DETERMINING THE VALUE OF OPEN-SOURCE INTELLIGENCE FOR PUBLIC SAFETY

GIOVANNA CIOFFI

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN COMMUNICATION AND CULTURE
YORK UNIVERSITY
TORONTO, ONTARIO

September 2023

Abstract

The intent of this work is to determine the regard to which open-source intelligence (OSINT) is an effective tool for emergency management, especially in relation to public safety. This work seeks to accomplish this through meeting the following objectives: (1) examining OSINT from a public safety perspective, (2) identifying potential challenges and barriers that may limit an analyst's use of OSINT tools and techniques; (3) exploring the changing nature of threats to national security and identifying how OSINT may provide a direct means of assisting with mitigation, prevention, preparation, response, and recovery; and, (4) by understanding how government analysts are training in OSINT collection and methodologies. The methodological approach to this research is qualitative in nature, focusing on case studies, tool exploration, and Access to Information and Privacy (ATIP) requests. Although the results of this work concluded that OSINT can be regarded as an effective tool in maintaining public safety, it also raised concerns regarding legislation, policy, training, and technical infrastructure that must be addressed if OSINT were to remain effective. All this in consideration, the results of this work were impacted by limitations in access to information. Given the sensitive nature of some collection procedures and the overstrained ATIP request portal, a number of documents were not made available for evaluation and analysis - likely a result of classification level, security requirements, or the overall time it takes for ATIP coordinators to make these documents available for public disclosure. Moreover, in consideration to the results of this work, two main recommendations were provided: (1) *Legislation & Policy*: clarification of legislation, effective policy development, ongoing communication, and oversight; and (2) *Training & Technical Infrastructure*: establish clear OSINT tiers, establish mandatory training plans, and establish standardized methods to account for attribution. Together, these two recommendations can further strengthen the OSINT capability as a whole and ensure that it continues to be an effective tool for public safety.

Dedication

To my family and friends. You have always been there, no matter what.

Thank you.

Acknowledgements

Success is never achieved in isolation. If I have ever succeeded, it has been by standing on the shoulders of giants. This work would not be possible without the patience, assistance, and support of my supervisor and dissertation committee: Dr. Livy Visano, Dr. Walter Perchal, Dr. Lorne Foster and Dr. Jay Goulding. Words cannot express my gratitude for your guidance and all you have done for me. I will carry your wisdom and guidance with me on every journey hereafter. I only hope that one day I may be able to impart on others what you have on me. A special note of gratitude is extended to Dr. Kevin McCormick and Dr. Awalou Ouedraogo, respectively serving as External and Internal Examiners. My heartfelt appreciation for the time and commitment you have generously invested.

I am also grateful to my colleagues, classmates, and friends for their moral support, feedback, and editing. Paul Hassanally, thank you for your formatting assistance. Without you, I'd be shopping for a new computer. Chris Oh, thank you for checking in and helping me to flush out ideas. You have a way of calming souls. Daryoush, thank you for being who you are. Never change. Tiffany, without you this small town would be lonely. Thank you for all the random pop-ins, chats, and people you've brought into my life.

Finally, to my family who has given me the space to do all I needed to in order to make this possible. Your belief in me has kept me going through this journey. I would also like to thank our two dogs, Fudge and Pio, who – in exchange for treats – guarded the doorway while I worked and reminded me when to take breaks. We will deeply miss you both. The love and joy you have brought into our lives will never be forgotten.

Table of Contents

List of Tables

[1] STANAG: Standardization Agreement

List of Figures

Chapter 1: Introduction

1.1. Background

Open-source intelligence, commonly referred to as OSINT - is an intelligence discipline that has gained much attention over the last two decades. This is especially the case with the advent and popularity of social networking websites, which have changed the ways in which people interact and the information that they consume. As these social networking platforms have enabled the exchange of communication and information globally, they also began to be used as recruitment tools and a means to incite violence around the world. As this trend became more apparent, there was a requirement for the intelligence community to evolve its OSINT capability to focus on online collection that would disrupt extremist intentions and cater to public safety.[2]

Although OSINT has become a rather contentious topic to discuss, there is limited research on how it may be employed to contribute to public safety within a Canadian context. Additionally, OSINT as a whole has suffered negatively as a result of government agencies and departments using the capability inappropriately (Tunney, 2019). Accordingly, the rationale of this work seeks to acknowledge past assertions and evaluate the capacity of OSINT to contribute to public safety. The case studies herein facilitate an enhanced understanding of public safety and they equally contribute to the formulation of much needed recommendations.

1.2. Motivation - Statement of Purpose

The purpose of this work is to understand the extent to which open source intelligence may contribute to public safety. This work will attempt to discover how various tools and techniques may be used to aid analysts in the intelligence community. Additionally, this inquiry seeks to secure information on training and policy documentation from various government agencies and departments in the form of Access to Information and Privacy Requests. The evidence therein will advance an

[2] It is important to note here that the mention of public safety in this work is largely in reference to the intelligence community's efforts to counter terrorism and crime in general.

understanding of how analysts within the intelligence community are trained to use OSINT and how policies guide their actions.

1.3. Significance of the Research

As a whole, this work is significant in that it is contributing to an area of study that has received limited attention. Therefore, the results of this work have the potential to ground an understanding of how open source intelligence may be used to benefit public safety as it directly and indirectly relates to Canada and Canadians. Further to this, the results of this work also have the potential to effect change in how OSINT is conducted, especially in consideration to ethics, policy, and legislation. Moreover, it is within the scope of this work to identify issues with the employment of OSINT and make recommendations to counter those issues and increase its effectiveness, thereby having the potential to impact the capability as a whole.

1.4. Problem Statement

Consistent with the statement of purpose, this particular study is needed to facilitate an understanding of how / if OSINT can be used to contribute to public safety. Currently, the literature available focussing on the Canadian context in consideration to Canadian policy and legislation or ethics and training is woefully limited or non-existent. Additionally, it has been noted that there has been considerable criticism about the intelligence community's efforts to collect information from open sources, but research regarding open source's impact on the intelligence community and public safety as a whole - especially within a Canadian context - is quite limited. Therefore, this work will attempt to alleviate this problem by considering the use of open source intelligence for public safety, within the Canadian context.

1.5. Research Aim, Objectives and Research Questions

Research Aim: To understand how OSINT may contribute to public safety.

Research Objectives:

Objective 1: To examine OSINT from a public safety (e.g. national and international security) perspective.

Objective 2: To identify potential challenges and barriers that limit an analyst's use of OSINT tools and techniques.

Objective 3: To explore the changing nature of threats to national security, and identify how OSINT may provide a direct means of assisting with mitigation, prevention, preparation, response, and recovery.

Objective 4: To understand how analysts are training in OSINT collection and methodologies.

Research Question:

To what extent is OSINT an effective tool for emergency management, especially in relation to public safety?

Given its broad range of applicability, is OSINT an effective capability to support public safety, especially in combination with other intelligence disciplines? However, if legislation is not in place to address issues pertaining to the collection, use, retention, and dissemination of intelligence derived from OSINT, then will this capability inevitably turn into a liability for government agencies, and the people they intend to protect?

The thesis is addressed by answering the following questions:

Sub-Research Questions

1. What is OSINT, and how might it be useful to mitigating threats against public safety?

   The first sub-research question aims to understand what OSINT is and what it is not. Once this has been clarified, understanding OSINT methodologies and frameworks will assist in identifying how it may prove useful to mitigating and responding to threats against public safety.

2. What is the current value of OSINT?

   The second sub-research question seeks to understand the current value of OSINT in detecting emerging and ongoing threats. This question will be answered by exploring multiplied advanced OSINT strategies and tactics, in addition to observing how they have been used in an investigation.

3. What are some current and emerging capabilities?

   This sub-research question will discuss emerging capabilities as they directly relate to the use of OSINT. Such capabilities will include, but not be limited to social networks, online maps and geolocation tools, exploitation techniques, the surface, deep, and dark web, image analysis, and big data analytics.

4. What may be some deterrents to progress?

   This question will address some of the deterrents associated with utilising OSINT techniques as an investigative tool and methodology for identifying threats to public safety. Discussion topics will include training and tradecraft barriers, cultural deterrents, resourcing, and partnerships. Additionally, this question will also include consideration to the risk associated with not utilising OSINT as a means of providing public safety.

5. What are the legal, ethical (training + mental health), and policy challenges to employing OSINT?

   Challenges exist throughout processes and organizations. In order to determine if the use of advanced OSINT techniques impedes or strengthens public safety, it is important to address the challenges that may impact analysts and organizations as a whole. This question will therefore examine legislation that governs how the public and private sector may utilise publicly available information. Additionally, it discusses training, mental health (i.e. exposure to traumatic content), ethics, and policy regulations as they are relevant to using OSINT in the interest of public safety.

6. What does the future of OSINT look like?

   This sub-research question is highly dependent on what the future of the web evolves into. This evolution will not only trigger an evolution of collection strategies and tactics, but will also require a reassessment of the policies and authorities in place to regulate the collection and storage of data. This question aims at addressing what role OSINT will have in the future of emergency management and public safety.

1.6. Overview of Chapters

Chapter 1: Introduction. This chapter will introduce the topic to the reader, specifically discussing the background and motivation of the work. Additionally, the problem statement, research question, significance, thesis overview, and research objectives will be identified here.

Chapter 2: Literature Review. This chapter will provide an overview of the current literature on OSINT, within the scope of this work. Specifically, it will provide a summary of what open source intelligence (OSINT) is, especially as it relates to publicly available information (PAI). From here, a discussion on the current value of OSINT will be made by acknowledging how the Internet and World Wide Web are ever present in our daily lives. Once this is established, a review of the literature pertaining to the capabilities of OSINT will be conducted, specifically as it relates to social networks, mapping software, the different layers of the Web, and big data analytics. After this, the literature review will then discuss the deterrents to progress that OSINT may face as it is being employed as a capability. This discussion will include the risks in the utilisation of OSINT, as well as the advantages and disadvantages of using OSINT. Furthermore, the literature review will close with a final section of the legal, ethical, and policy limitations for OSINT collection and use, which will include a review of relevant legislation, in addition to pertinent ethics and policy documents.

Chapter 3: Theoretical Framework. This chapter will discuss the theoretical framework in which this work is viewed. As such, it will introduce concepts of Network Theory, and more specifically Social Network Analysis. After key concepts are provided, this chapter will then evaluate how the theory fits into this work, and will then discuss its specific relevance to the research.

Chapter 4: Methodology. This chapter will discuss the methodology and workflow of this work. In particular, this chapter will discuss the methodological approach, the method of data collection, the method of analysis, and finally the evaluation and justification of methodological choices undertaken to answer the research question.

Chapter 5: Results and Discussion. As a result of the qualitative nature of this work, this chapter has combined both the results and discussion into one chapter. Portions of this chapter have been separated into sections in order to organize and isolate the case studies from one another for better clarity. The case studies in this chapter focus on using OSINT for a particular task or output, including situational awareness and horizon scanning, locating a person, group, or location, verifying information, social network analysis, and access to information and privacy (ATIP) requests. This chapter concludes with a discussion and concluding remarks, including a section on recommendations for employing OSINT.

Chapter 6: Conclusion. This chapter will summarize the main research question, the overall argument, and key takeaways. Additionally, it will also reiterate any important limitations of the research and provide relevant recommendations. In closing, this chapter will explain how this research has contributed to knowledge in the field, and will also discuss the future of OSINT.

Chapter 2: Literature Review

2.1. Introduction

Since its inception, open source intelligence has evolved to encompass numerous meanings to its practitioners, critics, and bystanders. These meanings have developed alongside the advances in technology and tradecraft in various sectors spanning academia, national security, social interaction, and marketing, to name a few. As its meaning has evolved, so has its capacity to become a powerful research method to collect and produce a plethora of data and information to develop knowledge toward any given subject matter of one's curiosity. For this reason, open source intelligence may be viewed as a knowledge enabler for those who regularly digest large amounts of data and information.

On account of this work's scope being focused on public safety, this section will serve to administer a review of the current literature available as a means to provide a preliminary understanding of the regard to which OSINT may be considered an effective tool for public safety. In consideration to its various interpretations, Section 2.2 of this chapter will review what open source intelligence is - especially as it relates to other intelligence disciplines - and how it is collected. Followed by this, Section 2.3 will review the current value of OSINT principally in consideration to the increased digitization of society and how nefarious actors have taken advantage of society's increased interaction online. Section 2.4 will then review the capabilities that enable OSINT to be a powerful investigative methodology, and Section 2.5 will review the deterrents to progress that both challenge and limit the implementation of OSINT in the intelligence community. Finally, Section 2.6 will review the relevant legal, ethical, and policy documents surrounding OSINT.

At this time, open source intelligence is still a relatively new topic of exploration within academia, with the vast majority of peer reviewed literature occurring in the last ten to fifteen years. In consideration to this, the timeliness of this topic is on par with the development and modernization of policies and regulations that govern how both the public and private sectors interact with and utilise publicly available information and data. In view of this, there remain numerous gaps in literature that have either not been addressed or are limited in quantity and scope. The literature review aims to identify those gaps, highlight relevant discourse on this subject, and critically examine the current

research on open source intelligence with a view to situate a baseline knowledge on the matter at hand.

2.2. What is OSINT?

Open source intelligence - more commonly known as OSINT - is an intelligence discipline that "addresses the search, collection, processing, analysis, and use of information from open sources that can be legally accessed by any individual or organization" (Evangelista, et al, 2020, 347). This definition is consistent with how Williams and Blum (2018) describe OSINT. However, they add an additional factor that is key to intelligence analysis: the importance of addressing a specific intelligence requirement, which defines both the need for the collection of information, and the production of intelligence (Williams & Blum, 2018, 8; Arena, 2016, 1). Within the discipline of OSINT, there are two key terms that, while often used interchangeably, are quite distinct from each other. These terms are "open sources" and "publicly available information." To distinguish between the two, their definitions are as follows:

> Open Sources: This includes information that is publicly available, but is not necessarily information that should be "open," or that one has intended to share past a specific cadre of people.
> Publicly Available Information: Sometimes abbreviated as PAI, publicly available information includes any "data, facts, instructions or other material published or transmitted for general public consumption; legally observed by any individual." (U.S. Department of the Army, 2012, 1-1).

While these two concepts connote the two main terms of data and information within OSINT, it is important to note that OSINT, alongside other intelligence disciplines, rarely occur in isolation of each other. In other words, speaking in terms of definition, intelligence disciplines "are sometimes driven more by the unique regulatory authorities of intelligence collection agencies than by the distinct differences between the collection methods or the material itself" (Williams & Blum, 2018, 7). Moreover, as technology and collection methods continue to develop, the marked differences between each discipline will increasingly blur.

Perhaps one of the most frequent instances of disciplinary overlap occurs while utilising OSINT collection methodologies. As depicted by Williams & Blum (2018) in Figure 2.1, OSINT

often overlaps with a number of intelligence disciplines. For example, as commercial satellites continue to develop the capability to provide overhead imagery similar to classified collection platforms, MASINT,[3] GEOINT,[4] and OSINT will have more overlap. Additionally, as OSINT continues to be used in conjunction with social media and advertising technology (Adtech), overlap will occur with the collection methodologies of both HUMINT and SIGINT. Similar to HUMINT, the data collected from social media or Adtech may situate insight from a particular individual or group with unique access. Or, similar to SIGINT, data collected from those same sources may involve electronic collection of a number of records to identify critical interactions and communications that are of interest (Williams & Blum, 2018, 8).

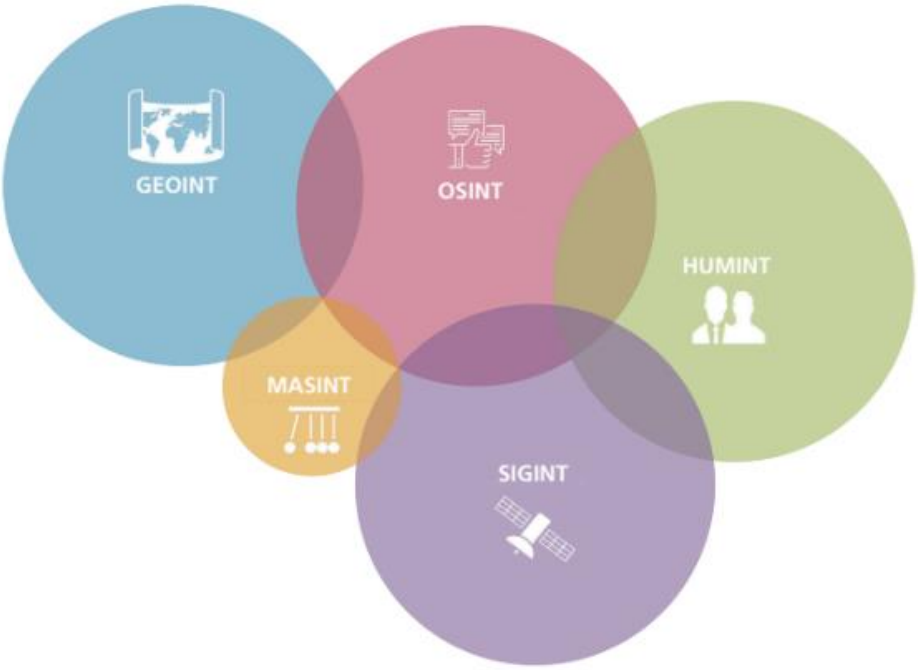

Figure 2.1: The Overlapping Nature of Intelligence Disciplines
Source: Williams & Blum, 2018, 9
Note: The image is not intended to show *all* the ways in which intelligence disciplines overlap. Instead, it is designed to depict the blurring lines between the different intelligence disciplines.

Open source intelligence has been used in a variety of ways by a number of agencies. While some may not consider it to be a major part of their collection methodology, others acknowledge its increasing significance. According to Schaurer & Storger (2013), for the Central Intelligence Agency

[3] "Measurement and Signature Intelligence (MASINT) is scientific and technical intelligence obtained by the qualitative analysis of data (metric, angle, spatial, wavelength, time dependence, modulation, etc.)" (Department of National Defence, 2001, 7).
[4] Geomatics intelligence, sometimes referred to as geospatial intelligence.

(CIA) open source intelligence during the Cold War "became the leading source of information about the adversaries' military capabilities and political intentions, including any warning and threat forecasting" (Schaurer & Storger, 2013, 53). During the Cold War, OSINT was already a well-established methodology where security officials analysed magazines, books, newspapers, radio, and TV broadcasting daily as a means of solidifying the targeting for their collection efforts. Conversely, with the emergence of the Internet and the World Wide Web, most intelligence communities remained hesitant in their appreciation of information emanating from these sources for two reasons:

> "(1) Intelligence agencies seek an informational advantage through covertly dealing with secrets. Relying on open information and its respective copyright restrictions runs counter to that idea.
> (2) In most cases it is more difficult, risky and expensive to apply clandestine methods in order to acquire secret sources, thus giving the impression that those sources must be of higher value than open sources, confusing the method with the product or mistaking secrecy for knowledge" (Schaurer & Storger, 2013, 53-54).

While these may be legitimate claims of concern, the two reasons above expose an anchoring bias that has historically set an over-reliance on what has traditionally been more familiar (classified collection), thereby under-valuing unclassified collection. In spite of this, open source intelligence has come a long way since the Cold War, and has contributed immensely to situational awareness, security investigations, and public safety.

Contrary to Schaurer & Sorger's (2013) claim, Cameron Colquhoun (2016) has argued that post World War II, OSINT as a discipline had largely become a forgotten and "unfashionable" term up until the mid-2000's, and was seen as a less desired discipline especially when compared to "the secret world of agent-running and signals intelligence" (Colquhoun, 2016, np). Colquhoun (2016) further argues this point by illustrating how in the context of government analysis, the discipline of OSINT appeared to not have any true effect during the Cold War or 9/11. However, during the Green Revolution of 2009 - which saw an increase in non-violent protests for more democratisation in Iran - suddenly everything changed; millions of young Iranians coordinated their activities over the Internet, shared viral content, and encouraged people to support or join their cause. Internet use in Iran had increased "from 34% in 2008 to 48% in 2009 … mobile phone subscriptions went from 59% to 72% of the population … [and] 60% of posts on Twitter referred to Iranian politics" (Colquhoun, 2016,

np). Suddenly, the internet was surging with information with what would later be termed *citizen journalism*.

Nearly two years later, the Arab Spring would sweep through the Middle East and North Africa (MENA) region and would lead to fundamental changes in the political systems of a number of countries. According to an Associated Press report, the Central Intelligence Agency's Open Source Center (CIA OSC) - set up in response to the events of 9/11 - "began paying particular attention to social networking websites in 2009, when Facebook and Twitter emerged as primary organizing instruments in Iran's so-called 'Green Revolution' " (Fitsanakis, 2011, np). According to Doug Naquin, the Open Source Center's Director at the time, the CIA OSC had "predicted that social media in places like Egypt could be a game-changer and a threat to the regime, but had been unable to foresee the precise development of Internet-Based social activism in the Arab world" (Fitsanakis, 2011, np; Central Intelligence Retirees' Association, 2007). On this note, Colquhoun (2016) argues this failure to be a result of an obsession with collecting intelligence from the "powerful elite" as opposed to exploiting troves of open source data (Colquhoun, 2016, np). Moreover, both Fitsanakis and Colquhoun insinuate a common theme in their work: that although social media platforms like Twitter and Facebook did not directly *cause* either the Green Revolution or the Arab Spring, they certainly contributed to the spread of revolutionary aspirations throughout the MENA region.

In their article discussing cyber intelligence and OSINT mitigation techniques against cybercrime threats on social media, Yeboah-Ofori & Brimicombe (2017) argue how OSINT is becoming increasingly vital in the modern war on terror. Their systematic research found that the use of OSINT - especially as it relates to social media - is vital in understanding terrorist network topologies, mining and analysing crime data, and furthermore, countering improvised explosive devices. This is of particular interest, especially as it relates to the collection of OSINT in the name of public safety. Similar to the role social media had in both the Green Revolution and the Arab Spring, its significance in contributing to public safety has proven its value today. For example, at the time of writing, multiple social platforms are being used to track Russia's invasion of Ukraine, especially with regards to troop movement, war crime investigations, and disinformation campaigns.

Echoing the previous statement from Yeboad-Ofori & Brimicombe (2017), Adams (2022) argues that "OSINT enables government agencies to collect and analyse information from the Internet relating to their criminal investigations, infrastructure and action planning, and tackling misinformation" (Adams, 2022, np). In regards to collection, Nihad Hassan (2018) identified three types of OSINT collection: passive, semi-passive, and active. As the most used type of collection, Hassan (2018) described the passive method to be one which collects information strictly through the use of publicly available resources. Additionally, semi-passive collection is slightly more technical in that it "sends limited traffic to target servers to acquire general information about them" (Hassan, 2018, np). Finally, the active method involves direct interaction, and presents a higher risk as there is an increased chance the target may become aware of reconnaissance efforts via a possible intrusion detection system, intrusion prevention system, or through other counterintelligence (CI) efforts (Hassan, 2018, np).

Conversely, the Royal Canadian Mounted Police (RCMP) broke down the use of open source collection into a three-tier structure that provides some additional clarity on OSINT collection methodology. They are as follows:

> "Tier 1: overt online activities (e.g. research, environmental scanning, inquiry, public engagement, etc.).
> Tier 2: passive online activities for investigations and intelligence performed in a manner that minimises identification to law enforcement (e.g. discreet, no engagement with a subject of interest).
> Tier 3: covert online activities in a manner that is explicitly concealed and safeguarded to prevent identification to law enforcement (e.g. engagement with a subject of interest is permitted with authority)" (Royal Canadian Mounted Police, 2021, np).

While some similarities between Hassan (2018) and the Royal Canadian Mounted Police's (2021) tiers of open source collection exist, there appear to be a number differences. While Hassan's passive collection and the RCMP's Tier 1 method both share similarities, differences begin to emerge thereafter. Both of Hassan's semi-passive and active collection methods are indicative of higher risk factors through technical and social interaction, while the RCMP 's second and third tier appear to exercise more safeguards, which serve to protect the collector and the agency itself. Additionally, a degree of vagueness is apparent especially in Hassan's description which mildly suggests an element

of cyber operations. On the contrary, the RCMP appears to focus efforts on managing attribution. Overall, the varying levels of OSINT collection tiers within the literature are indicative of a lack of standard terminology across both the public and private sectors. Moreover, this presents an opportunity of improvement for the discipline itself on the grounds that if both the public and private sector share a standard understanding for what type of collection methodologies are acceptable, then a clear and ethical process of collection can occur. The focal point for this would be to create a clear, living policy document which outlines what *is* and *is not* permitted, and that can be updated regularly alongside the evolution of collection methodologies, technology, and the World Wide Web.

