**OSINT Assignment**

You have nine questions in total (in red). Your assignment is in two parts - reading and conducting your own investigation. Number your questions, and make sure that your screenshots are understandable.

**Part 1 – Open Source Reading**

**(35 pts.)**

Read Chapter 1 from [*Open Source Intelligence Methods and Tools: A Practical Guide to Online Intelligence*](https://drive.google.com/file/d/1vWxczgnCo-ThUYDLY3avBGzhUgHhtBRs/view?usp=sharing), by Hassan and Hijazi, and answer these questions:

1. **Describe the differences between *data*, *information*, and *knowledge* and give an example of each (your reading gives examples, but provide your own.)** **10 pts.**Data is considered to be a collection of truths conveying something without further clarification or examination. An example of data is “98% of the students passed test.” Information is defined as a kind of data that has been interpreted accurately to provide a valuable significance within a specific context. An example of information is Compared to last week, the percentage of students that passed the biology test grew from 86% to 98%. Knowledge is considered to be a blend of information, experience, and insight that has been learned or deduced after some experimentation. Knowledge clarifies what your mind has documented in the past, and these records can aid in constructing sounder determinations concerning the future when encountering similar contexts. An example of Knowledge would be compared to last week, the percentage of students that passed the biology test grew by 12% since the last test only had a passing rate of 86% because of study sessions afterschool, enabling students taking this course to have a higher grade by the end of the semester.
2. **Describe the difference between *gray literature* and *gray information* and give an example of each (your reading gives examples, but provide your own.)** **10 pts**. Gray literature is when all publications can be acquired through traditional approaches such as going to bookstores or paying for a subscription; its anything that is published in public. An example of gray literature is speeches. Gray information people cannot acquire publications that differs from gray literature publications from traditional methods such as utilizing specialized services. An example of gray information is personal emails.
3. **Your text lists several benefits of OSINT. Select the one you think is most beneficial and discuss why**. **5 pts.** One benefit of utilizing Open Source Intelligence (OSINT), is that it is considered to be less risky than other intelligence gathering software or methods. When utilizing it for investigative purposes, users will not have to encounter much difficulty since sensitive information while using the software will not be manipulated or collected by unauthorized individuals or groups.
4. **Your text lists three challenges of OSINT. Select the one you think is the most challenging, and discuss why.** **5 pts.**  One challenge of utilizing Open Source Intelligence (OSINT) is that efforts made for users to utilize it are too time-consuming. This will be a long process that will make an investigation last longer. OSINT generates vast quantities of data, and it is the investigator or users’ responsibility to view each of data in order to verify that it is reliable and/or compare it to other confidential data for military and commercial purposes to guarantee its significance too.
5. **Suppose a whistleblower had a video about our military killing innocent civilians in another country, and they post that video to a public website. Should that information be used by other governments, researchers, and writers? Why or why not.** **5 pts**.  If a whistleblower has a video of our military killing innocent civilians in another country, and they decide to post it to a public website, governments, researchers, and writers have the right to utilize that information because it was posted on a public website and it is solely based on the content provided by and behavior displayed by the whistleblower.

**Part II – An Open Source Investigation**

Background Information

All OSINT investigations start with some question(s) that need to be answered about a target (suspect in a crime, a person of interest, a potential employee, etc). As Pastor Galindo et al. (2020) show in the diagram below (Figure 1), some of these questions can be (1) political, sexual, and religious preferences, (2) tendency to crime, (3) cached life (collecting and storing a target’s life events), (4) activity on the web, and (5) places visited. This infographic from Pastor-Galindo et al. (2020) can help you organize your thinking about collecting OSINT (as well as your Part 1 reading). These three phases are collection, analysis, and knowledge extraction.

|  |
| --- |
| **Figure 1**  **OSINT Workflow and the Intelligence Derived from It** |
| Diagram  Description automatically generated |
| “The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends”, *IEEE ACCESS,* JAVIER PASTOR-GALINDO , PANTALEONE NESPOLI , FÉLIX GÓMEZ MÁRMOL , AND GREGORIO MARTÍNEZ PÉREZ. 2020. |

In the **data collection phase**, there are at least eight possible techniques for finding information on someone: (1) real name, (2) user name, (3) search engines, (4) domain name, (5) IP address, (6) location, (7) social networks, and (8) email address. There is also a ninth technique not discussed by Galindo et al, and that is using a (9) image. In other words, we may have a person’s real name and that is what we use to find more data about that person. The person’s name is data to collect, but the name is also used to find other data - via data collection tools. So a name leads to an IP address, which then leads to a domain name, and so on. Imagine all you have of a target is a nickname -  “BuckeyT.” You start with this bit of data and try and collect more.

