cost effective**:
    - ✓ Collecting OSINT is generally less expensive compared with other intelligence sources.
    - ✓ For instance, using human sources or spying satellite to collect data is costly.
    - ✓ Small businesses with limited intelligence budgets can exploit OSINT sources with minimal costs.
  - ✓ **Ease of accessibility**:
    - ✓ OSINT sources are always available, no matter where you are, and are always up-to-date.
    - ✓ OSINT sources can be used by different parties in any intelligence context; all you need are the required skills/tools to harvest and analyze OSINT properly.
    - ✓ For example, military departments can predict future attacks by analyzing activities on social networking sites, while corporations can use it to build their new market expansion strategies.
  - ✓ **Legal issues**:
    - ✓ OSINT resources can be shared between different parties without worrying about breaching any copyright license as these resources are already published publicly.
    - ✓ Of course, some limitations apply when sharing gray literature.

- ✓ **Aiding financial investigators**:
  - ✓ OSINT allows specialized government agencies to detect tax evaders, for instance.
  - ✓ Many famous celebrities and some giant companies are involved in tax evasion, and monitoring their social media accounts, vacations, and lifestyles has a great value for a government inspector who may be chasing them for undeclared income.
- ✓ **Fighting against online counterfeiting**:
  - ✓ OSINT techniques can be used to find false products/services and direct law enforcement to close such sites or to send warnings to users to stop dealing with them.
  - ✓ This is a great advantage of OSINT, especially when fighting against counterfeit pharmaceutical and natural health products.
- ✓ **Maintaining national security and political stability**:
  - ✓ This might be the most important role of OSINT; it helps governments to understand their people's attitudes and act promptly to avoid any future clashes.
  - ✓ Wise governments utilize OSINT in their future strategies, especially for their domestic policies.

Challenges of Open Source Intelligence

- ✓ All intelligence gathering methodologies have some limitations.
- ✓ Some of the challenges that face OSINT gathering.
  - ✓ Sheer volume of data:
    - ✓ Collecting OSINT will produce a huge amount of data that must be analyzed to be considered of value.
    - ✓ Of course, many automated tools exist for this purpose, and many governments and giant companies have developed their own set of artificial intelligence tools and techniques to filter acquired data.
    - ✓ However, the tremendous volume of data will remain a challenge for the OSINT gatherer.
  - ✓ Reliability of sources:
    - ✓ Bear in mind that OSINT sources, especially when used in the intelligence context, need to be verified thoroughly by classified sources before they can be trusted.
    - ✓ Many governments broadcast inaccurate information to mislead the OSINT-gathering process.
  - ✓ Human efforts:
    - ✓ As we already mentioned, the sheer volume of data is considered the greatest challenge for OSINT collection.
    - ✓ Humans need to view the output of automated tools to know whether the collected data is reliable and trustworthy; they also need to compare it with some classified data (this is applicable for some military and commercial information) to assure its reliability and relevance.
    - ✓ This will effectively consume time and precious human resources.

LEGAL AND ETHICAL CONSTRAINTS

- ✓ Despite the great importance of OSINT, it has legal concerns.
  - ✓ For example, if someone acquires OSINT sources by illegal means to justify an honest case, how should the legal system handle it?
  - ✓ Another dilemma is when the OSINT sample is minimized or selected according to the collector's need. They could effectively discard important sources purposely in favor of bringing about a specific outcome.
  - ✓ Another concern is when some forms of hidden public information are collected and publicized widely as part of a scandal
  - ✓ What is the consequence for such things? What will be the effects on some groups or individuals when revealing such information about them? What are the moral consequences?
  - ✓ Over the past five years, many whistleblowers have stolen classified information from well-guarded agencies and institutions and published it online (Edward Snowden is a clear example).
  - ✓ Should we consider this information belonging to the public source? Of course, military departments around the world will be thirsty for such information, but should we use it—as individuals or companies—as a public source for our intelligence?

- ✓ Many corporations (Facebook and Google are examples) harvest a large volume of online user data for commercial intelligence; most of this data belongs to the user's actions and behavior online and cannot be used to recognize the user's real identity.
- ✓ For instance, there are two types of data that can be collected online:
  - ✓ **Sensitive personal information (SPI)** such as name, Social Security number, place of birth, parents' names, passport or ID number
  - ✓ **Anonymous information** such as technical information like your OS type and version, browser version, IP address, connected device location, and anything that is shared between more than one connected user To justify the collection, these corporations say they acquire only anonymous data, but what if this anonymous information has been combined with other sources to become SPI? How should such information be handled by the OSINT analyst?
- ✓ The final legal concern that we are going to cover is the reliance on **automated machines to collect and analyze OSINT information**.
- ✓ Can we trust the outcome of automated machines and treat it just like the data collected by humans? What if there is a software flaw in the tool that produces inaccurate output that leads to harmful consequences?
- ✓ How we can find a balance between using automated machines, which are necessary in the OSINT-gathering process, and remaining ethical?
- ✓ The limitations of OSINT in addition to its legal constraints should encourage its adopters to follow an individualized and tailored approach when using

REFRENCES

1. Juniperresearch, “CYBERCRIME WILL COST BUSINESSES OVER $2 TRILLION BY 2019” August 25, 2017. https://www. juniperresearch.com/press/press-releases/cybercrimecost-businesses-over-2trillion.
2. Gpo, “Public Law 109-163 109th Congress” August 25, 2017.https://www.gpo.gov/fdsys/pkg/PLAW-109publ163/html/ PLAW-109publ163.htm.
3. CIA, “Intelligence in Public Literature “August 25, 2017.https:// www.cia.gov/library/center-for-the-study-ofintelligence/csi-publications/csi-studies/studies/vol.- 56-no.-1/no-more-secrets-open-source-information-andthe-reshaping-of-u.s.-intelligence.html.
4. Fas, “Final Report of the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction” August 25, 2017. https://fas.org/irp/offdocs/ wmdcomm.html Chapter 1 The Evolution of Open Source Intelligence 20.
5. Gartner, “Gartner Says 8.4 Billion Connected ‘Things’ Will Be in Use in 2017, Up 31 Percent From 2016” August 25, 2017. https://www.gartner.com/newsroom/id/3598917/
6. Comsoc, “IDC Directions 2016: IoT (Internet of Things) Outlook vs Current Market Assessment” August 25, 2017. http://techblog. comsoc.org/2016/03/09/idc-directions-2016-iot-internetof-things-outlook-vs-current-market-assessment