Given the parameters of Hassan's (2018) passive collection and the Royal Canadian Mounted Police's (2021) Tier 1 collection methodology, this is likely to be one of the most common types of OSINT conducted. This is because it allows an analyst to gain as much information as possible without generating a high risk of attribution, which could result in a number of personal, operational, and information security issues. An example where this type of collection has been used was during the 2021 Capitol Hill insurrection. During this incident, participants utilised a number of resources online for both planning and mobilisation, and as a result were later identified by public citizens and investigators through the media they posted online (Adams, 2022, np). This is one example of how passive / Tier 1 collection had been used in order to collect information on insurrectionists. Further to this, industry leaders such as Flashpoint recognise that the social web can be leveraged in a number of ways for both domestic and foreign extremist movements and terrorist organizations, especially in regards to recruiting, spreading propaganda, and communicating (Flashpoint, 2022, np). Table 2.1 provides some brief examples of the types of information OSINT can assist in addressing and disseminating throughout the intelligence community (IC).

More recently, much of Ukraine's success in its fight against the Russian invasion has had a direct link to the use of OSINT during the course of the war, which has enabled Ukraine to withstand multiple Russian assaults. In regards to military actions, OSINT has enabled the Ukrainian military to track and exploit Russian military plans, movements, and operations (Smith-Boyle, 2022). For example, open source satellite imagery has provided information on areas that the Russian military has attacked and is also being used to investigate war crimes and show what the war looks like on the

ground in real time (Amos, 2022). In addition to this, Ukrainians have been able to eavesdrop on Russian communications via unencrypted radio waves and cell phones[5] which has enabled Ukraine to strongly challenge Russia's military (Myre, 2022).

Type of Emergency	OSINT Can Assist in Addressing:
Counter-Terrorism and Extremism	- Financing and propaganda distribution via social networks and dark web marketplaces - Recruiting, planning and mobilisation on covert networks - Network topology
Addressing Mis- and Disinformation	- Impersonation of public/private sector or personal social media accounts - The spread of false or misleading information about a government, individual, or event - Deep fake or misused images or videos that do not represent reality - Reposting false content on legitimate sources - Misleading phrases or hashtags that quickly gain popularity
Cybersecurity	- Breaches and cyber espionage: Data related to public sector technologies; ransomware attacks on critical infrastructure; data of government entities or individuals - Network attacks and take-downs: Distributed denial of service (DDoS) attacks on critical government systems - Botnets: Infect computer networks with malware or impact democratic processes by distributing disinformation
Transportation Security	- Alerts for location-based threats (i.e. major transportation hubs) - Inform security teams about tactics used to bypass security systems or commit attacks - Monitor for threats directly targeted at the security/public sector organizations themselves (i.e. horizon scanning) - Stay alert to vulnerable data that could compromise a transportation network's digital or physical security
National and Global Crises	- Where is the crisis occurring, what does the damage/impact look like, and what is likely to happen next? Where are resources required? - How are other countries responding to the crisis, and how can that inform our strategy? - What information needs to be disseminated? - Activities happening at the border or other pertinent geographical areas

Table 2.1: Types of Information OSINT can Assist in Addressing and Disseminating Throughout the Intelligence Community
Source: Flashpoint, 2022, np

[5] The literature notes that a number of Russians brought their personal cell phones into Ukraine. Once Ukrainians learned of this, all Russian phone numbers were cut-off from the Ukrainian network, so their phones stopped working. This prompted Russian soldiers to seize cell phones from Ukrainian civilians. When the Ukrainian government learned of this, they encouraged civilians to inform the Ukrainian State Service of Special Communications and Information Protection. Ukrainian civilians largely complied, which has led the Ukrainian government to know which phones had been seized by the Russians, and effectively used them as listening devices (Myre, 2022).

Nevertheless, as technology continues to evolve the World Wide Web will adapt alongside it, and therefore so too will the tactics, techniques, and procedures (TTPs) of adversarial groups that operate in online spaces. Whether the threats appear from state groups (e.g. Russia, China, Iran) or non-state groups (e.g. DAESH's Cyber Caliphate, white supremacy groups, insurrectionists), there is no doubt that online recruitment, financing, planning, and execution will continue to escalate as the social aspect of the web continues to facilitate its ease and obscure its attribution. As this happens, the intelligence community will not only have to adjust their collection strategies, but they will also require a more in-depth knowledge of how the Internet and the World Wide Web work.

While often used interchangeably, the Internet and the World Wide Web are two vastly different entities. According to Leiner, et al, (1997), the initial concept of the Internet emerged during the 1950s in the USA out of a recognized requirement for a communication system that emphasised robustness and the ability to withstand the loss of a significant portion of underlying networks (Leiner, et al, 1997, 3). Today, we experience a world that now extends into such an online information-rich environment that can be accessed through an array of devices.

Since its initial conceptualisation, the Internet has evolved into the network of computers that applications such as the World Wide Web (the Web) operate on. Sameena Misbahudin (2019) - a Technology Reporter for the BBC - describes the Internet as the miles of interconnecting roads that link cities together, while the Web is made up of what one would see on the roads - such as a variety of shops and houses. Furthermore, the vehicles on the road would represent the data that is constantly travelling - separately from the web - transferring files and emails all across the Internet (Misbahuddin, 2019). Of course, as society grows and the demand for the Internet and the Web increases, more "roads" will need to be built to facilitate the increasing amount of "shops" that are built in addition to the "vehicles" that traverse those "roads." However, it is not only the *amount* of such "roads, shops, and vehicles" that will evolve, but also the *way* in which they evolve will be vital as well.

The Web was introduced as a concept by Sir Tim Berners-Lee in 1989 (Getting, 2007; Kamel Boulos & Wheeler, 2007). According to Nupur Choudhury (2014) and Kamel Boulos & Wheeler

(2007), Berners-Lee's view of the capabilities of the Web was conveyed through three phases which correlate with three innovations: Web 1.0 (the web of documents; the read-only web), Web 2.0 (the web of people; the read-write web), and Web 3.0 (the semantic web; the read-write-execute web) (Choudhury 2014; Kamel-Boulos & Wheeler, 2007). In brief, the main differences between the three is that Web 1.0 targets the content creativity of the producer, Web 2.0 targets the content creativity of both its users *and* producers, and Web 3.0 targets linked data sets (Choudhury 2014, 8099). Table 2.2 details further differences between the three evolutions of the Web.

Web 1.0	Web 2.0	Web 3.0
1996 - 2004	2004 - 2016	2016 +
The Hypertext Web	The Social Web	The Semantic Web
Tim Berners Lee	Tim O'Reilly, Dale Dougherty	Tim Berners Lee
Read Only	Read and Write Web	Executable Web
Millions of Users	Billions of Users	Trillions + users
Echo System	Participation and Interaction	Understanding Self
One Directional	Bi-Directional	Multi-User Virtual Environment
Companies Publish Content	People Publish Content	People build applications through which people interact and publish content
Static Content	Dynamic Content	Curiously Undefined. Artificial Intelligence and 3D, The Web of Learning
Personal Websites	Blog and Social Profile	SemiBlog, Haystack
Message Board	Community Portals	Semantic Forums
Buddy list, Address Book	Online Networks	Semantic Social Information

Table 2.2: Comparison of Web 1.0, Web 2.0, and Web 3.0
Source: Choudhury 2014, 8099; Khanzode & Sarode, 2016, 7-8

The advantages and conveniences brought about through the Internet and the World Wide Web are numerous. However, it must also be acknowledged that the same technology that enables social networking, content distribution, and various other forms of interaction, is also capable of facilitating radicalisation, recruitment, incitement, and glorification of violent acts that threaten public safety (United Nations Office on Drugs and Crime, 2012). In other words, both the Internet and the Web can and have been used as both a *tool* and as a *target* for violent acts such as terrorism, hate crimes, and mass dis/misinformation. So, obtaining a deeper understanding of the functionality of both the Internet and the Web can furthermore assist analysts within the intelligence community to

ask deeper questions of the data and information they collect. Utilising open source intelligence techniques will moreover enable analysts to identify and answer intelligence gaps that may have previously been out of reach, thereby identifying its increased value as an intelligence discipline.

2.3. Current Value

According to Evangelista, et al (2020), the process of analysing open sources was historically laborious and time consuming. Today, the amount of publicly available information and data is growing so rapidly that it is oftentimes overwhelming for analysts. This is because the rate at which data and information are available has exponentially grown as a result of both the propagation of material, and the extensive sharing of that material by virtue of the availability and massive reach of the Internet and the Web (Evangelista, et al, 2020). As explained by Ali (2021), every two years, the volume of information and data available online effectively doubles. Perhaps one of the most useful images that helps put this into perspective is the Internet Minute wheel, also known as "Data Never Sleeps," created annually by Domo. Depicted in Figure 2.2 below, is a breakdown of how much data was globally generated *every minute* in 2021. In proportion to this, Domo (2022) has indicated that 63% of the world's population (over 5 billion people) have access to the Internet, which represents a 10% increase from January 2021. Additionally, projections from Statista estimate the global volume of data and information created, captured, copied and consumed to increase from 97 zettabytes in 2022 to 181 zettabytes[6] in 2025 (Taylor, 2022).

[6] A zettabyte is 1,024 exabytes, or 10^{21} bytes. For further information on data sizes, see Appendix 1.

Figure 2.2: Data Never Sleeps 10.0 (2022)
Source: Domo, 2022

When one considers the vastness of the Internet alongside all of the activity that concurrently occurs in multiple applications, it truly is an extraordinary thing to behold. As depicted in the image above, Domo (2022) compares the amount of data being generated by users of multiple applications and platforms every minute. Among all of this, there are online services and applications that have become somewhat of a staple in today's world, which moreover produce incredible quantities of both user activity and associated data (Ali, 2021). For example, some notable *per-minute* figures to consider alongside their 2013 numbers include the following:

- Google user queries: 2M in 2013 vs 5.9M in 2022
- YouTube hours uploaded: 48 in 2013 vs 500 in 2022
- Instagram photos shared: 3.6K in 2013 vs 66K in 2022

- Twitter tweets shared: 100K in 2013 vs 347K in 2022
- Facebook content shared: 684K in 2013 vs 1.7M in 2022
- Emails: 204M in 2013 vs 231M in 2022
- Global Internet Population Growth in billions: 3.4 (2016), 4.3 (2018), 4.5 (2020), 5.0 (2022) (Domo, 2022).

What Domo outlines above is a clear indicator of how online engagement via social media platforms, streaming services, peer-to-peer communications, and other activities have "increased hundreds and even thousands of percentage points" (Domo, 2022, np). Undoubtedly, the volume of data created every minute will continue to accelerate as modern life demands a digital presence. As a result of this, every like, tag, share, swipe, stream, email sent, or link clicked, tells a story of the user, and to a degree the community that surrounds them. But what does this mean in relation to public safety?

In the article titled "Terrorism and the Use of Social Media in Mumbai Attacks," Justin Nachsin (2010) discusses how social media had made a notable appearance over the course of the 2008 Mumbai attacks.[7] Nachsin (2010) makes particular note of how terrorists began to post and share images on social media detailing their attacks as they occurred (Nachsin, 2010, 15). Further to this, Nachsin illustrates that in the years preceding the Mumbai attack, terrorists had a limited means of disseminating their message - usually through underground media and newspapers or blogs (Nachsin, 2010, 16). In agreement with Nachsin (2010), Joel Vargas (2014) argues that the impact of social media from the 2008 Mumbai attacks shed light on its significance as a valuable tool to intelligence communities around the world (Vargas, 2014, 34).

Not long after the 2008 Mumbai attacks, an additional violent incident in April 2013 would occur which would shed further light on the value of publicly available information and open source intelligence. In their article, "Digilantism: An Analysis of Crowdsourcing and the Boston Marathon Bombings," J. Nhan, L. Huey, & R. Broll (2015) explore the outcome of the bombings.[8] In particular, they draw attention as to how the general public utilised online communities such as Reddit in order to assist law enforcement with their investigation. According to Nhan, et al (2015), although members

[7] On 26 Nov 2008, ten members of Lashkar-e-Taiba (LeT), (a vast jihadist group based in Pakistan) simultaneously attacked five locations in Mumbai with AK-47-type weapons, killing 140 Indians and 25 foreign tourists (Mahadevan, 2019).

[8] The Boston Marathon bombings occurred on 15 April 2013. Two bombs had exploded near the finish line of the Marathon killing three people and injuring over 170 others. The investigation that followed was one of the largest manhunts in US history lasting four days and consisting of the convergence of police, citizens and technology having a significant role in the real-time pursuit of the perpetrators (Montgomery, et al, 2013; Nhan, Huey, & Broll, 2015, 341).

of the Reddit community did not entirely resolve the investigation on their own, "their actions suggest the potential role the public could play within security networks" (Nhan, Huey, & Broll, 2015, 341). Similarly, Gary Marx (2013) suggests that online communities such as Reddit extend the 'eyes and ears' of police, especially when the requirement for efficient assistance is met with increasingly scarce resources that have been exhausted by newer communication and analysis technologies (Marx, 2013, 57; Nhan et al, 2015, 342).

As the literature thus far suggests, OSINT is increasingly valuable for its real-time investigative and analytical capabilities. In reference to counter-terrorism efforts and the global war on terror (GWOT), Akhgar, Bayerl, & Sampson (2016) agree that OSINT provides significant increase in situational awareness. This has been identified specifically when, in May 2013, one of the terrorists planning to bomb the Embassy of Myanmar in Indonesia was detected and foiled after the group's plans to execute the attack was revealed through a Facebook status update (Younas, 2014). However, as conveyed by Kleckova (2021), an OSINT analyst's work is often more elusive and abstruse, not limited to social networks, and is rarely as simple as finding a perpetrator's plans through a status update. Most often an analyst will require input from other analytical techniques and intelligence disciplines in order to corroborate intelligence.

Perhaps one of the most significant ways in which OSINT has been used in a counter-terrorism effort is in the conflict against DAESH.[9] According to Carter, Maher, & Neuman (2014), a significant number of DAESH fighters are active on social media spreading propaganda and battlefield information in order to provide details on what is happening on the ground to their supporters. Some examples of DAESH's use of social media is the group's posting of images, discussions of battles, the reporting of member's current locations and activities associated with battles, in addition to testimonials relating to events taking place within and outside of Iraq and Syria (Kleckova, 2021). Pursuant to this, Klausen (2014) identifies that in a reviewed selection of 563 tweets, approximately 40.32% of them often included content depicting the location of fighters,

[9] Also known as the Islamic State (IS), the Islamic State of Iraq and the Levant (ISIL), the Islamic State of Iraq and Syria (ISIS), and by its Arabic name al-Dawla al-Islamiya fil Iraq wa al-Sham (Daesh - also spelled Da'ish). The Arabic acronym will be used throughout this document.

attacks and battles, alongside the discussions related to them. Thus, as Tow and Yeo (2005) point out, Jihadists have made use of the Web as a means to communicate, spread propaganda, and interact with the world. They further argue that the value of OSINT is to be found in identifying and understanding the ideological concerns and thinking of key personnel in order to provide a type of 'strategic barometer' of the perpetrator's intentions - each of which can be found through analysing digital publications such as Al Qaeda's "Voice of Jihad" and DAESH's "Dabiq" (Tow and Yeo, 2005).

As mentioned in the previous section, the value of OSINT has been underscored through its operational level impact in the Russian invasion of Ukraine. According to Richard Baffa (2022), OSINT has a wide range of applicability during international conflict, providing information and data relevant to developing Indications and Warnings (I&W) for a number of incidents affecting the safety and security of civilians and military personnel. Specifically, in Ukraine OSINT has allowed both the Ukrainian military and regular civilians to track Russian military movements and operations (Smith-Boyle, 2022). As specified by Pranshu Verma (2022), Russia's invasion of Ukraine has been unravelling at a rapid rate over social media and has also greatly expanded the ranks of "hobbyist spies" and numerous others who keep track of such movements, ultimately having an impact on tactical planning. Further to this, Ukraine's minister of digital transformation - Mykhailo Fedorov - has detailed the significance of the community's work that resulted in the Ukrainian government creating an app called Diia, which allows people to upload geotagged photos and videos depicting Russian troop movement (Verma, 2022). While some use the Diia app regularly, many others have taken to Twitter to chronicle footage from the war. Simply conducting the search "#Ukraine" on Twitter will call up a number of accounts that track weapons and troop movement throughout Ukraine.

In a number of ways, the Russian invasion of Ukraine has been perceived as the first war that is quite similar to a live broadcast event with the OSINT community assisting through the dissemination of information (Puiu, 2022). It is significant to note here that OSINT is continuously undergoing metamorphosis and evolution to reflect the ongoing technological capabilities of the modern digital environment. As such, indicating its value today may undermine its utility in the near future. However, there is no doubt that OSINT has and will continue to show intelligence value both

for emergency management and public safety, especially as intelligent electronic devices and online applications continue to have such a prominent role in our lives. The following section will bring forward the literature on the current capabilities within OSINT, which will provide a more thorough understanding on the technologies and techniques most effective for open source intelligence.

2.4. Capabilities

The essence of any discipline of intelligence is that it is timely and accurate. Both of these characteristics are driving forces to the success of any intelligence-lead operation, crisis, or task. As such, the capabilities that manifest success within the open source environment are a combination of processes, frameworks, skill sets, behaviours, and tools. Although at this time there is not a significant presence of *peer-reviewed* literature speaking specifically to the capabilities of open source intelligence, this section will address that gap by identifying the aforementioned processes, frameworks, skill sets, behaviours, and tools which combined, form the basis of open source intelligence capabilities that are used by practitioners in the field.

2.4.1. Social Networks

Historically, a social network was a term that referred to one's network of associated relations with other people. As technology has become more central to our lives, social networks and social networking has taken on a meaning that is more often associated with online based social media websites where one connects with and develops a number of relationships with others, be they familial, cordial, or professional. Today, social networking via social media is one of the more popular activities one conducts online. According to Statista, as of 2021, 4.26 billion people utilise social media to connect with others online – a number projected to increase to approximately six billion in 2027 as internet infrastructure and the availability of affordable mobile devices reach under-developed regions (Dixon, 2022b).

Social networking sites quite possibly host the largest database of information that can be utilised by private investigators and government agencies. As of January 2022, the top five social

networks ranked by monthly active users were Facebook, YouTube, WhatsApp, Instagram, and Weixin/WeChat (Dixon, 2022a). Each of these platforms contain photos, videos, geospatial

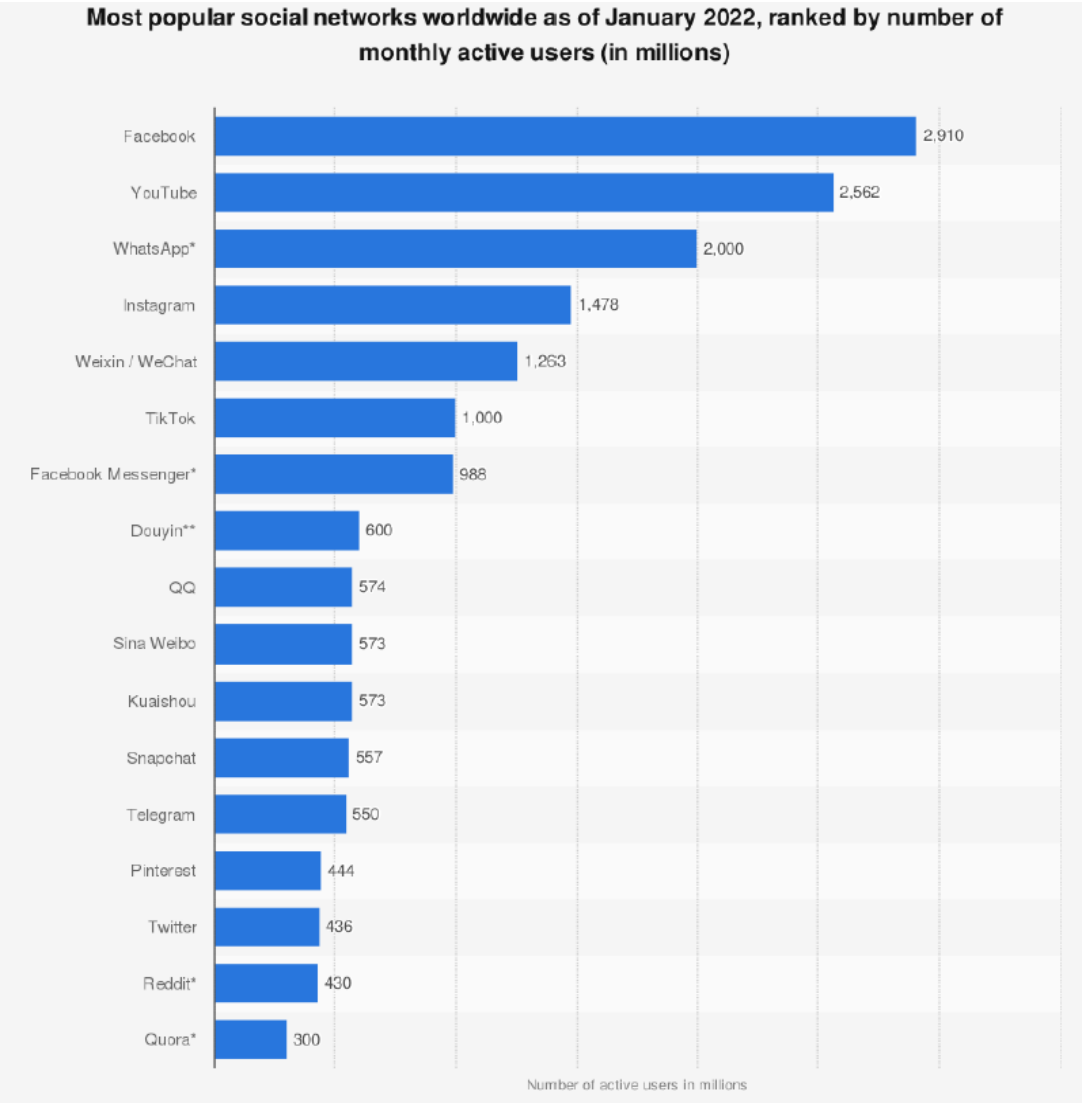

Figure 2.3: Most Popular Social Networks Worldwide (Jan 2022)
Source: Dixon, 2022a
*Platforms have not published updated user figures in the past 12 months, figures may be out of date and less reliable
**Figure uses daily active users, so monthly active user number is likely higher

information, comments, personal information, and personal associations. Considering this, it is understandable that social networking sites contain troves of information that may be able to assist in identifying threats. For example, in the LexisNexis article and infographic, "Social Media Use in Law Enforcement," survey respondents had provided scenarios where social networking sites were used to prevent or stop crimes such as thwarting an active shooter, actively tracking gang behaviours, and executing a number of outstanding arrest warrants (LexisNexis, 2014; LaSorsa, 2022; Byrne, 2022).

Furthermore, LaSorsa states, “For the private investigator seeking information on the behavioural circumstances of a subject, something as quick and easy as analysing status updates, check-ins and posted photos by the subject and their friends may provide the information necessary to conclude if a legitimate threat exists” (LaSorsa, 2022, 1). Pursuing this type of information via social networking sites therefore prove to be beneficial to investigations supporting public safety.

In consideration to the billions of registered social network users, one can imagine the vast amounts of data available and accessible to anyone who knows where to look. At the most basic level, anyone who conducts a search for a moniker, phone number, or email address for example, has conducted a basic open source investigation themselves (LaSorsa, 2022). These searches may represent a simple query in an attempt to find and connect with an acquaintance, but they can also be used in an investigative or influential capacity. Gritzalis and Stavrou (2016) identify eight capabilities by which social networking sites may be used. They are: (1) the identification and prediction of insider threats; (2) detecting the sentiment of a population; (3) the identification and prediction of public opinion (4) the identification and prediction of fluctuations in public opinion; (5) the detection of influential user(s) (a central node/hub) of a cause; (6) detecting the appropriate content and means of communication; (7) achieving public influence; and (8) optimising the efficiency of communication and influence tactics (Gritzalis & Stavrou, 2016, 8). It is important to note here that these eight capabilities can be used both offensively or defensively, and by anyone. Further to this, the use of social networking sites may be compounded when utilised in tandem with a number of other capabilities such as geolocation tools, Internet of Things devices, the Web (surface, deep, and dark), and data analytics tools.

2.4.2. Online Maps and Geolocation Tools

Satellite imagery has become an increasingly popular open source tool, especially given the level of detail and historical searches available at little to no cost to users. According to Michael Bazzell (2022), online map data has been used by a number of agencies before executing a search warrant for a residence because of the detailed views provided that allow for the examination of structures, doors, windows, obstructions, ingress/egress routes, alleys, and driveways. This level of

detail is compounded when you consider the various views, layers, and tools that the user is able to utilise. For example, street view, satellite view, terrain layer, and distance measurement are examples of just some of the many features that can be employed in an investigation and can be used freely on a number of sources such as Google Maps, Bing Maps, Yandex Maps, and Wikimapia.

In addition to online maps, there are also a number of tools that can be used for geolocation. According to Steven Harris (2020), one of the fastest and simplest means to geolocate an image is through utilising EXIF data,[10] which is the file that stores the metadata that is captured by most phones and digital cameras. In particular, the most important metadata here would be the GPS coordinates which record the location of where the image was captured. However, it is important to note that although this is a relatively simple method that reveals accurate information, one of its biggest issues is that most social networking sites, messaging apps, and web services remove GPS information from the images they process (Harris, 2020). As a means to overcome this, Harris indicates that analysts can use the extract,[11] research,[12] and verify[13] methodology as a technique to collect, examine, and verify a location of interest.