As you try and learn more, experimenting with tools is integral to this phase. Different targets have different types of information available online and some tools will be more beneficial than others. In this phase, you are collecting as much data as possible. Say an investigator has only a nickname of a leader of a local gang. In this phase, the investigator must collect additional information. As you move through this phase and your tools give you accurate pieces of data, you can start to move into the phase where you are analyzing the data.

In the **analysis phase**, the investigator tries to organize relevant data around three different entities: (1) people, (2) organizations, and (3) networks. Cybercrime investigators are most likely focusing on people and networks of people (to investigate them), while cybersecurity professionals are most likely focusing on networks of computers and organizations (to protect them). The idea in this phase is to start sifting through your data and putting the data together in a coherent way. You want to move from data to information. Through a search on social media, you find many people with that username. One of the users with that username has posted images that look like where “BuckeyT” operates. You are certain that this profile is the person you are interested in and so you then go and collect more data: who his associates are, what he posts about, etc. Finally, you have enough to get a real name and even more information. A good OSINT investigation may end up with pages of text, images, known associates, contacts, financial information, etc.  You don’t want all of this information – instead you want the information that matters for you or your employer.

Finally, you move from information to knowledge in the **knowledge extraction phase**. Taking the information you have and answering a question or coming to a conclusion about something is producing knowledge. Using our gang member “BuckeyT”, you may see that he associates with several hardened characters who have been arrested in the past for dealing heroin. Moreover, you have been able to find his real name and address – Brad Jenkins – and see that he dropped out of high school but has been seen around town in an expensive car. You conclude that BuckeyT is likely the leader of a drug trafficking enterprise. You have produced this knowledge through passive OSINT techniques that did not alert Jenkins to the investigation or put yourself in danger.

Collection Tools

There is a vibrant open-source intelligence community. You can learn more about this community (and most of the tools used) through the OSINT Open Source Intelligence Framework Landing Page - <https://start.me/p/ZME8nR/osint>, which looks like this:

|  |
| --- |
| **Figure 2**  **Screenshot of OSINT Landing Page** |
| Graphical user interface, text, application  Description automatically generated |

If you are thinking of breaking into the field, it may be worth your while to follow some of the professionals interested in OSINT:

|  |
| --- |
| **Figure 3**  **OSINT Professionals on Twitter** |
| Text  Description automatically generated |

There is a wealth of tools available for an OSINT analysis. I am showing the landing page and people to follow for OSINT-curious students who wish to dig deeper on their own time.

For this class, we will focus on a wonderful website made by Justin Nordine (@jnordine) called the “OSINT Framework”: <https://osintframework.com/>, shown in Figure 4.  The site organizes many tools you can use in an OSINT investigation.

|  |
| --- |
| **Figure 4**  **The OSINT Framework** |
| Diagram, schematic  Description automatically generated |

OSINT investigations are done to answer one or more questions about a person or suspect. However collecting personal information about a person – even if the information is public, and even if it is for educational purposes – is still not entirely *right*. Many people unrelated or not connected to that person would then be aware of that person’s contact information, their social media presence, and their life cache.

So this is how we will handle that problem:

* You will choose the person you will investigate. Make it a friend or a family member. Famous people will likely have their information scrubbed or hard to find, so you may want to avoid doing OSINT on them. However, you are welcome to try. You will refer to this person throughout as “the target”
* When you take screenshots (you will need to document your work) you will blur or redline the personally identifiable information you wish to hide.

These steps will allow you to demonstrate your investigative skills without sharing too much information about your family or friends.  Your assignment is in four parts: Introduction (10 pts), Data Collection (20 pts), Data Analysis (20 pts), Knowledge Extraction (10 pts).

Introduction

The first step is to think in your mind about **who** you want to investigate (the target). Now, think of a reason **why** you want to investigate them. This reason can be thought of as the question you want to answer in your OSINT investigation. What is that question or what are those questions?

 Possible questions are:

* Is the person financially stable?
* What is the political ideology of the person?
* Do they have a lot of friends?
* Does the person show good character?
* Is the person cheating on me?
* What is their employment history?
* Does this person have a criminal record?

1. **In your introduction, you need to write a paragraph that describes who you want to investigate (again, no names), why you want to investigate them, and what is the question you want to answer.** **(10 pts)** The target for my investigation is a cousin from my maternal side. The basis for this investigation is to know their political affiliations and have a deeper understanding behind their political beliefs. It could potentially steer me away from difficult topics that may cause uncomfortable discourse. Before I start an Open Source Investigation (OSINT), I am wondering if my target is a Republican, Democrat, Socialist, or Independent since some of our ideologies are similar in some aspects but vastly different in others. The question that relates to my investigation is “What is the political ideology of the person?”

Data Collection Phase

Start the data collection phase by using tools from the OSINT framework ([OSINT Framework](https://osintframework.com/)). As an example, **Figure 5** shows three separate tools - one collecting data from a social network, one from public records, and one from people search engines.