In addition to Harris' extract, research, and verify methodology, an additional technique sometimes referred to as chronolocation, which utilises light and shadows to determine a possible time frame, can be beneficial to geolocation analysis. According to Youri van der Weide (2020), one of the most popular tools to assist with this type of analysis is known as SunCalc, which allows researchers to "analyse the position of shadows and the sun at any given time and date, and at any given location" (np). Essentially, the user simply selects a date and location, and is then able to drag the sun icon in order to observe the position of the sun and any shadows at a specific time in location. Van der Weide (2020) indicates this technique has been successfully used in a number of investigations such as the Hadrut executions[14] that took place in Azerbaijan in October 2020.

---

[10] EXIF data may be collected through a variety of tools, some of which are browser based such as Firefox's Exif Viewer, Forensically, and Google Chrome's Exif Viewer Pro (Harris, 2020).

[11] Extract: Collect as many data and information points as possible from the images you are analysing (Harris, 2020).

[12] Research: Examine and research those data and information points in order to obtain additional information about the location (i.e. vehicle registration, architectural features) (Harris, 2020).

[13] Verify: Confirm your research findings through a verification process. Once you have a general location of where the image was taken, utilise existing images and/or mapping tools for comparison corroborate your findings (Harris, 2020).

[14] Additional information on how chronolocation had been used during this investigation can be found here: https://www.bellingcat.com/news/rest-of-world/2020/10/15/an-execution-in-hadrut-karabakh/

In summary, online maps, geolocation tools, and chronolocation assist OSINT investigations by providing the analyst with a methodology to determine the location of images and their approximate time of capture. This may be significant to OSINT investigations as it provides the analyst with the tools necessary to determine the authenticity of events in addition to a location of interest that can provide additional information and pivot points to the analyst(s) developing the intelligence picture of a particular event. From here, the investigation may make use of data and information from the surface, deep, or dark web for additional insight into the location and person(s) of interest.

2.4.3. The Surface, Deep, and Dark Web

The World Wide Web exists as a number of layers far beyond the search engines that index its content. Weimann (2015) explains that as the web has expanded in use, search engines such as Google and Bing were capable of retrieving static web pages, but incapable of retrieving dynamic[15] webpages. So, what was indexed and retrievable became known as the clear, or 'Surface Web,' and the pages that were unable to be indexed were coined as the 'Invisible Web' by Jill Ellsworth in 1994 (Burgees, 2007). In 2001, the term 'Deep Web' was used by Michael Bergman to better describe 'invisible' web pages that contained content that was (a) beyond the reach of traditional search engines; (b) retrieved only via keywords or a targeted query; (c) unable to be, or not indexed by traditional search engines; and (d) secured via username/password, registration or other types of codes (Weimann, 2015; Burgees, 2007).

Further to the above, Weimann (2015) describes the deepest layer of the web as the 'Dark Web', which consists of webpages that have been intentionally concealed, and generally contains illegal content that can only be viewed through the use of specialised browsers such as Tor (The Onion Router), I2P (Invisible Internet Project), and Tails (The Amnesic Incognito Live System) (Weimann, 2015; Lee, 2022). Figure 2.4 provides additional information for understanding the differences between the Surface, Deep, and Dark Web.

[15] A static webpage is on that is linked to other pages on the web, whereas a dynamic webpage is linked to a specific webpage that can only be retrieved through a targeted query (Weimann, 2015, 195).

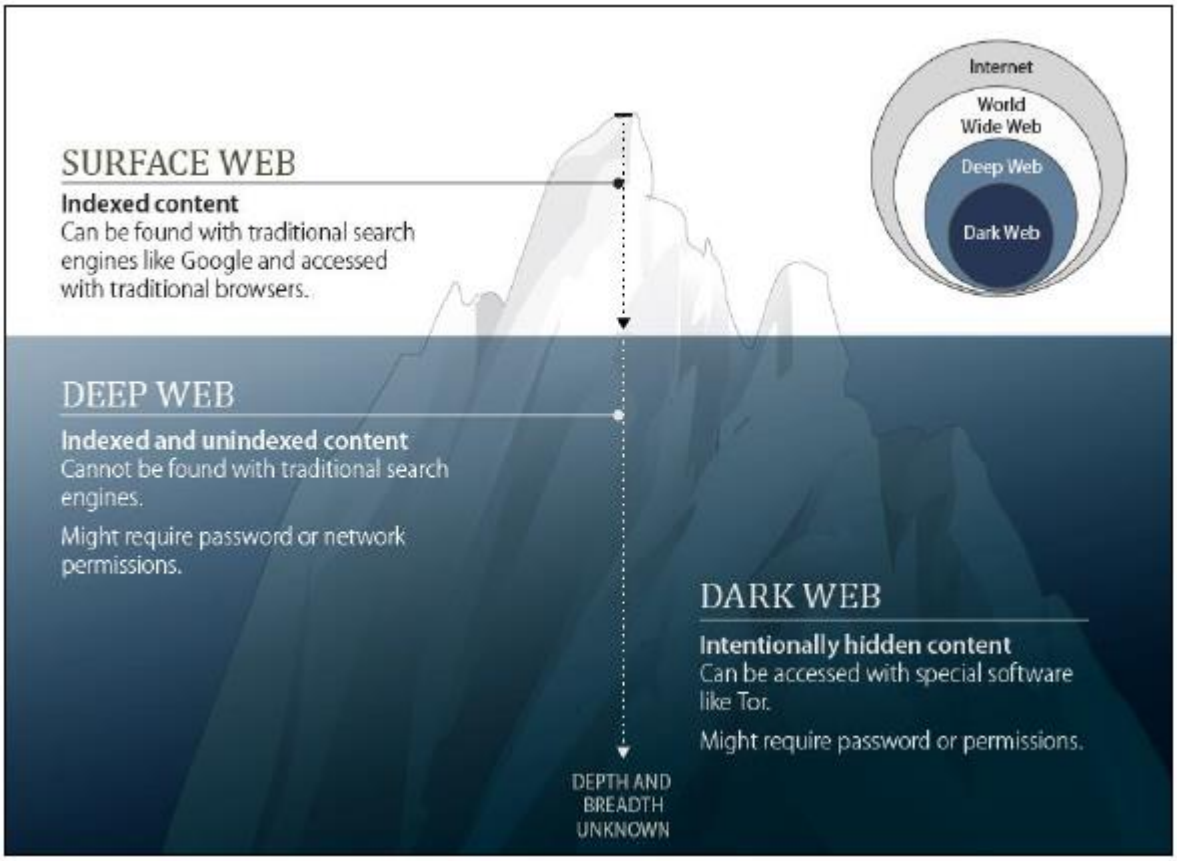

Figure 2.4: Layers of the Internet
Source: Finklea, 2017, 4

Similar to how malicious activity may occur on the Surface Web, so too can it occur on the Deep Web and the Dark Web. In her Congressional report, Kristin Finklea (2017) describes how malicious actors – from criminals to terrorists to state-sponsored entities - may leverage the Web as a forum for "conversation, coordination, and action" (Finklea, 2017, 9). While this is largely due to a number of factors including risk of detection, association, and attribution, it is also a result of malicious actors relying on technology to carry out their operations. Dilipraj (2014) explains the Dark Web as a type of "treasure trove" to malicious actors because it enables the distribution of information, training material, recruitment, planning, and the coordination of attacks. This is a result of the increased difficulty of policing the Dark Web.

In July 2015, FBI Director James Comey announced that DAESH was using the Dark Web as a means to inspire attacks throughout the world. Comey's statement also warned that DAESH's use of encryption and new forms of communication surpassed the level of the US government's current legal and technological capability to adequately keep up with the group (Weimann, 2015). While DAESH

activities had been closely monitored on the Surface Web, its activity on the Dark Web effectively hid the identity and location of its members and supporters. Following the Operation Paris campaign set in motion by the hacker collective, Anonymous, DAESH's media outlet – the Al-Hayat Media Center – broadcast explanations and links on Telegram (an encrypted communication platform) wherein supporters could access their forum on the Dark Web through the Tor browser (Weimann, 2015). The messages and links (hidden by the author) of the message was as follows:

> "The name of Allah the Merciful given the very narrow on the site # Asaddarat_klavh so that it is deleting any new domain after its publication announce the launch of the site on the "Dark web" * will work for users of the Tor and users of ordinary users of the Tor link: XXXXXXXXXX link ordinary users: XXXXXXX and we promise you that we are continuing to try to get a new range of normal and we will publish, God willing, when you get it next to the scope of the Tor. Glory be to God, His Prophet and the believers." (Weimann, 2015, 198).

According to Weimann (2015), this Dark Web site held links to DAESH's private messaging group on Telegram, as well as an archive of various types of propaganda type material such as their documentary-style film, *The Flames of War* (Weimann, 2015). Thus, as a result of the anonymity provided by Dark Web browsers such as Tor, this area of the Web has become a venue for malicious actors to distribute information. However, there are circumstances in which the Dark Web may in fact bestow benefits not only for malicious actors or those seeking privacy online, but also to those regularly working toward public safety: law enforcement, the military, and other government agencies (simplified from here forward as the intelligence community.

Anonymity on the Dark Web can be leveraged by the intelligence community as a whole. Kevin Poulsen (2013) explains how the FBI has been using a "computer and internet protocol address verifier" (CIPAV) since 2002 as a means of identifying potential suspects who conceal their location using anonymising services (e.g. Tor) or proxy servers in order to target malicious actors. Additionally, Donna Leinwand Leger (2014) explains that the intelligence community can also utilise common mistakes made by malicious actors, or flaws in technology in order to support their operations. This is precisely what occurred in 2013 when the FBI successfully disbanded the then

world's largest underground black market, the Silk Road[16] and shortly after, Silk Road 2.0[17] (Krebs, 2014).

Of course, defeating malicious actors on the Dark Web is not limited to the application of technology. For example, Andy Greenberg (2015) explains that some of the most reliable tools that the intelligence community may employ come from traditional crime-fighting means. For example, in March 2015, US federal investigators had "sent a subpoena to Reddit demanding that the site turn over a collection of personal data about five users of the r/darknetmarkets forum [a subreddit where users discussed anonymous online sales of drugs, weapons, stolen financial data, and other contraband]" (Greenberg, 2015). Consequently, as Greenberg (2015) explains, utilising subpoenas in this fashion may drive such discussions and activity to anonymous forums of the Dark Web, where once again the intelligence community will require technological solutions to identify users and locations.

In regards to more offensive operations, Finklea (2017) explains that the Tor Dark Web browser can be used as a means to conduct computer network operations that could involve initiatives not limited to intercepting or inhibiting malicious communications, psychological operations,[18] taking down a website, or conducting a denial of service attack. Although the majority of the intelligence community's detailed use of the Dark Web against malicious actors is classified, understanding how data analytics is used in open source intelligence can provide a significant grasp on some of the methodologies that may be employed when conducting investigations on the Surface, Deep, or Dark Web.

2.4.4. [Big] Data Analytics

The Internet is the largest database in the world. As previously discussed, the volume of data and information created online is expected to increase from 97 zettabytes in 2022 to 181 zettabytes in

[16] It was reported that FBI agents had exploited weaknesses in the code used to operate the Silk Road website. Those weaknesses were thereafter used to hack the website's servers which revealed the unique identifying addresses that the FBI's investigators could then use to locate and seize the servers (Leinwand Leger, 2014).

[17] Silk Road 2.0 was disbanded after the FBI discovered critical errors made by the website's proprietor who used his personal email address to register the servers. Agents were able to seize the servers which resulted in the website being shut down (Leinwand Leger, 2014).

[18] For example, disinformation or misinformation campaigns (Finklea, 2017).

2025 (Domo, 2022; Taylor, 2022). According to Puleri (2021), the research community has recently come to understand the progresses made in big data analytics[19] – especially as it relates to open source information – as a means to yield a proficiency to study social and behavioural dimensions of online activity.

According to Fisher, et al (2012), posting one's thoughts, opinions, social relationships, photos, videos, purchasing habits, et cetera., online leaves behind a detailed digital record of one's life that can be used by companies, government agencies, researchers, and scientists. As a response to this, big data analytics has become a major discipline that aims to visualise "the big picture from the minutia of our daily lives" (Fisher et al, 2012, 50). On its own, data analytics allows an analyst to collect, transform, and organize data with a view to inform decision making, improve analysis, draw conclusions, and amplify assessments (Google, 2023). Introducing big data to this acknowledges the multiple types of data,[20] in addition to the complexity in managing such voluminous information using traditional data processing tools that needed to account for the characteristics of big data.[21] (BasuMallick, 2022; Panimalar, Shree, & Kathrine, 2017, 330-332).

In regards to utilising open source intelligence and big data analytics, Pastor-Galindo, et al., (2020) discuss how such processes continuously expand upon the knowledge base of the subject of interest. Specifically, three principal use cases are identified including (but not limited to), social opinion and sentiment analysis, cybercrime and organized crime, and cybersecurity and cyber defence (Pastor-Galindo, et al., 2020, 10282-10283). On the matter of social opinion and sentiment analysis, Pastor-Galindo, et al., (2020) explain that the collection of user interests, preferences, interactions, and messages may be beneficial when applied to disaster management, which is a component of public safety (Pastor-Galindo, et al., 2020, 10282; Bello-Orgaz, et al., 2016). Regarding cybercrime and organized crime, the authors discuss how analytics and OSINT processes can be applied to

[19] "The size that constitutes 'big' data has grown according to Moore's Law. In 1975, attendees of the first VLDB (Very Large Databases) conference worried about handling millions of data points found in U.S. census information" (Simonson and Alsbrooks, 1975, 496-498; Fisher, et al, 2012, 52). In the context of information visualization, Shneiderman describes a dataset as big when it's too big to fit on a screen – at one item per pixel, most desktops would stop at a few million data points today" (Shneiderman, 2008, 3-12; Fisher, et al., 2012, 52).

[20] Multiple types of data may include: structured data, unstructured data, semi-structured data, geospatial data, machine or operational logging data, and open-source data (BasuMallick, 2022).

[21] Volume, Velocity, Value, Variety, Veracity, Validity, Volatility, Visualization, Virality, Viscosity, Variability, Venue, Vocabulary, Vagueness, Complexity, Verbosity, Voluntariness, Versatility (Panimalar, Shree, & Kathrine, 2017, pp. 330-332). See Appendix 2 for additional explanations.

identifying patterns and relationships between criminal acts (Pastor-Galindo, et al., 2020, 10282-10283). This in turn provides the intelligence community with the opportunity to exploit open source data, detect nefarious activity, track the activity of criminal and terrorist organizations that have been expanding their online presence, and possibly prevent or appropriately react to situations that threaten public safety (Larsen, et al., 2017; Dawson, et al., 2018; Ali, et al., 2019; Pastor-Galindo, et al., 2020, 10282-10283).

The final analytic use case that Pastor-Galindo, et al. (2020) discuss is that of cybersecurity and cyber defence, which largely focuses on information and communication technologies, sometimes abbreviated as ICTs (Pastor-Galindo, et al. 2020). Jang-Jaccard & Nepal (2014), assert that as such systems undergo continuous attacks from criminals and state actors, the availability, integrity, and confidentiality of those systems and services are at risk of being exploited. In relation to this, Nespoli, et al., (2018) note that it is crucial to defend ICT system infrastructure through the analysis and correlation of attacks, which can moreover assist with effective response. Additionally, in a similar manner, Quick & Choo (2018) contend that open source intelligence and data analytics can support forensic digital analysis as a means to complement the digital evidence that may be left by cyber incidents.

In addition to the use cases discussed above, Pastor-Galindo, et al. (2020), also indicate that big data analytics alongside open source intelligence may also be applied to scenarios involving social engineering, identifying deepfakes, disclosing fake news, natural language processing, social network analysis, et cetera (Pastor-Galindo, et al., 2020). With regards to these and the aforementioned use cases, making use of data mining and artificial intelligence techniques in addition to big data analytics and open source intelligence techniques will assist in developing an intelligence picture that benefits public safety through the following techniques and processes:

> **Correlation**: Kim, et al. (2018) describes this as the detection of the relation and connection between variables including people, data, and events, where relations are seen as valuable in revealing non-explicit associations that may be present in the dataset(s).
> **Classification**: This includes data that can be grouped by predefined categories in order to permit knowledge extraction through effective organization of information (Pournouri, et al., 2019; Deliu, et al., 2017).
> **Outlier detection**: This process applies analytic techniques to the datasets with a view to

detect anomalies that may identify hostile actors whose actions and behaviours diverge from the general population (de la Torre-Abaitua, et al., 2019).

Clustering: Considering a number of conditions, data is assigned into clusters (Azevedo, et al., 2019). Clustering may reveal various online profiles, network behaviour, or forms of attacks on individuals, organizations, or infrastructures without prior knowledge of the existence of such variance or diversity (also known as unsupervised learning - discovering patterns without human intervention) (Wang, et al., 2018a).

Regression: This technique is used to predict numeric facts or values as a means of predicting continuous outcomes, and whether or not changes observed in one (dependent) variable is associated with another (explanatory) variable (Pellet, et al., 2019; Castillo, 2021).

Tracking Patterns: Not to be confused with outlier detection, pattern recognition is used to detect the regularities in data (Wang, et al., 2018b).

Together, these techniques and processes can deduce complex and abstract associations within the dataset(s) that are otherwise not explicitly drawn out. However, while the application of such techniques to public safety shows considerable value, it is important to address a number of challenges.

In their article, Pastor-Galindo, et al., (2020), identify a number of challenges in using these techniques, especially in regards to researching and developing the knowledge extraction process[22] and privacy considerations. The authors contend that, "the extracted knowledge about a person, company or organizations may be specially sensible and its manipulation indirectly leads to ethical and legal problems"[23] and may potentially be misused to harm people or groups (Pastor-Galindo, et al., 2020, 10288). On a similar note, Cardenas, et al., (2013) indicates that privacy is increasingly relevant as requests for sharing data between industries and agencies emerge. This relevance exists on the basis that such sharing negates the "privacy principle of avoiding data reuse - that is, using data only for the purpose that it was collected" (Cardenas, et al., 2013, 75). Up until recently, maintaining privacy largely relied on the limitations of traditional collection, extraction, analysis, and correlation, which has to a great extent been overcome by big data analytics, which in turn has inadvertently made privacy violations easier (Cardenas, et al., 2013). As a result of this, Cardenas et al. (2013) argue that it is imperative to consider privacy principles when developing big data applications. It is also evident that big data analytics - especially as it relates to open source intelligence for public safety - is

[22] The knowledge extraction process is that which identifies, profiles, or monitors malicious criminals and organizations, and/or uncovers and attributes cybernetic incidents (Pastor-Galindo, et al., 2020, 10288).

[23] These issues will be further discussed in section 2.6 Legal, Ethical, and Policy Challenges.

challenging the landscape of security technologies and privacy concerns. Therefore, it is imperative for analysts and software developers to remain familiar with and value privacy in their endeavor to utilise data and develop applications that assist with open source intelligence operations. The following two sections will discuss the literature on the evolution of risk and vulnerability in using OSINT for public safety, followed by the legal, ethical, and policy challenges that are relevant to these matters.

2.5. Deterrents to Progress: Tradecraft, Culture, Resourcing, and Partnership Barriers

The previous sections have aimed to provide context into what OSINT entails, the value it provides to public safety, and several of the capabilities it encompasses. This section will explore the literature on the deterrents to progress in the intelligence community's engagement with OSINT as a collection method and investigative tool. Additionally, it will also examine literature on overcoming both institutional and capability barriers to progress.

Tradecraft, hereto referring to the tactics, techniques, procedures, methods, and technologies one may use within a profession, is in an ever evolving state. If stagnation were to occur, it is inevitable that such tradecraft would become obsolete. As such, it is crucial for analysts within the intelligence community to not only maintain, but to be anticipatory towards developments in both intelligence and counterintelligence defensive and offensive tradecraft. According to Janjeva, et al., (2022) barriers within the OSINT tradecraft include the verification of analytical rigour, a gap in OSINT tradecraft (e.g. training), and the verification of PAI for the evidential context.

On the matter of verification and analytical rigour, Janjeva, et al. (2022) insist on having rigorous processes in place as a means to evaluate the veracity of information output, especially as the range of sources and techniques applied to collect PAI are quite vast and constantly growing. As the digital environment is rampant with misinformation, disinformation, deepfakes, and data manipulation,[24] analysts attribute considerable risk to products and processes that do not undergo verification and revision, based on standards from across the intelligence community (Janjeva, et al.,

[24] An example that the authors use is one that is commonly found online, where footage of one particular incident (i.e. a riot with a violent incident) later on appears in a forum claiming it belongs to a different incident (i.e. a war crime) (Janjeva, et al., 2022).

2022). Related to this is the second OSINT tradecraft barrier that Janjeva, et al. (2022) identify - the gap in genuine OSINT tradecraft. The authors contend that this second gap is better explained through first distinguishing between three key types of assurance activities: (1) the validation of sources, (2) the verification of individual pieces of information,[25] and (3) the establishment of robust analytic tradecraft for generating and testing analytic judgements[26] (Janjeva, et al., 2022). Together, the authors illustrate these barriers as burdens that have been placed on the analyst as a result of a lack of "tried and trusted" methods for assessing veracity that other intelligence disciplines have overcome (e.g. reliability and confidence assessments for HUMINT sources and information[27]) (Janjeva, et al., 2022). The third and final barrier within OSINT tradecraft that Janjeva, et al. (2022) specify is the verification of PAI for the evidential context. Here, Janjeva, et al. (2022) acknowledges that the criticality of verification carries additional significance in both legal and evidential proceedings. They largely argue that maintaining possession of a defensible presentation which outlines the investigative steps taken[28] is a necessity for both transparency and to complete the evidential chain (Janjeva, et al., 2022).

In addition to the above OSINT tradecraft barriers, Janjeva, et al. (2022) also identifies issues in culture, resourcing, and partnership as further barriers that act as deterrents to progress. As opposed to the latter two, cultural barriers are perhaps one of the most biased-ridden deterrents thus far as they are directly related to an anchoring bias (how things have always been done). As a result of this, new techniques - especially those that demand time and money for training - may be difficult to justify especially if an analytics team is actively engaged in numerous operations. Specific to cultural barriers, Janjeva, et al. (2022) associate biases "favouring classified information and internal datasets,

[25] For example, geolocation. Online content is created in both time and space. So, it will take time to identify and contextualize clues within the content and corroborate it with other databases as a means to identify what is factual, and moreover determine whether or not any secondary analysis is required. An example of this would be to conduct a comparison of observed railway carriage numbers with a country's railway database) (Janjeva, et al., 2022).

[26] For example, creating evidence documents, red teaming, and peer review. Every OSINT investigation should contain an evidence document that includes screenshots, annotations, and transcripts that contain footnotes for every single data point. All assumptions and assessments within such evidence documents should then be interrogated by a senior member of the team to certify that all assumptions and assessments are beyond reproach. Shortly after, analysts external to the original investigative team should "red team" (take an adversarial approach) the document as a means to find any analytical mistakes or faults within the evidence. This will be increasingly significant as OSINT investigations are used in targeting processes and legal proceedings (Janjeva, et al., 2022).

[27] See Appendix 3 for source and information credibility scales to gain a better understanding of how HUMINT is evaluated.

[28] Hunchly is a popular and effective software that is known for its screen-recording software and detailed audit trails (Janjeva, et al., 2022; Hunchly, 2023).

approaches to training and upskilling, and the fragmentation of open source activity across government" as three central domains that shed light upon cultural deterrents to progress for open source intelligence investigations (Janjeva, et al., 2022, 31).

Moving forward to resourcing and partnership barriers, Janjeva, et al. (2022) link this to investment and procurement challenges, in addition to a lack of partnerships with external stakeholders. This is perhaps not surprising as new and/or developing capabilities require considerable resources and partnerships in order to come to fruition. In agreement with Janjeva, et al. (2022), time, money, and expertise will be critical resources that are especially relevant in this circumstance. Despite large portions of PAI not having an adherent cost, it will still require technical expertise to train, collect, analyse, secure, and store OSINT data - all of which has a financial and temporal cost. A significant factor to consider here is the value and cost of committing to advancing such a capability, and the comparison to the cost of not possessing the capability at all. This is especially pertinent to consider as the intelligence community is in an era where technological competitiveness and advancement is heavily reliant upon in modern society.

2.5.1. Risk in the Utilisation of OSINT

Alongside deterrents to progress are the risks and vulnerabilities that may produce a sense of unease to the utilisation of OSINT techniques. While many definitions of risk exist, this work will make use of the cyber risk definition as it is most relevant to OSINT. As such, cyber risk is defined as "the likelihood of damage to sensitive data, critical assets, finances, or reputation" (Kost, 2023, np). In view of this, cyber risk is calculated by accounting for "the identified security threat, its degree of vulnerability, and the likelihood of exploitation," and can be expressed as follows (Kost, 2023, np):

Cyber Risk = Threat x Vulnerability x Information Value (Kost, 2023)

In consideration to the risk equation above, examples of cyber risk may include data breaches, data leaks, cyber attacks, malware, ransomware, social engineering, phishing, and insider threats (Kost, 2023). While each of these are relevant to OSINT, it can be argued that transparency[29] and the

[29] For example, Janjeva, et al. (2022) makes an interesting contrast: where government stakeholders are generally apprehensive about transparency due to the concern of revealing tradecraft capabilities, the opposite is true for civil society stakeholders who without maximum transparency, risk having their credibility undermined (Janjeva, et al., 2022, 43).

exposure of tradecraft especially with regards to techniques, tools, and capabilities are risks worth exploring as well. Alas, at this time there is limited peer-reviewed literature at the unclassified level available to provide an informed review on the risk of exposure to public agencies, likely because it would reveal inherent vulnerabilities that would increase their risk of attack.

Moving forward however, as OSINT has begun to gain popularity amongst law enforcement agencies (LEAs), government agencies, and a number of private sector companies, a number of concerns relating to discrimination, function creep, data protection, and privacy have emerged. These concerns are largely a response to the facilitation of monitoring digital spaces which increases the visibility and exploitation of social life (Troittier, 2015). According to Troitter (2015), privacy advocates have noted that monitoring technologies - especially in relation to social media - appropriate entire populations and make them suspects. Troitter (2015) clarifies that the basis to this argument is grounded in issues of data retention and analysis, and how such retained data may result in function creep.[30] A key theme in Troitter's (2015) article is that data and information - especially on social platforms - is produced in a particular context, and as such, any collection of data should only be authorised for a particular purpose and length of time in order to avoid discrimination and false criminalisation. Some examples where this has been of concern include Los Angeles, California, USA, where Operation Laser[31] and PredPol[32] were active between 2011-2020 (Bhuiyan, 2021; Puente, 2019). Although no longer in use after having been proven inconsistent and racist, the programs essentially validated current policing patterns, and reinforced the need for police officers to patrol certain neighbourhoods and people over others, resulting in the over-policing of Black and Brown communities within Los Angeles (Bhuiyan, 2021). What Troitter (2015), Bhuiyan (2021), and Puente

[30] Function creep may occur when information or data collected is used for a purpose outside of its original specified purpose. For example, "a workplace may install a security system that requires employees to sign-in or sign-out of the workplace. The purpose of the security system is to prevent unauthorized access to a particular workplace. However, organizations may end up using this information about individual employees to track employee attendance" (Young, 2018).

[31] Launched in 2011, Operation Laser (Los Angeles Strategic Extraction and Restoration) used a data-backed approach to extract "offenders" from communities. It used historical information including data on gun-related crimes and arrests as a means to map out 'problem areas' for police officers to focus their efforts on. So, when the police target a specific location, it will naturally generate more stops, crime reports, and arrests, and so the subsequent crime data will lead the algorithm, risk assessment, or data analytic tool to continuously lead police back to the same area. In 2019, Mark Smith (LAPD Inspector General), announced that the criteria used within the Laser program was inconsistent. Operation Laser was discontinued in April 2019 (Bhuiyan, 2021).

[32] Short for Predictive Policing, PredPol was adopted by the LAPD in 2015 until April 2020. The software was devised as a means to predict when and where crimes would occur over a 12-hour period using 10 years of crime data (crimes, dates, locations, and times). Academics have argued that the theory behind the PredPol algorithm is flawed, citing the math as too simple to predict policing because it was only assessing where arrests have previously been made, and then sending police back to those locations (Bhuiyan, 2021; Puente, 2019).

(2019) bring forward sheds light on the risks that emerge when algorithms and data are outdated, incomplete, or not evaluated for bias. Processing such data within an algorithm and enhancing it with artificial intelligence will ultimately produce severe risk to the population. In consideration to this, while the overall effort might have been brought forward with the intention of decreasing crime, its process was inherently flawed and cast a negative shadow over the use of data in policing.

With respect to the above, additional risks in the utilisation of OSINT include those associated with privacy and data protection. According to Sandor (2020), the evolution of Web 2.0 not only created new opportunities for multidisciplinary intelligence collection, but it also instilled risks and vulnerabilities to both personal and organizational data exposures. Although users generally choose what they share on the Web, Sandor (2020) indicates that attention needs to be drawn to potential sensitive data disclosures even if the user restricts general access to their data. This is a critical element to regard as the social aspect of cyberspace makes the exploitation of the Web possible not just from a technical approach, but from a social vantage point as well (e.g. social engineering). As a means to counter this, Sandor (2020) proposes education, training, and well-determined organizational processes to be significant factors in the preservation of privacy and personal data online. Undoubtedly, if knowledge on these matters is taken seriously, the exploitation of data and privacy can be severely limited.

2.5.2. Advantages and Disadvantages

Alongside deterrents to progress are the advantages and disadvantages in utilising OSINT for public safety. As with any technique, advantages and disadvantages are inherently present in the employment of OSINT, such as in the collection, analysis, knowledge extraction, and dissemination phases of its use. Before understanding the advantages and disadvantages of using OSINT, it is beneficial to understand a typical OSINT workflow process. Figure 2.5 depicts a typical OSINT workflow and the type of knowledge that may be elicited from such a process.

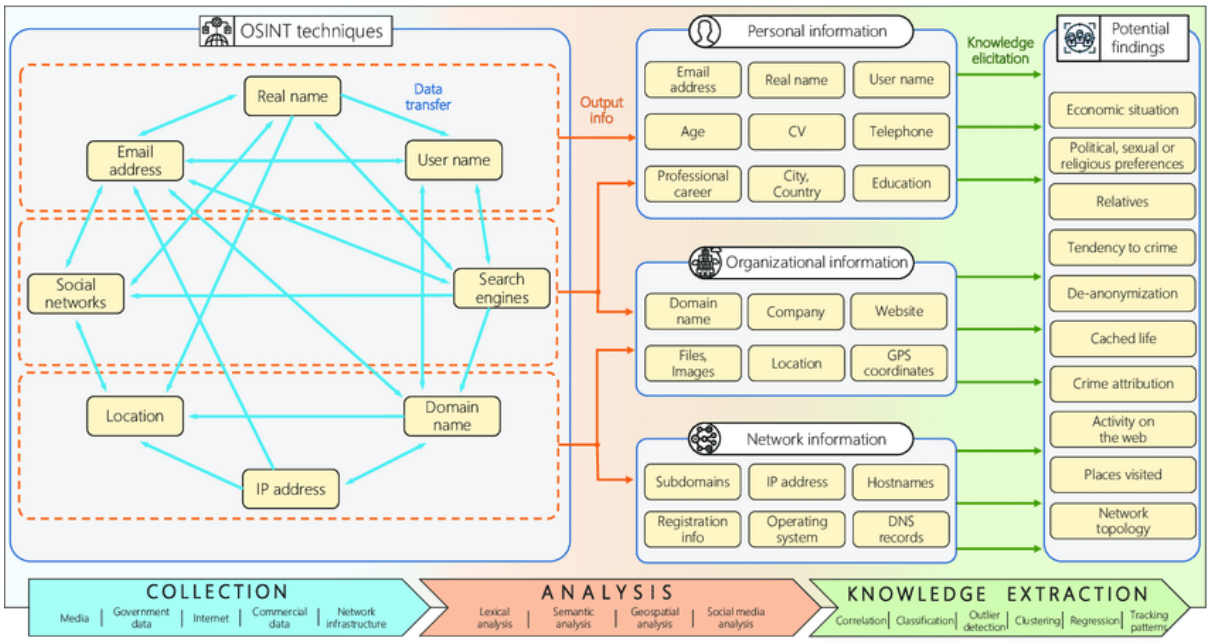

Figure 2.5: Principal OSINT workflows and derived intelligence - (For a larger image, please see Appendix 4)
Source: Pastor- Galindo, et al., 2020, 10287

As depicted in Figure 2.5, Pastor- Galindo, et al., (2020) show possible start points an analyst can focus on (e.g. Email address, username), and how each start point can be linked to the collection of additional data for analysis (e.g. network information). During analysis, the analyst may choose to use a variety of software as a means to extract information from that data (e.g. activity on the web), process it with other information and produce an intelligence assessment based on their findings. Once this is complete, the analyst will continue to develop the intelligence picture as necessary with additional data and informational points in order to have the most current understanding of their operational environment.

Over the course of an OSINT investigation, analysts will generally parse the Web for data related to their operation, with a view to transform basic information into actionable intelligence. According to Qusef & Alkilani (2022) and Cobwebs Technologies (2021),[33] several significant advantages to using open source intelligence include the following[34]:

(1) **Strategically advantageous**: OSINT is not as costly as other traditional forms of intelligence collection, namely Signals Intelligence (SIGINT) and Geospatial Intelligence

[33] Cobwebs Technologies is a leader in the global intelligence market serving national security, law enforcement and the private sector.
[34] Note these advantages are from the perspective of a private organization and may not completely apply to public agencies.

(GEOINT). Each of these have a poor return on investment (especially for smaller organizations) when one considers the cost of consultants and other high-priced tools (Qusef & Alkilani, 2022; Cobwebs Technologies, 2021);
(2) **Abundance of material**: OSINT databases have an abundance of data and information on a variety of topics that can be accessed at any time. Given the nature of the resources and tools used in OSINT, much of the information and data are updated regularly[35] (Qusef & Alkilani, 2022; Cobwebs Technologies, 2021);
(3) **Leakage detection**: OSINT tools may uncover flaws in IT applications, expired applications, accessible ports, detect compromised passwords, vulnerable networks, and operating system models (Qusef & Alkilani, 2022);
(4) **Legal**: The information has been made public with the consent of the original source, so it is entirely legal to gather the data you find[36] (Cobwebs Technologies, 2021);
(5) **Relevant to national security**: Time again, OSINT has proven to be a useful tool for matters relating to national security (Cobwebs Technologies, 2021);
(6) **Big picture view**: Using OSINT techniques can provide a general view of the investigation, which can assist in the creation of long-term strategic planning (Cobwebs Technologies 2021);
(7) **Money-saving**: Regular data collecting tools and techniques can be a large investment. Using OSINT requires little to no financial investment, as by definition it is publicly available information (Cobwebs Technologies, 2021).

To add to this, Pastor-Galindo, et al., (2020) share a similar perspective to Quesef & Alkilani (2022) and Cobwebs Technologies (2021) in that they also identify the abundance of material available as a significant advantage to using OSINT. However, they further add that OSINT data allows for high capacity of computing,[37] big data and machine learning,[38] the availability of complementary types of data,[39] and that the data can have a flexible purpose and wide scope[40] (Pastor-Galindo, et al., 2020, 10285). The main aspect to keep in mind with these advantages however, is the time to train and the capacity to purchase.

In relation to the initial seven advantages identified above, specifically that OSINT is strategically advantageous, it is important to consider that it is not designed to replace SIGINT or

35 This does not discount the fact that one must always be mindful of mis/dis information and social engineering.

36 This may be *technically* true, but there is much debate around the ethics of this, especially as the revelation of some information has the potential to cause harm to individuals (i.e. sexual preference, religious beliefs).

37 Technological advances in computer architecture, processors, and graphic processing units have enabled labour-intensive operations associated with collection, processing, analysis and storage, which has moreover allowed analysts to process and analyse large amounts of publicly available data at a faster rate (Pastor-Galindo, et al., 2020).

38 The proliferation of (and emerging) data mining and analysis techniques, in addition to automation through machine learning has the potential to create more intelligent and efficient investigation and decision making processes (Pastor-Galindo, et al., 2020).

39 Using OSINT techniques is not limited to OSINT data. The intrinsic nature of the process as a whole is broad enough to include external data that has not been obtained from open sources which can complement investigations. For example, classified information can be used to enrich OSINT investigations and profile a target (Pastor-Galindo, et al., 2020).

40 As a paradigm, OSINT can be used in multiple fields (psychologic, economic, journalism, security, etc.). Regarding malicious activity, OSINT could be used to monitor suspicious activity and people, detect radical influencers, investigate attribution of cyberattacks and cyber crime, enhance digital forensics, investigate numerous trends, among many others (Pastor-Galindo, et al., 2020).

GEOINT, but rather complement it. As such, it is important to consider that what Qusef & Alkilani are referring to above is likely most relevant to smaller organizations that do not possess the budget, policy, or authority in place to employ such technical assets to their full potential. On the matter of the second advantage, Qusef & Alkilani make note of the abundance of material available online, but do not account for misinformation, disinformation, social engineering, or deepfakes, which is an aspect of collection that is becoming increasingly relevant. Regarding leakage detection, this is an advantage that is dependent on known vulnerabilities and the technical background to understand the risk of compromised hardware and software. On the matter of OSINT being a legal collection method, there is room to argue some controversy here. Although by definition there is no warrant required to collect OSINT as it consists of publicly available information, it is important to consider the vast amount of information and data available that has been illegally collected and posted on the Web. Take for example, classified documents leaked on the Web (Wikileaks), and breached data (credit cards, passwords, sensitive documents, personal identifiable information, etc). At this time, the literature acknowledges a data breach as an illegal act, however it recognizes the acquisition of the data *after* it has been leaked to be both lawful and useful in OSINT investigations (Volkert, 2019). At the very least, this certainly showcases the "grey space" that one can find themselves in over the course of an investigation, and is likely viewed differently from organization to organization depending on their OSINT policy.

The fifth and sixth advantage are similar to the second advantage (abundance of material) in that analysts must be trained to identify information operations[41] and psychological operations.[42] Any deficiency in this could result in poor assessments that impact strategic, operational, and tactical decision-making which can ultimately cost lives. Finally, the seventh advantage identified OSINT as a money saving technique. While this may be true for rudimentary and some advanced OSINT

[41] NATO and the US Department of Defense define information operations as "the integrated employment, during military operations, of information-related capabilities in concert with other lines of operation to influence, disrupt, corrupt, or usurp the decision-making of adversaries and potential adversaries while protecting our own. Information operations are coordinated military activities in an information environment having specifically defined goals. They represent offensive and defensive measures focused on influencing an adversary's decisions, manipulating information and information systems. They also include measures protecting a country's own decision-making processes, information and information systems" (Joint Chiefs of Staff, 2014, GL-3; Vejvodova, 2019, 84).

[42] "Planned activities using methods of communication and other means directed at approved audiences in order to influence perceptions, attitudes and behaviour, affecting the achievements of political and military objectives" (Ministry of Defence, 2014, 1-1).

techniques, time is a critical factor in all investigations. As a means to being more efficient, it is likely that money will have to be spent on OSINT tools and datasets to streamline the process for effectiveness. Moreover, it is not the intention to denounce any of these advantages, but instead to draw attention to the interdependent variables that exist within the advantages and disadvantages of any tool or technique.

Acknowledging Figure 5.2, Ringin (2022) associates a number of disadvantages to using OSINT techniques for investigations including information overload, identity exposure, and misinformation. In addition to this, Ellis (2018) and Dekens (2020) add that the exposure to vicarious trauma while conducting OSINT is important to consider as well. While the list of disadvantages is not as lengthy as advantages, they should certainly not be discounted. Information overload and misinformation for example may prompt analysts to introduce tainted information into their analysis which may ultimately affect the intelligence picture and decision-making. Similarly, identity exposure is the disadvantage that can put the analyst, the investigation, and the agency at grave risk. Whether through tracking links on websites or by using unapproved software or login credentials, analysts may leave digital breadcrumbs that can allow malicious actors to collect information on them. In this regard, Ringin (2022) raises a critical point by identifying these as disadvantages as they can easily be overlooked, miscalculated, or underestimated.

Moving forward, Ellis (2018) and Dekens' (2020) recognition of exposure to vicarious trauma sheds light on the effect OSINT investigations can have on an analyst's mental health. OSINT investigations can regularly expose an analyst to traumatic audio or images displaying murder, abuse, and terrorism. This exposure can have lasting effects on one's mental health, thus it is rightfully listed as a disadvantage to conducting OSINT. The acknowledgment however, is widely observed in the OSINT community and has garnered a significant response as to how one can prevent, address, and identify vicarious trauma in oneself and others.[43]

While each of the aforementioned disadvantages are unique in their own way, there is also an interdependent factor that links them together: training. Training is particularly interesting on this

[43] For examples, please see Appendix 5.

matter since arguably it is the foundation that mitigates against each of the disadvantages above, which would moreover decrease the risk to the analyst and the intelligence community as a whole. At this point, there is a considerable delta in the availability of resources and courses that train against the mitigation of risk while using OSINT techniques. Although there are a number of OSINT courses available from several reputable sources such as OSINTTechniques.net, OSINT Combine, Toddington, and the SANS Institute, the vast majority focus largely on OSINT techniques with a small portion of training focusing on basic risk mitigation through the use of a VPN, virtual machines, or sock puppet accounts, for example. As previously alluded to, cost and training may perhaps be seen as a disadvantage as well. However, it is important to consider that every technique and tool has an associated cost, but brings additional investigative capability.

Taking into account all that has been discussed on the deterrents, risks, advantages and disadvantages of OSINT, it is important to scale these to determine if and how they may or may not justify the adoption of and utilisation of OSINT. While this concept will be explored in greater detail in later sections of this work, it is worth reflecting on as we move forward onto the next section, which will discuss the literature on the legal, ethical, and policy limitations of OSINT.

2.6. Legal, Ethical, and Policy Limitations for OSINT Collection and Use

Data protection and privacy is regulated by several legal, ethical, and policy documents. While some of these regulations are limited to a particular country, region, province, or state, they are also relevant to those who process the data of citizens who reside in those areas as a means to ensure that private data stays private. According to Block (2021), regulations governing data protection are not meant to prevent the processing of personal data, but instead are meant to balance a citizen's fundamental right to privacy with the requirement of data exchange within society. As a means to achieve this, 137 out of 194 countries have legislation in place that regulates data protection and privacy worldwide (United Nations Conference on Trade and Development, 2021). Some of these documents include, but are not limited to the following:

- **Canada**: Privacy Act, The Canadian Charter of Rights and Freedoms, the Criminal Code, National Security Act, and Personal Information protection and Electronic Documents Act (PIPEDA);
- **USA**: Foreign Intelligence Surveillance Act (FISA), American Data Privacy and Protection Act (AAPPA), California Consumer Privacy Act (CCPA), USA Freedom Act, Electronic Communications Privacy Act (ECPA);
- **Europe**: General Data Protection Regulation (GDPR);[44]
- **Other:** Terms and Conditions (e.g. of social networking sites), Copyright Law.

According to the United Nations Conference on Trade and Development (UNCTAD) (2021), 71% of countries have data protection and privacy legislation, 9% of countries have draft legislation in place, and 15% of countries do not have any legislation pertaining to data protection and privacy.[45] The UNCTAD (2021) emphasizes that privacy and data protection are becoming increasingly relevant as more social and economic activities occur online. The following subsections will review the regulations which are most relevant to the scope of this work.

2.6.1. PIPEDA, Bill C-11, and Bill C-27

Within Canada, data protection and privacy is regulated from two positions: protection from (1) the private sector[46] and (2) the government. For example, the *Personal Information Protection and Electronics Documents Act* more commonly known as *PIPEDA*, is such a document that exists to "govern the collection, use and disclosure of personal information in a manner that recognizes the right of privacy of individuals with respect to their personal information and the need of organizations to collect, use or disclose personal information for purposes that a reasonable person would consider appropriate in the circumstances" (*Personal Information Protection and Electronic Documentation Act*, 2020, 4). This act was legislated as a means to recognize that consumers and the environment in which they are immersed in are inherently distinct from consumers of the past (Penny, 2019). Data and information drive the twenty-first century, so it is no surprise that consumers of digital spaces

[44] GDPR applies to any company that collects and/or processes data on citizens of the EU, regardless of where they are located. (European Commission, nd).

[45] There is no data for the remaining 5% of countries (United Nations Conference on Trade and Development, 2021).

[46] PIPEDA does not generally apply to not-for-profit and charity groups or political parties and associations unless they are engaging in commercial activities that involve personal identifiable information. Additionally, municipalities, universities, schools, and hospitals are usually regulated by provincial laws in Canada, however in certain circumstances, PIPEDA may apply (Office of the Privacy Commissioner of Canada 2019).

should receive protection. Penny (2019) indicates three unique threats and challenges that the "information consumer" faces on a daily basis. First, is the matter that the consumer has become the product. A number of authors have examined this notion and have found that people not only consume information online, but that information about them is consumed as well. Both Srnicek (2017) and Zuboff (2015) have acknowledged how consumer information and data is collected, analysed and then monetized as a means to drive the digital economy - a term which Zuboff (2015) has coined as surveillance capitalism. This is especially of concern as the digital environment is rampant with deception, mis/disinformation, and manipulation. This, in combination with big data analytics and artificial intelligence, can enable private sector companies to influence, manipulate, or distort the information environment which can result in systematic consumer vulnerability (Calo 2014; Zuboff, 2019).

The second issue that Penny (2019) identifies is based on the notion that the online environment where information consumers are active is not simply a business environment, but also a democratic environment as well. When we consider social media, blogs, and news websites, we understand that the digital environment is not simply a business environment, but also a social and informational environment as well. We see this especially within social media platforms that typically serve multiple functions such as connecting with friends and family, following local or international news, and voicing one's own opinion (e.g. Twitter). Authors such as Boyd (2011) have referred to social media platforms as "networked publics" as an attempt to illustrate the indistinct line between public and private spaces. To add to this, Penny (2019) also draws a significant point to consider, describing how digital spaces such as social media platforms are a distinct amalgamation of both consumerism and democracy - noting that the majority of democratic spaces online are owned, operated, and therefore shaped and controlled by private sector interests.

The final challenge that Penny (2019) identifies is the idea that digital spaces - and therefore the information consumer - exists at a time of unprecedented distrust. According to Penny (2019), corporate and government failure to effectively address online threats[47] has essentially created an

[47] Such as data breaches, misinformation, disinformation, and online hate and abuse.

information environment that is corrosive by nature. Powers & Jablonski (2015) have referred to a similar notion as an "information-industrial complex" that leaves consumers with no choice but to accept such harms in order to engage online in a social or democratic fashion. Unsurprisingly, the failure to address online threats has resulted in mass distrust towards digital spaces - so much so that in the United States, fake news has been perceived as a greater threat "than terrorism, illegal immigration, violent crime or racism[48]" (Siddiqui, 2019; Mitchell, et al., 2019).

Although *PIPEDA* was enacted in 2004,[49] as a means to provide data and privacy protection to information consumers, it is evident that the information environment has changed at a more rapid rate than legislation can keep pace with. As a means to protect Canadians from modern threats in the online environment, the Government of Canada had drafted *Bill C-11: Digital Charter Implementation Act*, 2020, and had sat for its First Reading on 17 November 2020[50] (*Bill C-11*, 2020). According to Scassa (2020), *Bill C-11* was necessary for two reasons: first, *PIPEDA*'s weak compliance incentives were made evident by "egregious data security breaches and high-profile misuses of personal information" (Scassa, 2020, np). Secondly, is the "adequacy assessment under the European Union's General Data Protection Regulation [GDPR] … which prevents the flow of personal data outside EU borders unless a substantially similar level of protection is available in the country of destination[51]" (Scassa, 2020, np). Proportional to the *GDPR*, *Bill C-11* was meant to address these issues in ways that *PIPEDA* had not. Unfortunately, *Bill C-11* had expired in 2021 with an incomplete second reading in the House of Commons (*Bill C-11, 2021*).

At present relevance is Bill C-27: the Digital Charter Implementation Act, 2022.[52] According to the Innovation, Science and Economic Development Canada (ISEDC) (Innovation, Science and

---

[48] The Pew report surveyed the question, "% who say ___ is a very big problem in the country today." Made-up news/information scored 50%. Other surveyed answers and their score are as follows: drug addiction (70%), affordability of healthcare (67%), U.S. political system (52%), gap between rich and poor (51%), violent crime (49%), climate change (46%), racism (40%), illegal immigration (38%), terrorism (34%), and sexism (26%) (Mitchell, et al, 2019).

[49] PIPEDA received Royal Assent on 13 April 2000, and came into effect through a number of stages beginning on 1 January 2001. The Act was in full force on 1 January 2004 (Office of the Privacy Commissioner of Canada, 2013).

[50] Bill C-11: An Act to Enact the Consumer Privacy Protection Act and the Personal Information and Data Protection Tribunal Act and to make Consequential and Related Amendments to Other Acts. First Reading, 17 November 2020 (Bill C-11, 2020).

[51] GDPR will apply to Canadian entities who process or control data in connection with offering goods or services (with or without charge) to, or tracking any behaviour of people within the EU (Norton Rose Fulbright, 2018).

[52] The 10 Principles of the Charter are: (1) Universal Access, (2) Safety and Security, (3) Control and Consent, (4) Transparency, Portability and Interoperability, (5) Open and Modern Digital Government, (6) A Level Playing Field, (7) Data and Digital for Good, (8) Strong Democracy, (9) Free from hate and Violent Extremism, and (10) Strong Enforcement and Real Accountability (Innovation, Science and Economic Development Canada, 2022b). For additional details on each of these principles, please see Appendix 6.