**Figure 5**

A picture containing text

Description automatically generated

1. **To complete this section use six different tools to collect information about your target. Name the tool (e.g. “That’s Them”, “Followerwonk”, “Pipl”), take a screenshot of each tool as you are collecting information, and then discuss some of the information you have collected with the tool.** **(30 pts)**

|  |
| --- |
| Graphical user interface, application  Description automatically generated |

I utilized the people search engine, “True People Search” in order to find some information on my target such as their address, phone number, and date of birth.

|  |
| --- |
| Graphical user interface, application  Description automatically generated |

I utilized the public record website, “Info Tracer”, to find my target, Kyla’s email address.

|  |
| --- |
| Graphical user interface, application  Description automatically generated |

I utilized the email verification software, “Simple Email Reputation” and found that the email account was linked to Twitter and was compromised by a hacker for awhile.

|  |
| --- |
| Graphical user interface, text, application, chat or text message  Description automatically generated |
|  |

I utilized social networks to find the account of my target, Kyla, and found the people she was following such as friends, family, and political leaders.

|  |
| --- |
| A screenshot of a computer  Description automatically generated |

I utilized the username search engine, “Instant Username Search” to see that the cites Facebook, Instagram, Patreon, and Facebook have the username I typed in.

|  |
| --- |
| Text  Description automatically generated |

I utilized another social media platform “Facebook”, and found a repost of a former Supreme Court judge that passed away in 2020.

 Only **two** of these tools can be from the social network branch (Twitter, Facebook, Instagram, Tiktok, Snapchat, or YouTube.) If you have questions about this, please email me.  However, you can use more than one of the other types of tools (people search engines, dating sites, business records, etc.).

Remember, you want to anonymize the information (Figures 6 and 7). Then, you want to describe below the screenshot precisely what information you have collected from the tool. So, for Figure 6 you want to talk about whether it is the IP address, the email address, the phone number, or all three that you are collecting. Meanwhile, for Figure 7 (collected using Followerwonk), you may discuss the number of followers in a specific country. Some tools can give you a massive amount of data. For example, browsing someone’s Facebook profile will give you a wealth of likes, friends, etc. Some tools will give you a smaller amount - may be just a few lines about phone numbers and prior residences.

Data Analysis Phase

Now analyze the data. Remember, you can collect pages of data about someone. You can collect all their social media posts, all the places they’ve lived and the history of those places, all their relatives or associates, etc. But you need to turn that data into information by selecting the most relevant data linking them together in a summary.

1. **Given what you have collected, write a two-paragraph summary of the information you have collected.** **(15 pts.)**

The screenshots above display the process of me gathering intelligence to find what political party that the target, Kyla, follows. While using the search engine, I went to the website “True People Search” in order to find some information on my target such as their address which was 9720 57th Ave #3b, Corona, NY, 11368 and her most recent phone number (347) 642 5739. I also found her date of birth, October 1996. I utilized information that I obtained from “True People Search” by typing the target’s phone number of the target. Although there were many people that had a similar last and first name, I was able to match the information such as her address, date of birth, phone number on the public record website, “Info Tracer”, to find my target, Kyla’s email address, kylajohnson1012@gmail.com, This email address had the month of her birthday and presumably the date. The process above allowed me to gather the information I needed for the following steps.

The next steps enabled me to discover the answer to my questions. When I utilized the email verification software, “Simple Email Reputation” I put the information I obtained “True People Search” and gathered some intelligence on if the email, , kylajohnson1012@gmail.com was linked to an account that had been breached. I was able to see that the email was directly attached to a Twitter account that was compromised in 2018. While looking at who Kyla was following, I found evidence that would help come to conclusion about her political beliefs, such as following the Governor of New York, Chris Cuomo. I also utilized the username search engine to find another piece of direct evidence by using the software, “Instant Username Search” to see that the cites that had Kyla Johnson as their profile were on social media sites Facebook, Instagram, Patreon, and went to Facebook since I am unfamiliar with the other platforms. While typing her name in the search engine I was able to look through Kyla’s page, I found that there were some posts that were political. In one of her older posts, she had a photo of Ruth Bader Ginsburg that clearly showed what political ideology that she is more inclined to follow.

Knowledge Extraction Phase

Now turn the information you have collected into knowledge. Answer the question you posed at the beginning of your investigation in a few sentences.

1. **Explain how the information you collected helped you answer that question.** **(10 pts.)**

My question was about what political affiliation is my target, Kyla, inclined to. Throughout the process, I determined that Kyla is a Democrat. This was proven since the person she follows multiple times is Chris Cuomo is the former Democratic Governor of New York as known affiliations to another liberal platform CNN. Also, when I browsed through several posts on Facebook, I found a photo of Ruth Bader Ginsburg, a Supreme Court judge that was appointed while a Democratic President was elected to Office. Most of her court room decisions are based on Democratic ideologies such as Civil Rights.