Economic Development Canada, 2022a), the *Digital Charter* has been drafted by the Government of Canada as a means to bring Canada's privacy and data protection laws in line with international partners. Once the *Digital Charter Implementation Act* is passed, Part 1 will ultimately replace *PIPEDA* with the new *Consumer Privacy Protection Act* (Innovation, Science and Economic Development Canada, 2022a). Part 2 of this Act will see the establishment of an administrative tribunal[53] that will enforce the new law, and Part 3 will enact the *Artificial Intelligence and Data Act* to guide the responsible development of artificial intelligence within the Canadian marketplace (Innovation, Science and Economic Development Canada, 2022a). Further to this, the *Digital Charter* would also establish part of the current privacy law that regulates the use of digital records by the federal public sector as "stand-alone legislation under the *Electronic Documents Act*" (Innovation, Science and Economic Development Canada, 2022c). Although the *Digital Charter* does seem promising in increasing the data protection and privacy rights of Canadians, in order to be truly effective the Government of Canada must ensure the interpretations, applications, and enforcement of the *Digital Charter* are made clear. Additionally, the Canadian government must also heavily invest in the Personal Information and Data Protection Tribunal in order for it to truly enable privacy enforcement. A failure to do so will likely see large delays accumulating from unnecessary confusion, constrained resources, and a mass number of filings. However, as the *Digital Charter* remains in session within the House of Commons, *PIPEDA* still represents the legal framework for private sector entities to abide by for any data and privacy related matters (Bill C-11, 2021).

2.6.2. The *Privacy Act* and the *Canadian Charter of Rights and Freedoms*

In Canada, two main bodies of legislation serve to protect individuals within Canada against privacy violations from the Canadian government: The *Privacy Act* and the *Canadian Charter of Rights and Freedoms*. The purpose of the *Privacy Act* serves to "extend the present laws of Canada that protect the privacy of individuals with respect to personal information about themselves held by a

[53] The Personal Information and Data Protection Tribunal will serve to enforce the Consumer Privacy Protection Act. Specifically, it will review the Privacy Commissioner's recommendations to impose monetary penalties for certain violations of the Act. Additionally, the Tribunal will act as an accessible means for both individuals and organizations to seek review of any decisions the Privacy Commissioner has made (Innovation, Science and Economic Development Canada, 2022c).

government institution and that provide individuals with a right of access to that information" (*Privacy Act*, 1985). Of particular significance are the following sections within the *Privacy Act*:[54]

- Section 2: Purpose
- Section 3: Definitions
- Section 4: Collection of Personal Information
- Section 5: Personal Information to be Collected Directly
- Section 6: Retention of Personal Information Used for an Administrative Purpose
- Section 7: Use of Personal Information
  Section 8: Disclosures of Personal Information
- Section 9: Record of Disclosures to be Retained
- Section 10: Personal Information to be Included in Personal Information Banks
- Section 11: Index of Personal Information

This is not to downplay the remaining sections of the *Privacy Act*, but rather to draw attention to the ones that are most relevant to utilising OSINT for public safety. The significance of Sections 2 and 3 are that they lay a foundation to the interpretation of the Act. For example, Section 2 indicates that the purpose of the Act is to *extend*[55] the present laws[56] of Canada that protect the privacy of an individual's personal information, and Section 3 defines key terminology, including "personal information," "personal information bank," and "government institution" among various others. Sections 4 and 5 relate to how the government may *collect* an individual's personal information, while sections 6-11 are concerned with how the government may retain, use, disclose, and index personal information.

These sections, paired with the *Canadian Charter of Rights and Freedoms* (1982) - especially Section 8, the right to be secure against unreasonable search or seizure - serve as the bedrock that protects the personal information of individuals within Canada. However, with the exception of several minor revisions, the *Privacy Act* has not evolved to reflect the modern digital environment. As indicated in Section 75 of the *Privacy Act* (1985), the administration of the Act must be evaluated on a permanent basis, and a report to Parliament including a statement of recommendations must be submitted. Despite this, the *Privacy Act* has not had any substantive updates since it was introduced in

---

54 The headings of each section are provided here. For greater detail on these sections - including subsections and exceptions - please see the *Privacy Act*, 1985 available at https://laws-lois.justice.gc.ca/eng/ACTS/P-21/page-1.html#h-397177

55 Emphasis added.

56 For example, section 8 of the Canadian Charter of Rights and Freedoms indicates that "Everyone has the right to be secure against unreasonable search or seizure" (*Canadian Charter of Rights and Freedoms*, 1982).

1983 and revised in 1985. While the Office of the Privacy Commissioner of Canada has urged the Canadian Government to update the Act, its recommendations have not yet been heeded (Office of the Privacy Commissioner of Canada, 2022a). In their end of term statement, then Privacy Commissioner of Canada, Daniel Therrien, indicated that the state of privacy within Canada was in a condition of uncertainty (Office of the Privacy Commissioner of Canada, 2022b). In that same statement, Therrien identified a resistance to change and provided some examples of how new privacy laws should evolve, indicating that they should seek to:

- "Reintroduce the knowledge and understanding element of meaningful consent;
- Stipulate that personal information may not be collected for any purpose determined by an organization but only for "specific, explicit and legitimate purposes";
- Prescribe an objective standard for accountability, namely the obligation to implement a privacy management program "to ensure compliance" with the law;
- Authorize the OPC, like many other data protection authorities in Canada and abroad, to conduct proactive audits to verify compliance with the law; …
- Enable responsible innovation;
- Adopt a rights-based framework;
- Increase corporate accountability;
- Adopt similar principles in public and private sectors laws – given the increased reliance on public-private partnerships;
- Ensure Canadian laws are interoperable, internationally and domestically; and
- Adopt quick and effective remedies, including giving my office[57] the authority to make orders and to impose monetary penalties where warranted; …
- The Office of the Privacy Commissioner should have similar powers and be subject to similar appeal procedures as its provincial counterparts, if the federal office is to remain an influential and often unifying voice in the development of privacy law in Canada" (Office of the Privacy Commissioner of Canada, 2022b).

In addition to this, a review of the *Privacy Act* in 2016 identified twenty-eight recommendations[58] identified by the House of Commons Standing Committee on Access to Information, Privacy and Ethics, and eleven recommendations from the Privacy Commissioner of Canada (Parliament of Canada - House of Commons, 2016). Of particular interest are recommendations 14, 17, 18, and 20 from the Standing Committee and recommendations 2, 4, and 7 from the privacy commissioner. If implemented, these recommendations alone will greatly enhance privacy in Canada as they will legally obligate the government to: safeguard personal information,

[57] In reference to the Office of the Privacy Commissioner (OPC).

[58] See Appendix 7 for a complete list of the recommendations.

create explicit requirements for collection, conduct privacy impact assessments, consult the Office of the Privacy Commissioner of Canada on legislation and regulations with privacy implications, and, conduct ongoing five-year parliamentary reviews of the *Privacy Act*. At a time when it is increasingly easier to collect, analyse, retain, and share large amounts of personal information, it is also a time when increased risks of privacy violations may occur as a result of complacency. In addition to the recommendations listed above, Commissioner Therrien had the following to say regarding the risk of not updating the Act:

> "In the public sector, these consequences include, first, risks of data breaches that are not properly mitigated; second, excessive collection and sharing of personal information, which may affect trust in government; and more specifically, third, a reduced trust in online systems that may undermine the government's efforts to modernize its services and coordinate its digital communications with Canadians.
>
> Some governments have already moved forward to strengthen their privacy protection frameworks, most notably the European Union. There is a risk, in my view, that if European authorities no longer find Canada's privacy laws essentially equivalent to those protecting EU nationals, commerce between Canada and Europe may become more difficult" (Parliament of Canada - House of Commons, 2016, pp. 3-4; Standing Committee on Access to Information, Privacy and Ethics, 2016, 2).

Commissioner Therrien's remarks illustrate the necessity to revise the *Privacy Act* in order to sustain personal information and privacy rights in an environment that is steadily reliant on sharing information in a networked domain. It is critical that at such a time we do not disregard this necessity as its impacts will expand past Canadian borders.

As OSINT becomes increasingly relevant to public safety, so too will privacy laws be central to information collection, analysis, retention, dissemination, and sharing. With data-driven systems and technologies increasingly being employed, it is crucial to develop a means by which Canada can extract benefit from publically available and/or open source information and data while having a clear understanding of an individual's fundamental rights within a democratic society. As reflected in the recommendations, there are numerous gaps within the *Privacy Act* that leave Canadians vulnerable to privacy violations. Much of this has largely been observed as a result of the COVID-19 Pandemic, and others as a result of the incorporation of commercial applications within public institutions. In a number of ways, the pandemic altered one's daily routine: we began to attend Doctor's appointments

online, and children and adults were learning in a digital environment. These actions established new vectors of risk to privacy that are not covered in the Privacy Act or PIPEDA as neither legislation offer protection within the context of public-private partnerships, nor do they mandate application developers to "consider Privacy by Design, or the principles of necessity and proportionality" (Office of the Privacy Commissioner of Canada, 2021, np). As a result of this, doctor-patient confidentiality, contact-tracing applications, vaccine passports, e-learning platforms, etc., were not truly protected by any privacy framework within Canada.

However, public-private partnerships have been a contested issue prior to the pandemic as depicted in the Office of the Privacy Commissioner's (OPC) investigation of the Royal Canadian Mounted Police's (RCMP) use of the Clearview AI platform. This platform essentially scraped billions of images on the World Wide Web, and allowed the RCMP to compare and match these images against its database. In its investigation, the OPC had found that the aggregation of biometric information occurred without the knowledge or consent of the individuals, and was therefore illegal and a violation of the *Privacy Act* (Office of the Privacy Commissioner of Canada, 2021). In addition to this, the investigation also disclosed extensive gaps within the RCMP's systems and policies "to track, identify, assess and control novel collections of personal information through new technologies" (Office of the Privacy Commissioner of Canada, 2021, np). This investigation alone sheds light on the benefit of the previously mentioned 7th recommendation from the Privacy Commissioner- that government agencies should deliver a privacy impact assessment for new programs to the OPC prior to any implementation. Nonetheless, as the World Wide Web and other digital applications become more prevalent in society, it is clear that Canada's legislative framework governing privacy will need to change in order to protect Canadians.

2.6.3. Ethics and Policy

One of the main concerns with utilising open source information is the degree by which data and information should be regarded as public or private. According to Townsend & Wallace (2016), one of the main arguments that shape the perspective of researchers and analysts is that users of Web

2.0 platforms have agreed to the terms and conditions of the domains they use - some of which include clauses on how the user's data may be accessed and examined by third parties. In their research, Townsend & Wallace (2016) interviewed researchers on this notion and discovered a consensus among them in support of the use of public data with the justification that through agreeing to the terms and conditions, users have acknowledged that their data may be acquired by third parties. However, it is widely known that most users do not read the terms and conditions in full, and so this poses some debate over this perspective. For example, Boyd & Crawford (2012) claim that it is difficult for third parties to vindicate their activities as ethical solely on the basis that the data is publicly accessible. This therefore raises questions toward whether users of Web 2.0 maintain a reasonable expectation of privacy through the use of private accounts, password protected/invite-only groups, and forum discussions (Hewson & Buchanan, 2013; Hewson, Buchanan, et al., 2021).

Also in relation to the matter of public and private data is the concept, risk of harm. In 2012, Markham and Buchanan of the Association of Internet Researchers had proposed that the risk of harm to Web 2.0 users is increased when dealing with sensitive data that could lead to "embarrassment, reputational damage, or prosecution" (Markham & Buchanan, 2012; Townsend & Wallace, 2016, 7). However, it is important to consider that the risk of harm may not be limited to a specific user, but also to that particular user's network through friends, followers, likes, shares, and comments. While it may be relevant to investigate links between the user and their network, it is therefore important to ensure that both relevant and irrelevant information on associated users is treated in accordance with the relevant privacy laws in place throughout the conduct of the investigation and thereafter.

2.6.4. Conducting Ethical OSINT Investigations

While conducting ethical open source investigations for public safety may be challenging, it is entirely possible through the implementation of processes and principles that could be adapted into an organization's standard operating procedures (SOPs). In this regard, Charles Brown (2023) has advised on the use of four principles:

(1) **Use publicly available data where possible**: By nature, open source investigations use open source data and information that are either freely available or publicly licensable.

(2**) Ensure that investigations are targeted**: Indiscriminate collection of data and information can provoke risk in the form of non-compliance with privacy and data legislation and regulations, and may also result in reputational damage. The direction of clear research parameters is a means by which to avoid this. In the intelligence community, this can be achieved by adhering to the intelligence cycle and maintaining an intelligence collection plan.

(3) **Ensure decisions are made by humans**: There is an increasing unease surrounding the idea of non-human decision making (e.g. artificial intelligence). Although technology is vital to investigations, it is more ethical to utilise it for the more "manual" parts of an investigation, such as data collection and mapping networks. Essentially, technology should assist the analyst in obtaining the information required in order to make an assessment in a timely manner.

(4) **Remain accountable through recording all actions taken**: Recording your collection process allows an analyst to show and justify processes taken. Additionally, it can be used to support an investigation if required by law enforcement (Brown, 2023; Blackdot Solutions, 2021).

Ethical principles such as these enable analysts to be accountable to their research and collection, which is becoming more relevant as data is more easily accessible. Fundamentally, ethical principles are those which serve as a moral compass in how one engages in their duties.[59] In the same way the Hippocratic Oath guides medical practitioners on the "beneficence, non-maleficence, justice and respect for the patient's autonomy with its two rules of confidentiality and veracity," so to should ethical principles guide analysts to ensure collateral harm is not unduly brought upon innocent civilians (Rancich & Gelpi, 1998, 147). In order to do this, the challenges that prevent analysts from producing ethically sound work must be addressed.

When interviewing analysts and journalists about ethical challenges they perceive in their work, Kenausis (2021) identified a number of themes. The first theme was identified as an existing tension between speed and quality, especially as it relates to the pressure to quickly publish findings and assessments without the ability to thoroughly review any analysis (Kenausis, 2021). A second theme was centered on privacy and safety, especially in regards to protecting sources from harm that could be realized as a result of photos or posts that could identify a specific user as the source of the

[59] A number of professional communities (e.g. medicine, law, engineering, etc.) that maintain power over others have ethical norms and codes that guide their actions, maintain trust, facilitate accountability, and moreover steer actions toward moral good (Lorhrke, et al., 2021).

information (Kenausis, 2021). Third is the theme of unintended consequences as it relates to a concern that findings and assessments could be used to drive political narratives or inadvertently aiding nefarious groups[60] (Kenausis, 2021). Finally, the fourth theme is vicarious trauma, where a number of analysts and managers had identified their concern toward the psychological and emotional impact as a result of frequently observing violent material, including the ethical implications of tasking their colleagues to analyse such material (Kenausis, 2021). Together these themes represent key challenges that analytical teams are confronted with on a regular basis, and therefore show that ethical dilemmas within open source research are quite common. For greater awareness on some of the major ethical challenges encountered during open source investigations, several excerpts[61] from the Loehrke, et al., (2021) interviews with analysts and journalists have been provided below, with additional excerpts provided in Appendix 8:

> Background: The Ethical Journalism Network and the Stanley Center for Peace and Security interviewed 8 journalists and 20 analysts with a view to obtain examples of circumstances in which they or their teams had felt uneasy about publishing findings from their open source investigations (Loehrke, et al., 2021, 5).
>
> Excerpt 1: Diplomacy, Crises, and Conflict
>
> In these interviews analysts had discussed situations when they were concerned about how something they published might inadvertently create diplomatic tension, or result in an escalation of pressure during ongoing crises (Loehrke, et al., 2021, 5).
>
> Excerpt B: North Korean Missile Base
>
> "There was this North Korean missile base. Hardly secret. The base had been there a while. I wrote an update on the site, based on some recent imagery. The report ended up getting some media coverage, reframed, and put in context with ongoing diplomatic efforts. This was during some sensitive months between the US, South Korea, and North Korea.
>
> The media coverage of the report ballooned. The story got big enough that the US and South Korean governments commented on the missile developments.
>
> I thought it was useful analysis. Nothing flashy. But it took on a life of its own after we published." (Loehrke, et al., 2021, 6).
>
> Excerpt 2: Unintended Consequences
>
> In some situations, analysts may inadvertently discover information that, if published, could likely assist nefarious actors, or result in collateral damage to systems in place that are designed to sustain a common good. For a number of reasons, these situations can get difficult to recognize (Loehrke, et al., 2021, 6).

[60] For example, an investigated subject may be able to discern monitoring methods, equipment, technological / analytic capability, data sources, etc. (Kenausis, 2021).

[61] In the interest of space, only one excerpt from each topic is shown here. Please see Appendix 8 for additional excerpts.

Excerpt A: Tracking Terrorists in Africa

“I was using satellite imagery to track a terrorist organization in Africa. I identified some activity but also inadvertently discovered a US and allied military operation. I shared it with a contact, who asked me to not publish because it’d compromise the operation.

More than happy to do that. Not a problem. That’s why we do reviews like that.” (Loehrke, et al., 2021, 6).

Excerpt C: Russian and Syrian Bombing Sorties

“We obtained cockpit recordings of Russian and Syrian pilots on bombing sorties. Using some open source and geospatial mapping, documenting evidence and cross-corroborating data, we could validate the tapes. And show Russian involvement in strikes on hospitals in Syria.

We reported it. They bombed one of the hospitals again. We reported it, again. Our story was held up to the Russian ambassador in a UN Security Council meeting. Soon after, Russia withdrew from an additional protocol to the Geneva Conventions and, later, the UN deconfliction system.

Obviously, they’re aware now that there’s skilled groups like us who can call them out for war crimes using evidence that they can’t deflect like they used to.” (Loehrke, et al., 2021, 6-7).

Excerpt 3: Privacy

As open source information makes it easier to identify individuals and sources, it is getting much more difficult for teams to protect the identity of their sources. This issue is compounded by the growing requirement to consider the safety and security of sources, subjects and bystanders (Loehrke, et al., 2021, 7).

Excerpt A: Protecting Sources

“We have to be extremely careful when using bystander video in our reporting.

Immediately after an event—say an explosion or some act of violence—it’s fairly common to find firsthand video of the event posted on social media. It can be immensely valuable for verifying a story and adding visuals. But by showing video evidence and how we verify a story, it can endanger the person who uploaded the video.

We were doing a story about a missile strike in a country in the Middle East. We found video of the strike and got in touch with the person who took the video. We wanted to include the video in our reporting. But based on the video, it wouldn’t be hard to figure out which building, apartment, or window our contact was standing in when filming. That could get the person arrested or bring harm to a family.

In this case, we didn’t publish the video with our reporting.

But you have to anticipate this stuff. You can ask your source, “Is this okay? Can we do this?” You can ask the same questions internally. But it doesn’t change the reality that a wrong choice could put someone in danger.” (Loehrke, et al., 2021, 7).

Excerpt 4: Interaction Between Analysts and Journalists

Some open source analysts may work with journalists as expert sources as a means to ensure the perspective the journalist is providing is valid, precise, and within the correct context. However, things can still be misinterpreted. In the interviews, analysts and journalists both

> acknowledged their lack of coordination on their shared responsibilities if ethical challenges emerge (Loehrke, et al., 2021, 8).
>
> Excerpt A: Misinformation
>
> "You have to be really careful with what you're saying and not.
>
> There was a recent report about an explosion at a nuclear facility in Iran. Some people in the non-proliferation community used geospatial data to validate the location. Those analysts did nothing wrong. It was fine.
>
> But a reporter picked up the story of the explosion and then just sprinkled in, like, "Oh, by the way," an international inspector was detained last year with an assertion that they detected nitrate residue on the person's hands.
>
> That reporting was horribly irresponsible. All the facts were correct but presented in a way that creates the false impression that an international inspector had planted a bomb in a nuclear facility. That's how misinformation starts, and it could erode trust in the essential work that inspectors do." (Loehrke, et al., 2021, 8).

Together, these excerpts identify a small portion of the ethical challenges that are ever present in open source investigations. This however, should not invalidate open source analysis as a method for contributing to public safety, but should instead draw attention to the requirement for ethical considerations to be taken into account alongside analysis. While this often falls upon a single analyst's responsibility, these excerpts clearly illustrate the benefit of peer deliberations to identifying multiple stakeholder effects - something a single analyst should not be tasked with or navigate in isolation.

2.6.5. Navigating Ethical Challenges

Navigating ethical challenges in open source investigations appear to be gaining more attention as such methods grow in popularity. According to Kenausis (2021), only a handful of organizations possess a code of ethics to assist in decision making. The evidence cited here is specifically in reference to analysts within international security policy fields and those tracking non-proliferation, compared against "more-established fields" such as journalism, where the former often have little to no ethical guidelines in place (Kenausis, 2021).

In their interviews, Loehrke, et al. (2021) have established that while analytic training is predominately available, training on ethics is not. Their interviewees had discussed the availability of technical training and workshops focused on data analysis, platforms, sensors, and types of data, but

during such training, coaching on ethics within open source investigations was simply not present. As a result of this, Loehrke, et al. (2021) indicate that analysts have largely adopted an informal means to navigate any ethical dilemmas within their work (Loehrke, et al. 2021). Commonalities of these exist between Kenausis (2021) and Loehrke, et al. (2021) that identify analysts substituting alternative processes for ethics. For example, Kenausis (2021) and Loehrke, et al. (2021) both observed that (1) analysts will often reach out to trusted peers for their opinion on a situation and (2) analysts often treat accuracy and transparency as an adequate substitute for ethics. While peer review in addition to the pursuit of accuracy and transparency are key to any investigation, they are both inherently distinct from the evaluation of any ethical implications that may arise from publishing an investigation. It is apparent that within the literature on ethics in open source investigations, that this distinction has not been made clear enough, which is likely a result of a lack of clear and widely accepted ethical guideline for open source research.

In their work, both Kenausis (2021) and Loehrke, et al. (2021) reference two ethical frameworks that may be utilised as an ethical standard in the open source community. These frameworks are the Markkula Framework for Ethical Decision Making and the Berkeley Protocol on Open Source Investigations. Kenausis (2021) maintains that incorporating ethical guidelines such as the Markkula Framework will provide analysts with the preparation required to confront ethical dilemmas in their investigations. In short, the Markkula Framework is centered on five essential principles:[62] (1) Identify the ethical issues, (2) Get the facts, (3) Evaluate alternative actions, (4) Choose an option for action and test it, and (5) Implement your decision and reflect on the outcome (Markkula Center for Applied Ethics, 2021). Combined, these principles are intended to provide analysts with a means to explore and identify ethical courses of action while considering six ethical lenses[63] to understand the varying ethical dimensions at play (Markkula Center for Applied Ethics, 2021).

[62] Please reference Appendix 9 for further details of the Markkula Framework.

[63] The Six Ethical Lenses which the Markkula Framework utilises include: (1) The Rights lens, (2) The Justice Lens, (3) The Utilitarian lens. (4), The Common Good Lens, (5) The Virtue Lens, and (6) The Care Ethics Lens (Markkula Center for Applied Ethics, 2021). Descriptions of each one can be found in Appendix 9.

Akin to Kenausis, Loehrke, et al. (2021) also recognize the lack of ethical guidelines among the open source community. However, they suggest an alternate ethical framework to provide analysts with the tools required to confront ethical issues: The Berkeley Protocol on Digital Open Source Investigations. Co-published by the Office of the United Nations High Commissioner for Human Rights (OHCHR) and the Human Rights Center at the University of California, Berkeley, School of Law (2022), this document emphasizes professional,[64] methodological,[65] and ethical principles[66] that should be administered while identifying, collecting, preserving, analysing, and presenting open source information for criminal and human rights investigations. At its core, the Berkeley Protocol speaks to the proliferation of publicly available data and the capacity to which it can have as presenting evidence of human rights violations, war crimes, disinformation campaigns, and other global incidents. In reference to the Protocol as a whole, Renshaw & Chaplin (2020) assert that the framework is "mindful of the pace of technological development … avoids prescriptive, tool-specific rules, and instead proffers a legal, ethical, methodological and security-based framework for those carrying out open-source investigations" (np). As a result of these considerations, the Berkeley Protocol arguably enhances an analyst's ability to conduct ethical open source investigations with modern techniques and skill sets that can adapt to technological change. With consideration to this, it is reasonable to assert that the Berkeley Protocol is an ethical framework that is more fitting to open source investigations than the Markkula Framework, which is more theoretical in nature. However, the Markkula Framework does provide a solid philosophical base that can serve to ground an analyst's ethical thinking and mindset.

[64] Professional Principles: Accountability, Competency, Objectivity, Legality, and Security Awareness (Office of the High Commissioner for Human Rights, and the Human Rights Center at the University of California. Berkeley, School of Law, 2022, 11-15). For additional details please reference Appendix 10.

[65] Methodological Principles: Accuracy, Data Minimization, Preservation, and Security by Design (Office of the High Commissioner for Human Rights, and the Human Rights Center at the University of California. Berkeley, School of Law, 2022, 11-15). For additional details please reference Appendix 10.

[66] Ethical Principles: Dignity, Humility, Inclusivity, Independence, and Transparency (Office of the High Commissioner for Human Rights, and the Human Rights Center at the University of California. Berkeley, School of Law, 2022, 11-15). For additional details please reference Appendix 10.

2.6.6. Policy

In consideration to the sensitivities associated with open source investigations, it is vital for private and public organizations to develop a policy that clearly outlines how investigations are carried out. According to Michael Bazzell (2022), this type of document should at the very least outline the approval process, a minimum training standard, and an appropriate use policy. Additionally, it would also be appropriate for such policy documents to outline the tiers at which investigations may occur. For example, recall from section 2.2 that the Royal Canadian Mounted Police (2021) had developed the three tier structure. By separating investigations into various tiers, the RCMP ensures that its analysts are better prepared to understand the arcs within which they are legally, ethically, and professionally able to investigate. Furthermore, these tiers can assist in ensuring that only properly trained analysts are assigned to conduct work at the level to which they are trained and authorized, which ultimately protects both the analyst and the organization.

In addition to the above, it is important to consider that definitions of actions such as "friending" or "following" may also be required within an open source policy document as it will assist in defining the types of interaction that are acceptable. According to Bazzell (2021), headings to consider within a policy document include, but are not limited to: Statement of Intent, Definitions, Appropriate Use, Approval Process, Training, Retention/Audit, Agency Forms, Acknowledgement of Technological Advancement, Revision History, and an Appendix. Of course, depending on the type of investigative work being conducted, it is likely for analysts to come across abusive material or material which may indicate a crime of a specific nature which the analyst's organization does not have a mandate to investigate. In this regard, organizations should consider how this type of material should be dealt with and engage with the relevant organizations beforehand so a practice is in place and is easily executable in a short period of time. Apart from Bazzell's (2021) Child Sexual Abuse Material (CSAM) policy - which aims to provide a baseline procedure for the proper handling of illegal digital material - at the time of writing, there does not appear to be ample research available to review on this matter. This is likely a result of the sensitive nature of open source policy documents, and the lack of academic discourse on this specific topic. However, in such situations it is critical to

be mindful and anticipatory of what one may find online, and therefore consider appropriate "actions on"[67] discovery guidelines.

2.7. Conclusion

This section has reviewed the literature on open source intelligence, especially as it relates to the value it brings to investigations associated with public safety and the capabilities associated with varying open source tools. Additionally, it also explored tradecraft, cultural, resourcing, and partnership deterrents to progress, and the legal, ethical, and policy limitations of utilising OSINT.

As a whole, the key takeaway of this literature review is that the span of the Internet and the World Wide Web have enabled open source investigations to be an effective and efficient tool for real-time investigative and analytical capabilities. Regardless of this however, the literature has drawn its largest gaps from the existing legal, ethical, and policy issues regarding current data and privacy laws. Thus, the use of open source intelligence techniques must be applied to an appropriate ethical framework standard to the work environment in which it is carried out. This in particular is crucial to uphold not only while privacy and data protection laws are revised and updated in Canada, but also so that any evidence collected as part of an open source investigation proceeds with the evidential requirements necessary to be deemed admissible in court.

In relation to this dissertation, my research will address these gaps through two approaches. First, this work will conduct case studies that will situate the reader to understand how OSINT may be used for situational awareness, finding a person, group, or location, and for verification. Secondly, it will also address these gaps by conducting Access to Information and Privacy (ATIP) Requests to several government agencies in order to understand the policies and authorities that govern their OSINT collection capabilities. Additionally, where available, the ATIPs will also address the means of training by which various government departments use to equip their analysts with the proper tactics, techniques, and procedures for conducting OSINT investigations.

[67] "Actions on" are rehearsed actions that allow an individual to quickly react to a situation they encounter. Any "action on" should describe what the analyst should do, how they should do it, and how to communicate the situation up, down, and across their team.

Determining the value of open source intelligence for public safety is an area of academic discourse that is becoming more relevant and significant as databases of information become more widespread, and therefore inherently more exploitable. This research will contribute to new knowledge by identifying issues and vulnerabilities relevant to OSINT collection and by recommending approaches to mitigate against those issues and vulnerabilities. The intent of this work is to contribute to awareness of how OSINT can be used as an effective tool that can have an impact in how public (and private) agencies can contribute to public safety. Moreover, this work has drawn on a number of existing theories and methods in order to build a framework for this research to ensure the intent of the work is met. The Intelligence Cycle[68] and social network analysis[69] are in many ways foundational to how this research was conducted, how the results were interpreted, and how recommendations were approached.

[68] The intelligence cycle is a cyclical process that models how intelligence analysts carry out their directed tasks. As per the National Defence Land Force Information Operations Intelligence Field Manual, the cycle contains four steps: 1) Direction; 2) Collection; 3) Processing; and 4) Dissemination. All four phases of the cycle may occur concurrently and be revisited if required (National Defence, 2000).

[69] See Chapter 3 - Theoretical Framework.

Chapter 3: Theoretical Framework

3.1. Introduction

Modern interaction occurs on a social-technological domain. How we connect, communicate, share our thoughts, and keep informed is largely serviced by applications and platforms such as Facebook, Twitter, Instagram, TikTok, and Reddit. As these applications and platforms are used, they increasingly become databases of information on the people, places, and things that we interact with on a daily basis. Bearing this, and considering the initial explanation of the Internet and World Wide Web in Section 2, our online interactions can be represented by networks. When forensically examined, these networks are able to provide a significant amount of information about a person, place, or incident of interest that can significantly inform analysis in ways that traditional intelligence disciplines cannot. In consideration of this, this dissertation will utilise network theory as the lens which this work will observe the problem through. As such, this chapter will identify key concepts and provide an evaluation of theory and how it is relevant to this research.

3.2. Key Concepts

Network theory is a derivative of graph theory, and is employed in numerous fields including sociology, physics, computer science, public health, and neuroscience. Its applications are vast and include the distribution of the electric grid, logistical networks, social networks, the Internet, and the World Wide Web. Within network theory are various means of analysis. This work will largely focus on social network analysis as this perspective provides a manner of analysing a network's dynamics through a social lens characterised by nodes and links. This section will introduce a minimal set of concepts used in social network analysis.

Networks can be both simple and complex in nature. In their most basic form, a network represents a relationship, or rather, a linkage of some fashion. In essence, networks are dynamic objects that are constantly evolving and changing as a result of the activities or decisions of the components they are made from (Watts, 2003). In more formal terms, a network is made up of nodes (also called objects) and a description of the relationship between the nodes (Kadushin, 2004). For

example, Figure 3.1 depicts a very basic network with two nodes (N1, N2), and a linking relationship. In this example, N1 and N2 may represent two people. Relationship 1 may represent two people who are in the same class together. Relationship 2 may represent a single directed relationship between N1 and N2, but not between N2 and N1 (e.g. you may follow a celebrity on Instagram, but they do not follow you back; or, you may supervise someone who does not have any other subordinates). Relationship 3 may represent a symmetrical relationship between N1 and N2 (e.g. you are friends with someone on Facebook, and they are friends with you). These arrows - hereby called links[70] - represent the direction of the relationship within a network.

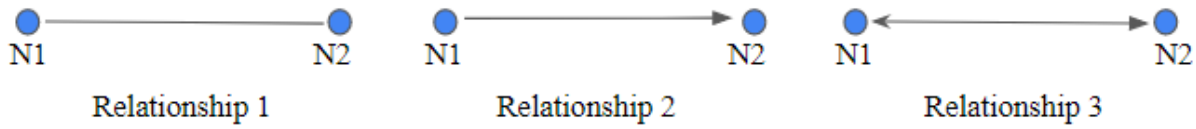

Figure 3.1: Link and Node Relationships in Networks
Source: Kadushin, 2004

When describing a network, it is also important to describe the distance between nodes. For example, Figure 3.2 represents a transitive relationship between three nodes. In this network, there are two steps between N1 and N3. However, since N1 has a reciprocal node to N3, this network can be said to be transitive (balanced) because N1, N2, and N3 are directly linked together.

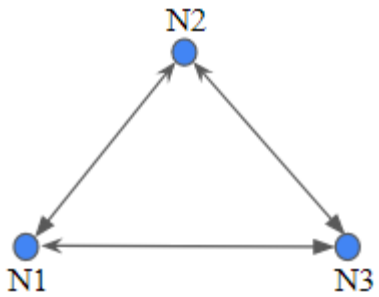

Figure 3.2: A Transitive Network
Source: Kadushin, 2004, 6

[70] "Edges," "ties," "connection," or "arcs" are synonymous with "Link" and used interchangeably throughout network theory and graph theory.

By examining the relations between nodes, social network analysis aims to explain social phenomena that may not be otherwise apparent. With the above basic concepts in mind, the following are a list of common and relevant terms which are applicable to the scope of this work.

Common Patterns Amid Ties

- **Direct Link**: A path that directly links nodes together (Kadushin, 2004, 35) (e.g. Node 1 is directly connected to Node 2 in Figure 3.1, Relationship 3).

- **Indirect Links**: A path that links a relationship via another node that is not directly connected to the initial node. (e.g. Node 1 is connected to Node 3 via Node 2 as depicted in the following illustration below)

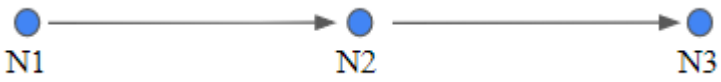

Figure 3.3: An Indirect Network
Source: Kadushin, 2004, 35

- **Multiplexity**: A link between two actors that services multiplicity interests. For example, two people who work together, but also happen to be friends are said to have a multiplex relationship (Kilduff & Tsai, 2003).

- **Strength**: The strength of a tie (link) is often defined in combination of time, intimacy, and reciprocity. Strength may range from weak (infrequent and distant links) to strong (frequent, lost-lasting) (Granovetter, 1973).

- **Direction**: Links may have an arrow to indicate the direction of the relationship, or a means by which nodes can reach other nodes (Kilduff & Tsai, 2003).

- **Symmetry (reciprocity)**: Symmetry and reciprocity represent balance in the network. For example, a network that has a high degree of reciprocity represents symmetric ties between nodes. So, if N1 likes N2 and N2 also likes N1, the link is symmetric (reciprocal). Networks that have a low degree of reciprocity may indicate a hierarchical network, with some nodes receiving likes that they do not reciprocate (Kilduff & Tsai, 2003).

- **Bridge**: A bridge is a concept within Granovetter's explanation of the strength of weak ties. Here, Granovetter asserts that weak ties "are disproportionately likely to be bridges as compared to strong ties, which should be underrepresented in that role. This does not preclude the possibility that most weak ties have no such function" (Granovetter, 1982, 130). Bridges may also serve as conduits, carrying information and influence to clusters that would otherwise not receive such influence (Kadushin, 2012, 58).
    - **Strength of Weak Ties**: Granovetter had hypothesized that important information and opportunities flow from nodes (people) that have limited links (Kadushin, 2012, 52). According to Granovetter (1982), "Our acquaintances ("weak ties") are less likely to be socially involved with one another than are our close friends ("strong ties"). Thus the set of people made up of any individual and his or her acquaintances

will constitute a low-density network (one in which many of the possible ties are absent), whereas the set consisting of the same individual and his or her *close* friends will be densely knit (many of the possible lines present)" (Kadushin, 2012, 56; Granovetter, 1983, 201-2).

- **Structural Hole**: A lack of connections, or links - opposite of density (Kadushin, 2012, 52). "A gap between two actors or two clusters of actors (A and B) that can be spanned by another actor (C) who may, thereby, become the only member to belong to both A and B (if these are clusters), or who serves as the only intermediary between A and B" (Kilduff & Tsai, 2003, 136).

Common Measures of Individual Actors

- **Node (synonymous with vertices, entity, actor)**: A set of objects that are connected together to form a network (Nykamp, 2022; Kadushin, 2012). See Figure 3.4 below.

- **Link (synonymous with arc, edge, tie, connection)**: Refers to the relationship between nodes. Links may also be directed or undirected (See Direct Link and Indirect Link above) (Nykamp, 2022). See Figure 3.4 below.

- **Hub**: A node that has a large degree (many links) (Nykamp, 2022). See Figure 3.4 below.

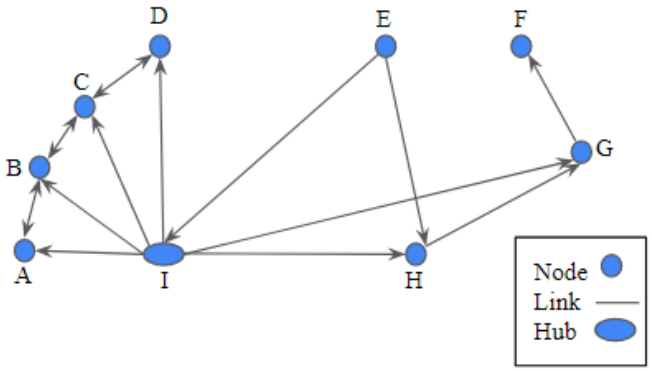

Figure 3.4: Nodes, Links, and Hubs
Source: Author rendered
Note: The Hub is enlarged in this network for illustration purposes, and may not necessarily be enlarged in other network depictions.

- **Degree**: Degree centrality refers to the number of links which a node has within a network, and is usually broken down into two components: indegree centrality (the popularity of the node; the amount of people who ask an individual for advice) and outdegree centrality (the amount of people that person might give advice to). A higher value indicates a greater degree of centralization over a central point(s) (Kilduff & Tsai, 2003).

- **Closeness**: Closeness centrality is the extent by which the most direct path that connects one node to each of the other nodes in a network is short instead of long. This type of measurement is only relevant for a fully-connected network that has no isolated nodes. A closeness centrality score that is high conveys a node that can access multiple other nodes and can therefore be said to be relatively independent from the control of other nodes (Kilduff & Tsai, 2003).

- **Betweenness**: Betweenness centrality is focused on how a node functions as a link for other nodes not directly connected to it - a sort of "go-between" or intermediary for other nodes in the network that creates the shortest connecting path between nodes (Kilduff & Tsai, 2003).

- **Distance**: The distance across a network between nodes. If these nodes represent people connected to a focal node, then they make up the "Interpersonal Environment," which is a key concept of the "Small World" network theory (Kadushin, 2012, 52).
    - **Small World Network**: A network comprising short distances that link a single node to all other nodes in a given network (Kadushin, 2012, 52).

- **Role/Position**: Applies to how nodes may relate to other nodes within a given network (Kadushin, 2012, 53).

- **Eigenvector**: The scale of which a node is connected to other nodes that are highly central (Kilduff & Tsai, 2003).

- **Centrality (Popularity)**: The centrality of a node is the degree to which a node inhabits a central position within the network. This may include maintaining links to other nodes (degree centrality), the ability to reach other nodes (closeness centrality), connecting nodes that otherwise have no direct links (betweenness centrality), or having links to centrally located nodes (eigenvector centrality) (Kilduff & Tsai, 2003, 130).

- **Prestige/Diversity**: In reference to the prestige or diversity of positions a single node can access (Kadushin, 2012, 257).

- **Brokerage (Broker)**: A node that has links to a variety of other links in other clusters that also has access to ideas/information that are otherwise not present in their own cluster (Kadushin, 2012, 103).

- **Distance (Closeness)**: Distance from one node to another is a requirement of measuring betweenness. Distance is typically a measurement of the shortest (non-redundant) path between nodes (Kadushin, 2012, 304).

- **Path**: A trajectory of nodes and links (edges) contained in a graph (Bockhold & Zweig, 2017).

Common Measures to Describe Networks

- **Size**: The number of nodes within a network (Kadushin, 2012).

- **Component**: Also referred to as component subgraphs, a component is a segment of a network that is disconnected from the rest of the network (Tsvetovat & Kouznetsov, 2011).

- **Connectedness**: A network is said to be connected when each node has the ability to reach any other node directly or indirectly through an intermediary. In essence, there are no isolated nodes (Kilduff & Tsai, 2003).

- **Cohesion**: Two nodes within a clique are said to be structurally cohesive because both are constrained by the composition of their group. Irrespective of linkages to each other, two nodes

that are reachable through other nodes can be pressured to "think and behave in similar ways" (Kilduff & Tsai, 2003, 58).

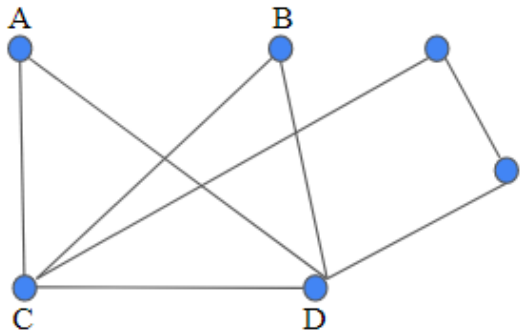

Figure 3.5: Nodes A and B are Structurally Equivalent. Nodes C and D are Structurally Cohesive
Source: Kilduff & Tsai, 2003, 59.

- **Density**: The number of links within a network (opposite of structural hole) (Kadushin, 2012). Density is calculated as the number of links in the network divided by the maximum number of links that are possible in the network (Kilduff & Tsai, 2003).

- **Centralization**: The degree to which a network is centralized over one or several nodes (Kilduff & Tsai, 2003).

- **Transitivity (balanced)**: A type of balance theory that anticipates relations among three nodes to be complete. So, if N1 considers N2 and N3 as friends, then it is expected that N2 will also observe N3 as a friend, therefore creating a transitive triple. Transitivity is "assessed as the number of complete transitive triples divided by the number of triples for which the addition of one missing link would make them complete" (Kilduff & Tsai, 2003, 136).

- **Homophily**: "The tendency of actors to interact with, and share the opinions and behaviours of, other actors similar to themselves on such dimensions as ethnicity, age, educational attainment, gender, etc" (Kilduff & Tsai, 2003, 134). Having common social attributes (Kadushin, 2004).

- **Mutuality/Reciprocity**: Mutuality implies that relations are reciprocal and any asymmetry within a relationship has little to no consequence (Kadushin, 2012).

- **Propinquity**: A subset of homophily that indicates that nodes are inclined to form links to other nodes (people) that are geographically close to each other (Kadushin, 2012).

- **Clustering**: A clustering coefficient measures the quantity of nodes that are linked with each other. In essence, it can be said to measure "mutual trust" between friends. It is important to note that in large networks clustering may be difficult to interpret (Tsvetovat & Kouznetsov, 2011).

- **Clique**: A clique refers to a group of nodes that has two characteristics: (1) All nodes have a direct link with all other nodes in the group, and (2) there is no node outside the group that any of the nodes have a link to (Kilduff & Tsai, 2003).

The definitions above represent common key terms associated with network theory and social network analysis. It is important to note that these terms are not exhaustive, and that there are numerous others pertinent to describing patterns among ties, measures of individual actors, and measures to describe networks. They do not appear here because they are out of the scope of this work, not because they are irrelevant to social network analysis or network theory as a whole. On account of this, the next section will evaluate network theory and social network analysis

3.3. Evaluation of Relevant Theories

It is difficult to ignore the interconnectedness of society and the networks that exist among us as part of our routine everyday life. For example, our social network (physical or digital), the electrical grid, the transportation network, logistics networks, the telephone network, the Internet, and even the World Wide Web. Each represents a network that is intertwined and therefore interdependent on each other. If there is no electricity, there is no fuel. If neither are available, there is no transportation of goods, no Internet, Web, et cetera. So then, why are networks so ubiquitous? Estrada & Knight (2015) have argued that 'being networked' is an essential attribute of complex systems (e.g. society). In consideration of this, it is reasonable to state that as societies grow more complex, the networks that enable them to function will simultaneously become larger and more complex. However, this growth is not limited to individual actors or nodes. The relationships - or links - that define a society are what bind networks together and help us to understand their evolution. For example, network links may represent physical links,[71] physical interactions,[72] ethereal connections,[73] geographic closeness,[74] mass/energy exchange,[75] social connections,[76] conceptual

[71] For example, a road network, the cables that link the Internet, or neural and vascular networks within our bodies (Estrada & Knight, 2015).
[72] For example, an interaction determined by physical force (Estrada & Knight, 2015).
[73] For example, an intangible connection as in the World Wide Web (Estrada & Knight, 2015).
[74] For example, cells connected together in tissue, or connecting countries on a map (Estrada & Knight, 2015).
[75] For example, trade networks, reaction networks, metabolic networks, food webs, et cetera (Estrada & Knight, 2015).
[76] For example, friendship, familial, or collaboration (Estrada & Knight, 2015).

linking,[77] and functional linking[78] (Estrada & Knight, 2015, 2) According to Estrada and Knight (2015), each of these concepts are not completely isolated from each other, and thus warrants the necessity to interpret a network from various points of view. Such various perspectives have been brought forward by several key theorists, which will be discussed below.

The emergence of social network analysis initially emerged from the work of Stanley Milgram who in 1967 conducted an experiment in an attempt to understand how people are connected to each other (Freeman, 2004). During this experiment, Milgram had tasked random people with an exercise that required them to forward a package[79] to one of their acquaintances who they assessed to be able to reach a specific target individual (Milgram, 1967, 60-67). Over the course of his research, Milgram had discovered that most of the participants were linked by six acquaintances, which led to the popular phrase "six degrees of separation" still in use today (Milgram, 1967; Ressler, 2006). Moreover, Milgram's experiment had contributed a new perspective on how to observe interaction and connectedness between people.

A second significant approach in social network analysis is the work of Mark Granovetter's network structures. In 1973, Granovetter claimed that a person's "strong ties" (e.g. family and close friends) were not as significant as their "weak ties" (e.g. acquaintances) when attempting to find employment - which initially brought forward the notion "the strength of weak ties" (Granovetter, 1973). In his subsequent research, Granovetter (1983) further explained that weak ties were much more likely to serve as bridges than strong ties as they connect to social circles that are distinct from their own, thereby acknowledging the value of diversity in relation to greater access to information. According to Granovetter (1983),

---

77 For example, citation networks (Estrada & Knight, 2015).

78 For example, part of a machine activates the function of another machine; one region of the brain functionally connected to another (Estrada & Knight, 2015).

79 The original instructions were as follows: "Each person who volunteered to serve as a starting person was sent a folder containing a document, which served as the main tool of the investigation. Briefly, the document contains: (1) the name of the target person as well as certain information about him. This orients the participants toward a specific individual. (2) A set of rules for reaching the target person Perhaps the most important rule is: 'If you do not know the target person on a personal basis, do not try to contact him directly. Instead, mail this folder … to a personal acquaintance who is more likely than you to know the target person…it must be someone you know on a first-name basis.' This rule sets the document into motion, moving it from one participant to the next, until it is sent to someone who knows the target person. (3) A roster on which each person in the chain writes his name. This tells the person who receives the folder exactly who sent it to him. The roster also has another practical effect; it prevents endless looping of the folder through participants who have already served as links in the chain, because each participant can see exactly what sequence of persons has led up to his own participation" (Milgram, 1967, 64).

> "Acquaintances, as compared to close friends, are more prone to move in different circles than oneself. Those to whom one is closest are likely to have the greatest overlap in contact with those one already knows, so that the information to which they are privy is likely to be much the same as that which one already has" (Granovetter, 1983, 205).

In comparison to their research in 1973:

> "Linkage of micro and macro levels is thus no luxury but of central importance to the development of sociological theory. Such linkage generates paradoxes: weak ties…are here seen as indispensable to individuals' opportunities and to their integration into communities; strong ties, breeding local cohesion, lead to overall fragmentation. Paradoxes are a welcome antidote to theories which explain everything all too neatly" (Granovetter, 1973, 1378).

Within a decade of research, Granovetter determined the significance of diversity within social circles. Furthermore, their work on network structures was foundational to social network analysis as it provided a new perspective on the significance of "weak ties" and the role in which they might have in one's life. In regards to public safety, such a perspective can assist in the analysis of extremist recruitment campaigns and hostage rescue, thereby providing a potential method of mitigation against possible threats that can emanate from such means.

Constructed from concepts of Granovetter's strength of weak ties argument and Milgram's notion of six degrees of separation, Duncan J. Watts established the small world hypothesis. In short, Watts' hypothesis states that while the majority of networks are highly clustered, they are also far-reaching (Watts, 1999). To add to this, in Watts's small world hypothesis, nodes are tightly interconnected, and each one has a weak tie that is able to link to any other node in the network within a few degrees[80] (Watts, 1999). Thus far, this hypothesis has been adopted by numerous research fields in an attempt to understand the organization and functioning of several networks including the spread of information, the spread of AIDS, and even the collapse of financial markets (Buchanan, 2002). Furthermore, Watts' small world hypothesis allows one to observe a network through an asymmetric approach as opposed to a traditional hierarchical approach which can prove useful when analysing asymmetric threats. Ultimately, this enables an analyst to acutely consider a networked approach,

[80] For example, someone's network of friends is usually tightly connected, sharing common friends, backgrounds, and possibly overlaps. But, despite a "clustered" core (think Milgram's six degrees of separation), one node - or a person - can still reach a complete stranger through a small number of connections (Ressler, 2006).

which can moreover provide critical information on significant nodes (hubs) within a network, and how that network functions - all of which can be used to inadvertently fracture an organization.

In consideration to the theories above, social network analysis has immense applicability to the intelligence community tasked with public safety. According to Wheaton and Richey (2014), analysts have used social network analysis to identify key people within an organization,[81] identify strategic agent networks, and associate new agents and information flows throughout networks. They further argue that social network analysis can also be used in combination with other analytical practices including but not limited to Geographic Information Systems and Intelligence Preparation of the Battlefield[82] as a means to produce reliable predictive analysis (Wheaton & Richey, 2014). However, social network analysis does have its limitations, especially as it applies to data overload.

As previously implied, the Web is very large with significant amounts of data flowing through it. So, if one were to input troves of data from multiple social networking sites it would inevitably be rather difficult to find persons of interest if one does not already have a start point. Essentially, there would be too much "white noise" to sift through and the effort would not be very efficient (Wheaton & Richey, 2014). As a means to overcome this however, analysts should separate their datasets into research specific investigations. This should not limit an analyst's ability to search other data sets for relationships and similarities, unless the software they are using does not permit them to search concurrently while already in an investigation. In this case, the solution would simply be switching or upgrading software. Moreover, using social network analysis as a "catch all" will likely not result in any meaningful analysis. Instead, looking at particular datasets for specific information and removing (filtering out) unnecessary entry points is what will take an analyst from an overwhelming large network graph to one that is organized and meaningful.

An additional disadvantage with social network analysis includes the possible introduction of bias (European Commission, 2014). At this time, bias is understood to be a disadvantage because the analytical output is highly dependent on which data (nodes and links) are included within the analysis

[81] For example, by determining a node's degree centrality.

[82] Intelligence Preparation of the Battlefield (IPB) is a means of defining the operational environment. "IPB is a systematic and continuous process by which enemy capabilities, vulnerabilities and probable courses of action (COAs) in a specific geographic area are determined" (Department of National Defence, 2001).

(European Commission, 2014). This can be a result of an analyst's ability to obtain the necessary data sets, and a subjective interpretation of the research question. A means of mitigating this however can be sought through careful framing and consideration of the research question beforehand (e.g. group discussion, brainstorming), peer review, and including a caveat to analysis wherever an analyst may perceive missing or absent data.[83]

3.4. Relevance to Research

Network theory, more specifically social network analysis will largely be used as a basis for exploring and interpreting how meaning can be generated to inform analysis. On these grounds, it may also be used to critique and assess the context in which this theory is most useful, to determine if the theory has greater significance if used in a unique way, or, if social network analysis can be used by analysts to develop hypotheses for their research and overall assessments. It is important to note here that it is not the intent to conduct complex mathematical experiments with network theory or social network analysis, but instead to consider how the concepts of the theories can be used to explore analysis as it relates to public safety through open source intelligence. In essence, using networks as information maps.

In relation to public safety, analysts may be confronted with datasets that will require them to discern patterns, connections, and relationships in order to assist with answering questions to the intelligence problem. These questions may range from the size and structure of a group (e.g. centralized, decentralized, hierarchical), its evolution (e.g. path to leadership to identify possible replacement leaders), its pattern of activity (e.g. meetings on a certain day/time), and so on. Through further analysis using a network lens, an analyst can obtain additional information from datasets including:

- If the network forms a connected structure (e.g. how many nodes interact with each other);
- If the nodes represent a Gaussian-like distribution[84] or if they are skewed with some nodes having a larger number of links, while others have few;

[83] For example, if an analyst is only able to obtain datasets for a specific timeframe, location or other variable, then it is important to acknowledge how missing datasets make skew the interpretation. Additionally, throughout their analytical process an analyst must consider: known knowns, known unknowns, unknown unknowns, and unknown knowns.

[84] I.e. a Bell Curve.

- If there is a central node (hub) within the network, and the percentage rate of interaction that includes this node;
- If there is a node that is critical to the passage of information to the rest of the network;
    - Is the node a critical part of the network?
    - How will the network exist or evolve without this node?
    - How resilient is the network?
- The most and least steps (degrees) that one node is separated from others;
- The average degree between nodes;
- If the network is transitive (e.g. If N1 passes information to N2 and N3, is there a high chance (>33%) that N2 also passes information to N3?);
- If and where there are any potential bottlenecks in communication;
- If there is a pattern among the nodes (e.g. are the links random or do they follow or form a specific pattern?);
- If the nodes with a higher number of links usually interact between themselves, or if they prefer interaction with nodes that have very few links?;
    - Are the nodes disassortative[85]?
- If there is a central core of nodes that dominate the network;
- If the network can be divided into two (or more) disjointed groups of nodes (Partially adapted from Estrada & Knight, 2015).

In essence, each of these questions have the potential to piece together fragments of a "story" that may not otherwise be apparent if not considered through a network lens. In addition to this, using a visualization of the network can also provide the opportunity to manipulate it as a means to test where its critical nodes are located, and therefore where its strengths and vulnerabilities exist. The relevance of analysing a network can be paramount to understanding and defeating some of the most dangerous networks that threaten public safety. Furthermore, maintaining a network lens to map information provides an alternate perspective to research and investigations.

Data is a network. Combining social interaction, one's online presence, and publicly available information can have profound implications to finding and/or understanding a person or location of interest. For example, if someone has 100 "friends" on a social networking website - and none of them are a friend of the other - then two steps removed from this, that person has access to 10,000 people (100x100), and three steps removed, that number increases to 1,000,000 (100x100x100) (Kadushin, 2012, 21-22). This concept of data as a network can also be taken a step further for investigative purposes. For example, if one has a Facebook account, then, it is likely that they also

[85] A node is disassortative when links between hubs are suppressed, and assortative when links between hubs are enhanced (Hao & Li, 2011, 1). In this manner, nodes with a lower degree are more likely to be linked to a node with a higher degree.

have a presence on Instagram or other social applications such as LinkedIn, TikTok, SnapChat, and / or Twitter. If an individual has a presence on social networking sites, it is important to look at who they are connecting with (friends, family, colleagues, influencers), and if there are multiple profiles present. In turn, this also means that the person has at least one email address, and a smart device to connect them to the Internet. If they have access to the Internet, they may also be using it for things outside of social networking such as banking, school, work research, et cetera. Observing this type of activity alone can be critical to locating a missing person or profiling / locating a person of interest. Moreover, this research will acknowledge that networks are present everywhere, and that society - digital or physical - is in itself a complex social network.

Overall, the relevance of network theory and social network analysis to this research is the perspective they provide to understanding the complexity and interconnectedness of society and information. The theories provided herein are not offered as a representation of network theory or social network analysis as a whole, but instead are intended to provide a basic understanding of the theories in order to show their utilisation as part of an analyst's thought process in their investigative research and analysis. Moreover, it is also a means to understand networks as information maps. This in turn is key to evolving an analyst's practice and the implementation of alternative research techniques in the intelligence community. In their book, *Linked*, Barabasi (2003) asserts that the Internet is often observed as human in its creation - "akin to an organism or an ecosystem demonstrating the power of the basic laws that govern all networks" (Barabasi, 2003, 8). In accordance with this statement, finding relevance within network theory and social network analysis is a matter of dissolving a reductionist mindset, and exploring data with an open mind.

Chapter 4: Methodology

This chapter will discuss the methodology that produced the fifth chapter of this work. The focus point of the problem is the research question initially stated in Chapter 1, but included here for ease of access: *In what regards is OSINT an effective tool for emergency management, especially in relation to public safety?* In regards to the results, this chapter will discuss the methodological approach, the method of data collection, the method of analysis, as well as the evaluation and justification of methodological choices in answering the research question.

4.1. Methodological Approach

The overall approach to the research was qualitative in nature, focusing on case studies, tool exploration, and access to information and privacy (ATIP) requests. These approaches assisted in answering the research question through granting the flexibility of the utilisation of free OSINT tools, publicly available information, and government records. The aim of such a passive collection methodological approach was to explore an under-researched topic, and show the relative ease with which an analyst can freely obtain data and information that can assist in their OSINT investigations. This aim was achieved through the use of primary and secondary sources of information, and through following a widely accepted methodological approach to OSINT case workflow (see Figure 4.1). Where some case studies required the manipulation of variables (e.g. temporal analysis of maps by month and year, manipulation of a JPEG to verify authenticity), others utilised descriptive data and observations (e.g. location descriptions, images) as a means to identify a location, verify an image, or maintain situational awareness.

The rationale behind this methodology was that it provided an effective means of collecting publicly available information in an ethical manner with minimal risk to personal security, the rights of others, or violation of legislation. These were of significance to consider as using publicly available information for open source investigations can potentially result in collateral damage to an individual or an organization if not conducted properly. Further to this, the assumptions that were central to undertaking this methodology included the availability of publicly available data and information that

would assist with the interpretation and verification of the availability of historical map data, locations, and images.

At the time of writing, this methodological approach was deemed suitable to answer the research question because it provided the most efficient and effective means of gathering the necessary information. This is a direct result of the vast amount of data and information that is available on the World Wide Web, and is amplified by the amount that will be created today, tomorrow, and thereafter. However, this is not to say that all gathered data and information did not go through a verification process to ensure validity and reliability. The processes used to prove validity and reliability have been documented within each individual case study, and where appropriate (e.g. Case Studies 5.3.2 and 5.3.3) were recorded utilising Hunchly, a tool that tracks and timestamps URLs during an online investigation.

Further to the above, ethical considerations were taken into account before and after each case study was complete. Although the intent was to ensure no identifying information was included, Case Study 5.3.2. does include images of hostages prior to their execution. While no names are included, there is a marginal risk that these images may be upsetting to the respective family of the hostages. However, given the availability of these images on the Web and the lapse in time since the incident, the assessed risk was determined to be low. In the same case study, a second ethical consideration occurred during the investigation and had to be managed appropriately. This occurred when the location was discovered to be a former U.S. military base. Given that the revelation of such information may result in matters pertaining to operational or personnel security, the name of the base has been redacted from the case study. Despite all information retrieved being publicly available, the United States military has not publically released the name of the base alongside its location online, and therefore neither will this case study. Although replacing this case study was a viable option, following through with its inclusion in the results allowed for a visual representation of managing ethical decisions within OSINT investigations, which moreover is valuable in and of itself.

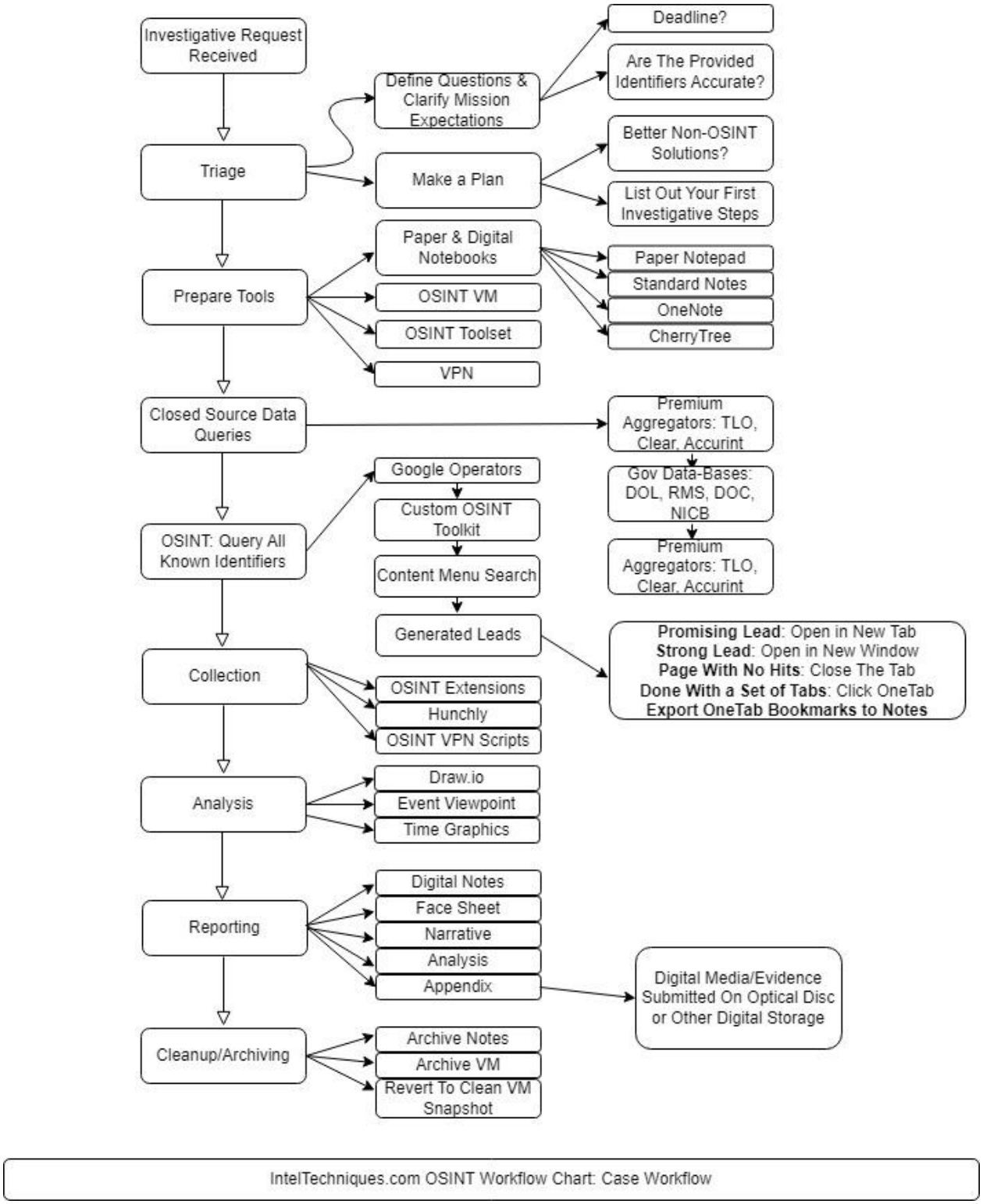

Figure 4.1: OSINT Workflow Chart: Case Workflow
Source: Bazzell, 2021, 606.

4.2. Methods of Data Collection

Each of the case studies relied on existing publicly available data that resided on the Web. The means of collection however, varied between each case study. As such, the method of data collection will be explained for each case study individually.

4.2.1. Data Collection: Case Study 5.2

Case Study 5.2 largely dealt with determining how OSINT could be used for horizon scanning and maintaining situational awareness of current and emerging threats related to public safety. Given this, data collection did not occur in the traditional sense, and as such, the data collected was not included in the case studies. The methodology to determine how OSINT could be used for horizon scanning and situational awareness was determined by evaluating several open source (free) news aggregators such as Feedly, Inoreader, Google News, Google Alerts, and Talkwalker. However, the case study only includes insights on Google News, Google Alerts, and Talkwalker. This is in consequence to the free versions of Feedly and Inoreader having limited feeds, limited alert capabilities (e.g. push notifications), and limited search parameters, among other deficiencies that classified them as inadequate for conducting horizon scanning and maintaining situational awareness without incurring a cost.

The search parameters entered for each aggregator were similar in that they included one or more keywords (e.g. Proud Boys; Canada + Terrorism; ISIL | Islamic State | ISIS | DAESH), and were delivered "as-it-happens." Common search parameters between the aggregators included the following:

- Canada + Terrorism
- US + Terrorism
- Europe + Terrorism
- Russia + Ukraine
- Wagner Group
- Haiti
- China
- Al Shabaab + Somalia + Kenya
- Burkina Faso
- Mali + Terrorism
- Sudan
- ISIL | Islamic State | ISIS | DAESH
- Afghanistan + Taliban
- Proud Boys
- Worm | Malware | Virus | Phishing | Ransomware

Google Alerts Search Parameters:

- How often: As-it-happens
- Sources: Automatic (News, Blogs, Web, Video, Books, Discussions, Finance)
- Region: Any Region
- How many: Only the best results
- Deliver to: email address

Talkwalker Search Parameters:

- Result Type: Everything (News, Twitter, Blogs, Discussions)
- Language: All languages
- How often: As-it-happens
- How many: Only the best results
- Delivery: Email

Search parameters within Google News differ slightly from Google Alerts and Talkwalker as Google News is not designed to push alerts to a user's email address, but instead allows the user to follow topics, save searches, and save stories. In consideration to this, the search parameters were similar, but not identical to the alerting aggregators. The Google News search parameters were as follows:

Google News Followed Topics

- Proud Boys
- Haiti
- 2022 Haiti Crisis
- Afghanistan
- Taliban
- Middle East
- Islamic State of Iraq and the Levant
- Al Shabaab
- Somalia
- Kenya
- Ethiopia
- Jama'at Nasr al-Islam wal Muslimin
- Burkina Faso
- Niger
- Mali
- Sahel
- Africa
- Russo-Ukraine War
- Russian Invasion of Ukraine
- Wagner Group
- Ukraine
- Europe
- China
- US
- Canada
- Computer Security

Google News Saved Searches

- Canada + Terrorism
- US + Terrorism
- Europe + Terrorism
- Afghanistan + Taliban
- Mali + Terrorism
- Al Shabaab + Somalia + Kenya
- Horn of Africa

Once the search parameters were determined, the output of the aggregators was assessed for their ability to provide quality information and alerts from credible sources in a timely manner. The outputs were assessed by their ability to answer several questions, including the basic "who, what, where, when, why, and how" type questions for a given situation. Additionally, the aggregators were

evaluated by their ability to provide enough information from multiple sources that an analyst could use to develop an assessment on a current threat, or whether or not an emerging threat was developing. Overall, three out of five aggregators met this threshold, and as such, Google News, Google Alerts, and Talkwalker were presented in Case Study 5.2 within the results chapter of this work.

4.2.2. Data Collection: Case Study 5.3

Case Study 5.3 is unique in that it contains three case studies within itself. These three case studies served to identify open source tools and techniques that could be used to find a person, group, or location. Case Study 5.3.1 largely served to identify tools and where one can find these tools. Case Study 5.3.2 located the site where DAESH hostages were executed. Case Study 5.3.3 located a moving convoy with "dashboard camera" footage.[86]

As previously stated, Case Study 5.3.1 mainly served to identify open source tools. Generally speaking, the method to find these tools would usually be accomplished by conducting a simple Google search for open source tools. However, these tools were discovered early on in the endeavor of this work, and were simply noted to be included for use in a case study. Overall, they were included because these tools are useful for new and experienced OSINT analysts, and are updated regularly.

In regards to Case Study 5.3.2, this case study was the most complex of the three because it involved the use of multiple skill sets (e.g. Boolean searches, mapping software, reverse image searches) and situational awareness (e.g. desert terrain, war and its effects on terrain). To begin, this image was selected because it provided enough foreground and background available to use in its location analysis. For example, it contained a pool with a distinctive shape, decorative bricks, shrubbery that is identifiable to a position on satellite imagery, a building with a distinctive roof, and a fence providing privacy around the pool. Additionally, it also presented a significant challenge; the

[86] Case Study 5.3.3 was originally a "capture the flag" (CTF) scenario from an OSINT CTF website, https://milosintctf.com/challenges. Although the footage does not originate from a true dashboard camera, per say, the original image was retrieved from Google Maps, and still served its originally intended purpose.

macro location of Mosul, Iraq, has a substantial amount of damage from ongoing conflicts in the region. Given the nature of the investigation, much of the methodological approach is also revealed within the results section as a means to show how one could reach a location from an image as a start point. However, for the purpose of clarity, it will be briefly discussed here.

To initiate the investigation, a visual analysis on the image was conducted in order to list any observed objects that could be used to confirm the location of the image later on. After this, a reverse image search of the original image was conducted, but did not return any information of significant value. From here, a search for "ISIS Pool Iraq" returned reporting that corroborated the execution. While reading through the reports, it was noted that an Iraqi news agency listed al-Faisaliya as the location of the swimming pool. A search for this area was conducted, and it was discovered that al-Faisaliya was in fact within the boundaries of the area of interest (Mosul, Iraq). From here, a map study of the area was conducted that proved unfruitful, and so a search on Google was conducted for locations that may have outdoor pools (e.g. palaces, hotels). Unexpectedly, this search resulted in images that were pivotal to the investigation because it had returned images of American soldiers taking part in a New Year's celebration in 2006 at FOB[87] XXXX. This image was compared to the original image, and was revealed to have multiple similarities, and provided additional visual perspectives that the original image in the investigation did not provide because of the angle from which it was taken. These similarities were noted visually in the results section. From here, it was determined that both locations were the same.

Since one of the websites that the U.S. images originated from made reference to a former palace of Saddam Hussein located north-west (NW) of Mosul, it made sense to pivot the investigation to that area. So, the next step was to conduct a detailed search and map study of palaces within the area of NW Mosul. While at first the searches revealed palaces in the area, no pools matching the description were visible. After some time, the map study was moved to Google Earth Pro where an additional map study was conducted using imagery from 2015. Studying the same areas from Google Maps, this search proved to be extremely beneficial as the uniquely shaped pool was found. When

[87] Forward Operating Base.

compared to the original image in the investigation, the listed observed images from the initial steps in the investigation were located, compared, and matched between the 2015 Google Earth Pro imagery, and the original image. Furthermore, an additional historical analysis on Google Earth Pro was conducted, which revealed that the pool was filled in 2017, which is why it was unobservable on Google Maps.

The task of Case Study 5.3.3 was to find - to the nearest 10 km - the location of a military convoy approaching Ukraine from Russia, assessed to be travelling to the Aviation Museum in Luhansk. Similar to Case Study 5.3.2, the investigation was initiated by making a list of visible objects, which in this case included traffic signs, tree lines, a T-shaped intersection, unique markings on the road, electrical distribution lines, and a culvert. After noting these objects, the first step was to determine the rough location of the image. This was done by translating the blue placard from Cyrillic to English using a Cyrillic/English Alphabet. After translation it was determined that черников translates to Chernikov. This revealed that the image was taken in the vicinity of Chernikov, Russia. A search on Google Maps produced a location in Russia near the Ukraine border confirmed as Chernikov, Rostov Oblast, Russia. From here, a search for the driving route between this location and the Aviation Museum in Luhansk, Ukraine was conducted. Utilising the Russian portion of this route, a visual search for a T- intersection located 2 km from Chernikov was completed, revealing one major intersection between Chernikov and the E-50 highway. This location was then assessed in Street View where the distinctions between the original photo were compared to the view of the T-intersection on the E-50 highway. The locations appeared to be an exact match, and a calculation was made between the location of the convoy to the Museum, which was found to be 100 km after rounding to the nearest tenth of a kilometer.

4.2.3. Data Collection: Case Study 5.4

The aim of Case Study 5.4 was to identify methods of verification that can be used when conducting open source intelligence. In consideration to this, the case study brought forward a critical thinking criteria developed from the author's experience, and also provided several tools to verify and evaluate information. These included tools on the matters of ethical guidelines for evaluation, fallacies

& biases, information verification, image verification, deepfake verification, and disinformation toolsets. These tools have been collected over the course of this work, originating from multiple websites and recommendations.

Subsequent to this case study, a tool exploration of Forensically was included as a means to showcase the process of verifying the authenticity of an image. Similar to the previous case studies, much of the method of verification was presented in the results section in order to facilitate an understanding of how the tool functioned. This particular example utilised the original stock photo on Forensically. The reason for this is that it allowed for the easiest method of replication while diminishing the possibility of inconsistencies (e.g. resulting compression or flattening of the JPEG from multiple uploads/downloads) from occurring in the reproduction of this case study. Nonetheless, the data was collected through adjusting the settings within the Forensically toolset, which moreover showed how and when the photo was manipulated.

4.2.4. Data Collection: Case Study 5.5

Case Study 5.5 sought to explain how OSINT and Social Network Analysis can potentially work together to serve the needs of an analyst in an open source intelligence investigation. As a result of possible ethical risks, the case study did make use of any real social networking data, but instead presented faux network data that represented the communication structure of a small insurgent cell. Since the intent of this case study was to identify how social network analysis may be applied in an OSINT investigation, using a faux dataset did not impair the aim of the case study. Instead, it situated a simple explanation of how an analyst can use social network analysis to identify and exploit vulnerabilities.

4.2.5. Data Collection: Case Study 5.6

Case Study 5.6 is distinct from the others in that it involved access to information and privacy (ATIP) requests. On 9 and 10 November 2022, ATIPs were requested from the Department of National Defence (DND), Transport Canada (TC), Financial Transactions and Reports Analysis Centre (FINTRAC), Global Affairs Canada (GAC), Canadian Border Services Agency (CBSA),

Royal Canadian Mounted Police (RCMP), Canadian Security Intelligence Service (CSIS), and Public Safety Canada (PSC).[88] The original request sought records within the timeframe of 2017-2022 pertaining to training, policies, and risks regarding open source intelligence. To date, responses from DND, FINTRAC, Transport Canada, Public Safety Canada, and the Canadian Security Intelligence Service (CSIS) have been received, while the remaining agencies have asked for additional time to gather the requested documents. Where documentation has not been received from the respected departments in a timely manner, reference to the Acts that legislate their powers, duties, and functions has been made. These Acts were accessed and collected from the Justice Laws Website.

4.3. Methods of Analysis

The method of analysis that was used to interpret the results included textual, content, and thematic analysis. While textual analysis was largely used in Case Study 5.6 for understanding the agencies responses, and interpreting legislation, it was also used in conjunction with content analysis during the interpretation of results of the remaining case studies. Thematic analysis was especially important in identifying patterns within images, but was also used in conjunction with content and textual analysis during the process of investigations, especially when assessing locations.

In specific regards to textual analysis, this method was used to interpret and record the data and information from images within the case studies. This type of analysis then allowed for multiple questions to arise and patterns to be discovered. For example, textual analysis was used to relate the information provided on U.S. military websites to information provided by Middle Eastern news agencies in order to assess the location of the DAESH execution site. Additionally, it was also used to determine the most viable tools for assessing emerging threats and maintaining situational awareness, as well as identifying methods for information verification. Moreover, as mentioned above, it was also an effective means of approaching and interpreting legal documentation from Case Study 5.6.

Similar to textual analysis, thematic analysis was very useful in processing images within Case Studies 5.3.2, 5.3.3, and 5.4. By using this approach in analysis, themes and patterns were

[88] These agencies were selected based on the list of participating departments that were accepting ATIP requests at the time. This list can be found at https://atip-aiprp.apps.gc.ca/atip/welcome.do

quickly identified, which furthermore complemented information gained from a textual analytical perspective. Moreover, the combination of these two approaches assisted in the corroboration of information, which was significant to the overall verification of information.

Finally, a contextual analytical approach proved useful in each case study as each one presented its own image, language, or problem that had to be categorized and approached in a specific manner in order to evoke meaning. For example, Case Study 5.3.2 required a contextual approach to time in order to identify an execution site. Case Study 5.3.3 required a linguistic analysis to geolocate a convoy, and Case Study 5.4 required a critical approach to data and information in order to identify misinformation, disinformation, fake news, and manipulated images. Case Study 5.5 required a contextual approach in order to analyse the flow of communication with the removal of specific hubs, and the additional network effects such a removal may cause. Moreover, although used in different contexts within each case study, content analysis served to reveal information that was otherwise obscured.

To conclude, the combination of these three analytical approaches were beneficial to the overall analysis of the results as they each provided a unique perspective in the approach of the case studies. Further to this, the methods of analysis provided a means to make assessments where appropriate, which moreover sufficiently completed each case study.

4.4. Evaluation and Justification of Methodological Choices

The methods discussed above were selected and evaluated on a number of criteria. Primarily, they had to be free (or incur a minimal charge[89]) and ethical. This was important to consider because oftentimes financial restrictions limit the exploration and/or expansion of a capability or interest, and unethical collection can result in inadmissible evidence and misinformation. Secondly, they had to be functional with limited OSINT background. Similar to financial restrictions, the exploration and/or expansion of a capability can be met with discontent and forestalling based on assumptions on technical complexity and time associated with training. Finally, they had to be effective in producing

[89] For example, the cost of an ATIP is $5, and Hunchly is $129/year, but can be trialed free of charge for 30 days.

actionable results,[90] or provide additional pivot points[91] that can either be used in an investigation, or by a different intelligence discipline.[92] Together, meeting these criteria assisted in addressing the objectives, and answering the research question of this work.

In addition to this, the OSINT Workflow Chart (Bazzell, 2021, 606) in Figure 4.1 was employed as an investigative workflow guide as a means to provide structure to the investigation. In consideration to Bazzell's OSINT experience and publications, they are one of the more prominent published practitioners in the field. As such, his methodological workflow is widely accepted as an ideal start point for OSINT investigations. Of course, the workflow chart is not a step-by-step guide to finding answers on the Web. Instead, it is designed more to be a fluid workflow to initiate planning, collection, and analysis. As such, some aspects of the chart may not be applicable to all investigations, but kneading through the workflow can initiate a thought process that is conducive to the overall investigation. For these reasons, the OSINT Workflow Chart was evaluated as a valuable asset to the methodological workflow of the case studies.

As a matter of course, there are multiple means and methods that can be used to conduct an OSINT investigation. In consideration to this, it is not to say that other tools and techniques were not suitable for this research, but rather that the methods selected were the most efficient and effective approach to addressing the research question and objectives. More specifically, the approaches taken have resulted in the acquisition of publicly available information - or in some cases, information that was made public (e.g. ATIPs) - that shows how open source information can be used to produce intelligence in matters relevant to public safety. Moreover, this has contributed to a new perspective on the use of open source intelligence for public safety, especially in regards to training, techniques, ethics, and policy.

Acknowledging the indicated, it is important to recognize the obstacles and limitations that were encountered in this approach. Most impactful to the results section was the inability of a number of government agencies to respond to the Access to Information and Privacy request in a timely

[90] For example, evidence that is relevant to positive identification or capture.
[91] For example, an additional person of interest.
[92] For example, HUMINT.

manner. Once submitted, the responding agencies had 30 days to provide the requested records, or provide a reason as to why an answer could not be produced in that given time period. At the time of writing, three of eight agencies had responded in a window spanning 30-90 days, while the remaining others had requested additional 90 or 180 day extensions, with the possibility of further extensions. While this impacted the ability to obtain the requested documents, an approach to solution was made by locating the Acts which governed the direction of the respectful agencies on the Government of Canada, Justice Laws website. Although this limited the specific information that could have been obtained, the Acts did provide legal grounds by which the particular government departments must operate within, which moreover provided sufficient information for analysis.

Perhaps common to all investigations, all required data and information are seldom located in a single place. On these grounds, minor obstacles were encountered in locating information (e.g. Case Study 5.3.2), but they were overcome by utilising alternative tools, such as Google Earth Pro. While it must be acknowledged that temporal map analysis would be available in several map databases (e.g. ArcGIS Pro, Esri), this would have incurred a cost, which was not in line with the criteria of the methods discussed at the beginning of this section. Moreover, while this was a technical obstacle to the investigation, it was minimally limiting and easily overcome.

Chapter 5: Results and Discussion

5.1. Introduction

The results in this chapter are based on case studies outlined in the methodology section, and included in each appropriate section herein. In consideration to the qualitative nature of this research, this chapter has combined both the results and discussion into one chapter. Although discussion will occur throughout the sections, the greatest detail is provided in section 5.7.

As a result of the unique nature of open source investigations, some of the methodologies described in the previous section will be described herein with visualization where appropriate. This process will allow for a more thorough understanding of the results and discussion.

5.2. Situational Awareness and Horizon Scanning

Case Study: Situational Awareness and Horizon Scanning[93]
Determine how OSINT can be used for horizon scanning and maintaining situational awareness of current and emerging threats relating to public safety.

Maintaining situational awareness and knowledge of emerging threats are both key to the effectiveness of emergency management and public safety. The critical elements to these functions are the availability of timely, reliable, and actionable intelligence of areas, situations, and / or persons of interest. In this regard, OSINT can complement current capabilities situated to confront the intelligence requirements of a current or emerging event. This is because OSINT allows public agencies to covertly surveil persons, areas, or situations of interest on social media, forums, chat groups, et cetera. The ability to conduct such actions has immense benefits for enhancing public safety through combating organized crime and terrorism. However, it is important to include here that covert techniques - whether active or passive in nature - must be appropriately deployed by trained individuals, and auditable in accordance with current policies and legislation. Otherwise, any evidence discovered may not only be at risk of being inadmissible in court,[94] but may also be an infringement

[93] Here, horizon scanning refers to the early signs of potential emerging threats.
[94] Or to the legal team of targeting committees on deployed operations.

on Charter and / or privacy rights - each of which could jeopardize an operation, put lives at risk, and impact the ability of an agency to employ such capabilities in the future.

In consideration to the case study question, OSINT can be very effective in providing and maintaining situational awareness and assessing emerging threats to public safety for both domestic and international issues and incidents. This is largely a result of the nature of the Internet and the World Wide Web being comparable with networks of information streams. These streams may be in the form of social networking websites, blogs, news articles, journals, chat forums, image analysis websites, et cetera. Pulling data and information from each of these sources enables a researcher to obtain vast amounts of resources, and to pivot to other sources within the network[95] that they can use toward their overall assessment of a situation. While there are a number of tools and means of maintaining situational awareness and assessing emerging threats to public safety, this case study will examine those which are free at the time of writing.

Perhaps most accessible to OSINT researchers are Google Alerts[96] and Google News.[97] Although both are distinct on their own, they each allow one to search for topics of interest using Boolean search operators (discussed in section 5.3.1.), which assists in filtering out irrelevant information. However, Google News will only allow you to "follow" a certain news stream, and will save any searches using Boolean search operators as a "saved search" instead of a stream to follow (Google News, 2023). One of the key differences between the two is that Google Alerts provides one with the ability to choose how often they get alerts[98] the types of sources,[99] the region,[100] the number of results,[101] and the delivery preference,[102] that is simply not present in Google News (Google Alerts, 2023). While this is not meant to discount the effectiveness of Google News, it is more so that each has different benefits and uses for particular cases.

---

95 For example, finding the original post, excluding circular reporting, determining the relationship and commonalities between pieces of data or information.

96 Available at https://www.google.com/alerts

97 Available at https://news.google.com/

98 Once a day, once a week, or as-it-happens.

99 News, Blogs, Web, Videos, Books, Discussions, Finance.

100 Allows you to pick which country produces the content.

101 "Only the best results" or "All results".

102 Via a Google email or RSS feed.

Similar to Google Alerts, is Talkwalker Alerts. However, unlike Google Alerts, Talkwalker also covers social media, and its searches are more customisable (Talkwalker, 2023a). In addition to its alerts function, Talkwalker also has a Social Search tool that allows one to monitor hashtags, mentions, and keywords (Talkwalker, 2023b). The benefits to this is that an OSINT researcher can utilise tools such as these to be alerted to domestic or international incidents that may pose risks to public safety for Canadians. For example, alerts about civil unrest in a particular country can alert specific organizations to plan for the possible evacuation of citizens from that country; an escalation in force can alert government agencies to plan for a possible deployment of military personnel; or, an imminent or sudden natural disaster can alert aid agencies to prepare for humanitarian assistance. Tools such as those mentioned can assist agencies to be proactive rather than reactive in emergency situations when time is of the essence.

Similar to how these tools can be used for situational awareness, they can also be a means of horizon scanning to assess the emergence of threats. For example, alerts can be created to provide information on escalating violence in a region; monitor activity associated with a pandemic (e.g. casualties); or monitoring environmental hazards that can impact the supply of food, water, and shelter. Each of these can result in various types of emergency situations over a period of time. Therefore, having a method to examine these types of threats and risks not only supports decision making, but is fundamental to mitigation, prevention, preparedness, response, and recovery.[103]

Overall, OSINT can be used to maintain situational awareness and conduct horizon scanning through the use of tools that notify the researcher to alerts of particular interest. Additionally, a researcher can also utilise the "follow" function on news streams to complement their analysis of ongoing or emerging situations that may present as future threats. From here, an analyst or collator can parse through the alerts, and disseminate a daily / weekly intelligence update or send out intelligence reports or summaries on matters of greater concern.

While this section made reference to three particular tools, there are a number of others with similar capabilities. It is important to note that each analyst will have their own preference, pending

[103] Mitigation, Prevention, Preparedness, Response, and Recovery are five principles of emergency management

which functions serve them best. Although the aforementioned tools are free at the time of writing, it is important to not discount other paid tools, as in some cases they do have very powerful search parameters that may be more in line with what the analyst may require.

5.3. Finding a Person, Group, or Location

Case Study: Tools and Techniques
Identify tools and techniques that can be used to find a person, group, or location.

Finding a person, group, or location using open source intelligence techniques can present a number of complexities and complications. While some searches proved to be complex as a result of multiple elements that had to be considered within the investigation, others showed higher levels of complication as a result of a limited number of elements or evidence available for examination. However, the data and information collected were consistent enough to formulate intelligence analysis.

In addition to the above, multiple ethical considerations were taken into account for each of the case studies. As mentioned in previous sections, open source intelligence - similar to other research and investigative methods - can be invasive, and can therefore expose information that is not meant to be made public through collection, analysis and correlation.

5.3.1. Person, Group, or Location

Finding a person, group, or location online can at times be more difficult than a simple search for a person's name or a search function within a mapping application. This is especially true for attempts to find missing persons, nefarious actors, and small towns or villages. In such cases, advanced open source tools and techniques are required in order to locate the most relevant information. This section will therefore present various tools and techniques that have been utilised by a number of OSINT analysts. It is pertinent to note however, that preference for one or the other is fairly subjective and depends on the researcher or analyst's experience and overall end state.

Tools within the OSINT community generally fall under two categories: paid subscription or open source (free). While there are a number of subscription-based tools that are well established

within the OSINT community, the findings will largely focus on those that are freely accessible. This is largely because access to tools within the OSINT community may vary depending on access to resources (e.g. money, policy, equipment), and credentials (e.g. security, intelligence, policing, and other professional communities[104]).

First, there is Michael Bazzell's IntelTechniques tool index (Bazzell, 2023), which allows one to conduct an investigation using any one of the following resources:

● Search Engines	● Addresses	● Business & Government
● Facebook	● Telephone Numbers	● Vehicles
● Twitter	● Maps	● Virtual Currencies
● Instagram	● Documents	● Breaches & Leaks
● LinkedIn	● Pastes	● Live Stream Audio
● Communities	● Images	● Live Stream Video
● Email Addresses	● Video	● APIs
● Usernames	● Domains	
● Names	● IP Addresses	

For example, utilising Bazzell's Facebook tool, one can expedite the process of finding posts, photos, and videos posted by a specific User ID. Additionally, one can use a general search term and isolate a search to a variety of categories such as posts, people, photos, videos, marketplace, et cetera, or input the Username to obtain a list of friends, photos, reels, or check-ins (among others). Together, these search parameters can provide valuable information to the researcher, especially as they relate to the pattern of activity of the person(s) of interest. Similar search parameters can be found within the Twitter, Instagram, and LinkedIn search tool, while other tools provide more niche search parameters relating to Domain searches, IP Addresses, and Breaches & Leaks.

[104] For example, some vendors may only sell their software to government agencies or private sector companies that have an interest in security and intelligence.

Secondly, OSINT Combine's list of free tools (OSINT Combine, 2023) are updated regularly and include the following:

- Alt-Tech Social Search
- TikTok Quick Search
- Social Geo Lens
- Instagram Explorer
- World Social Media Platforms
- Snapchat Multi-Viewer
- Google Analytics ID Explorer
- OSINT Collection Schema
- Reverse Image Analyser
- WhatsMyName Username Tool
- Data Visualization Tool
- Reddit Post Analyser
- Dark Web Bookmark Stack
- OSINT Bookmark Stack

While OSINT Combine's tools share some similarity to Bazzell's, they are also unique in that they offer an Alt-Tech Social Search parameter and a Social Geo Lens. The Alt-Tech Social Search tool is useful as a number of alternative social media platforms and paste sites have recently gained popularity. Parler, Gab, Minds, BitChute, DLive.tv, Rumble, JustPaste.it, WrongThink.net, and 8kun are just a few examples that OSINT Combine's Alt-Tech Social Search tool will search across (OSINT Combine, 2023). This is crucial to an OSINT investigation as a number of malicious actors (e.g. far right extremists) have taken favour to Alt-Tech platforms as a result of being banned from prominent ones (e.g. Facebook, Twitter, et cetera), in addition to the security that some of these sites offer, and the lack of general awareness of their existence.

Separately, OSINT Combine's Social Geo Lens is also quite unique in that it allows the researcher to conduct a geo-search of posts made on Facebook, Instagram, Twitter, Snapchat, or YouTube. This can be useful to an OSINT investigation taking place over a particular area where social media users have taken to social platforms to report on the current situation of an emergency, incident, or event (e.g. tornado, terrorist attack, protests, riots, et cetera). The results from such searches can then be used to determine what resources are required to effectively and efficiently respond to the situation at hand. Moreover, this search can further be used to update personnel on scene to any change in the situation.

Third, is the OSINT Framework, which sets out to assist researchers in gathering information from free tools or resources (Nordine, 2023). Headings under the framework include:

- Username
- Email Address
- Domain Name
- IP Address
- Images / Videos / Docs
- Social Networks
- Instant Messaging
- People Search Engines
- Dating
- Telephone Numbers
- Public Records
- Business Records
- Transportation
- Geolocation Tools /Maps
- Search Engines
- Forums / Blogs / IRC
- Archives
- Language Translation
- metadata
- Mobile Emulation
- Terrorism
- Dark Web
- Digital Currency
- Classifieds
- Encoding / Decoding
- Tools
- Malicious File Analysis
- Exploits & Advisories
- Threat Intelligence
- OPSEC
- Documentation Training

Similar to Bazzell's IntelTechniques and OSINT Combine's list of tools, the OSINT Framework also includes free tools. However, it also consists of tools that a researcher may have to register for, or pay to use. Unlike the former tool sets that are created by their namesakes, the OSINT Framework instead points the researcher to available tools that have been commercially developed outside of the Framework. For example, to search a username, the Framework will point the researcher to websites such as *NameCheckr*, *Thats Them*, and *Instant Username Search* (among others). This is beneficial to the researcher as it provides additional options for searching a variety of topics that the former tools may not be capable of executing. While the downside to the researcher is the reliance on external tools that may or may not be current, the upside is that alternative options are available to counter any services that may be down. Further to this, the OSINT Framework is also unique in that it offers search parameters on a number of topics that are not covered by other tool sets including threat intelligence, terrorism transportation, and instant messaging, among others. Perhaps one of the most valuable informative tools listed on the OSINT Framework is OPSEC, which includes persona creation, anonymous browsing, privacy & cleaning, and metadata - each of which are fundamental aspects of any OSINT investigation.

Fourth, is Bellingcat's Online Investigation Toolkit which has vast resources under the following headings (Bellingcat, 2022):

- Image and Video Verification
- Social Media
- People (Phone, Email, Username) Searches
- Maps, Satellites, Streetview, & Location-Based Info
- Transportation Trackers
- WHOIS, IPs, & Website Analysis
- Companies & Finance
- Environment & Wildlife
- Non-English (Social Media Sites)
- Online Security & Privacy
- Archiving & Downloading
- Data Visualization
- Guides & Handbooks
- Academic Readings / Resources
- Miscellaneous

Known for investigating subjects of public interest, Bellingcat's OSINT Online Investigation Toolkit is constructed entirely of tools external to the organization. While most of the tools listed are free, some do require registration and payment. What is particularly unique about this toolset is the availability of tools for non-english social networking platforms such as Vkontakte (Russian social networking website), Bilibili (Mandarin video sharing website), and Line (Japanese social networking site) (Bellingcat, 2022). Also unique to this toolkit are the available tools for archiving and downloading investigative data (such as Hunch.ly, the Wayback Machine, and Archive.today), and data visualization tools (such as Gephi, Maltego, and Timeline Graphics Editor). In consideration to persons of interest and incidents that occur outside English-speaking countries, these tools provide the researcher alternate avenues of approach to the investigation, which can furthermore provide additional pivot points to build upon.

Similar to the OSINT Framework, Bellingcat's toolkit also provides tools for online security and privacy. However, it assumes an alternative approach with a focus more so on protecting oneself from online tracking, browser identification, URL scanners, and data encryption. As a result of the sensitive nature of investigations, and the unintentional attention it may bring to the analyst, such material assists in protecting the individual - and the organizations they may represent - moreover contributing to personal security, operational security, and information security.

Perhaps not considered to be a "tool" in a traditional manner, Dorking and Boolean search operators - sometimes referred to as Google hacking - provide an additional means of obtaining open

source data and information in a granular way. Essentially, these search operators enhance a general search query by allowing the analyst to look for very specific data and information that is indexed within a search engine such as - but not limited to - Google. Some examples of Dorking and Boolean search operators can be found in Table 5.1.

The benefits of the above mentioned search parameters is that they enable the analyst to search for precisely what they require from a search engine, without the need to sift through the "noise" of a general search query. However, given the nature of their granular searching capability, dorks and Boolean search operators can also be used to find system vulnerabilities and private information if used in an unethical manner. While this does not outweigh its usefulness to researchers, it is something to consider when planning out one's collection methodology as sometimes such discoveries are not intentional. In such cases, it would be beneficial for the researcher to keep a log that tracks relevant information that can be passed onto the appropriate authorities in order to mitigate and respond as required.

Although the aforementioned tools can at times be used on their own, they are much more informative when combined with OSINT techniques that analysts have trained upon. This is largely because utilising techniques[105] over tools[106] allows the analyst to understand how data and information are used online, whereas tools simply collect data and information based on a specific programmed code. Additionally, tools assist with developing a baseline operational picture, whereas techniques serve investigations in a more dynamic function, especially where interaction (e.g. messaging, social engineering, et cetera) will occur. Moreover, tools are not always readily available. Whether the result of platform changes,[107] expiration of credentials,[108] coding issues, et cetera, not all tools are always readily available to the researcher. For example, if the Facebook tool on Inteltechniques.com is removed or does not work, it is imperative for the analyst to be able to manually parse through Facebook for the data they require - and know how to do so in a time efficient manner. In order to

[105] Techniques refer to the procedures the researcher / analyst might use to collect data or information. It does not exclude the use of tools but rather focuses on insight and procedure.
[106] Tools refer to the specific applications or programs that the researcher / analyst might use to collect data or information. It does not include any specific process per say, but instead focuses on a specific input to generate a result that is obtained by that tool.
[107] E.g. A social networking site changing its policies may result in tools requiring an update. This may not happen in a timely manner.
[108] E.g. Domains expiring, websites taken down by developer or malicious actor.

Google Dork	Description	Example
cache:	Show the cached version of a website	cache:https://en.wikinews.org/wiki/Main_Page
allintext:	A search for specific text on a webpage	allintext:osint tools
intext:	A search for a certain character or words within a webpage	intext:“safe browsing”
allintitle: (same as allinposttitle:)	Same as allintext:, but specific to the title	allintitle:“osint tools”
intitle: (same as inposttitle:)	Same as intext:but specific to the title	intitle:osint
allinurl:	Pages with every keyword in URL	allinurl:osintcombine
inurl:	Similar to allinurl, but only for one keyword	inurl:osint
allinanchor:	Pages with every keyword in anchor text	allinanchor:science technology and society
inachor:	Search for exact anchor text on links	inanchor:“online privacy”
filetype:	Search for specific file types	filetype:pdf password -Search for PDF files and URLs ending in .pdf whose content contains the word “password”
site:	Search for a list of indexed URLs for a specific domain or subdomain	site:docs.google.com al shabaab -Search of docs.google.com for content containing the term “al shabaab”
link:	List webpages that have links pointing to the given homepage	link:bbc.com -Lists webpages that have links pointing to the BBC homepage
*	Wildcard	osint * - The wildcard search will return pages osint software, tools, certification, etc.
|	The logical operator represented the word “or”	isis | daesh | isil | is | islamic state -Returns searches containing either search parameter
+	The logical operator represented the word “and”	osint + security + risk -Returns searches containing each parameter
-	Used to avoid showing results containing certain words	science technology -society -Returns searches for “science” and “technology” without the mention of “society”
AROUND (#)	Search for the “word” within (#) of words of the “word”	osint AROUND (3) techniques -Searches for the word “osint” within 3 words of “techniques”
“ “	Exact Phrase	"This is the way" -Searches for the exact phrase

Table 5.1: Dorks and Boolean Search Operators
Source: Borges, 2021; Eilers, 2022; Southern Adventist University, 2022
Note: Each search parameter can be used in combination with each other.

limit any negative effects from this, it is important that investigations not be dependent on the availability of tools. In other words, tools should complement an OSINT analyst's technique, not replace them.

In summary, this case study has identified several key tools that can be useful to an OSINT analyst focusing on public safety. Specifically, it has identified toolsets from inteltechniques.com, OSINT Combine, the OSINT Framework, and Bellingcat, in addition to dorking and Boolean search operators. While each of these are unique in their own manner, this section also recommended that tools should not be favoured over individual training or technique, as some tools may not always be readily available. Further to this, relying on tools may anchor an analyst into conducting OSINT in a particular fashion, while integrating new techniques will create a more dynamic search methodology that can enable an analyst to find new, more efficient ways of collecting data and information.

While some discussion of ethical concerns has already been made, it is important to reiterate those concerns here. OSINT tools and the respective search operators assist analysts in finding relevant open source data and information for their investigation. While it is possible that individuals or teams use these tools for their own purview, proper training that explains the OSINT policy, data etiquette and ethics should be made available and reviewed at least on an annual basis in order to stress its importance. Moreover, tools such as hunch.ly should be made available in order to track investigations without the loss of data, but also to be reviewed by teams to determine how to better refine their OSINT methodology.

5.3.2. DAESH Prisoners Drowned in Pool

On 23 June 2015, multiple news agencies published reports about several executions in Iraq that DAESH had filmed. One in particular portrayed the death of five prisoners in a pool said to have been located in Mosul, Iraq. This case study will use geolocation methods to determine the location of the execution site in Figure 5.1.

Figure 5.1: DAESH Prisoners in Cage Being Lowered into Pool
Source: Hall, 2015

Prior to utilising OSINT investigative techniques, the following information had been obtained from visually analyzing Figure 5.1:

1. Pool has curved edges (unique shape)
2. Decorative vertical bricks surrounding the perimeter of the pool
3. Possible letter or decoration on wall of building
4. A brick type building with an overhanging roof
5. Trees and other shrubs in location
6. A brick wall perpendicular to the building with an overhanging roof
7. Pool ladder

Other known information:

- Approximate location: Mosul, Iraq
- Date of publication: 23 June 2015

Figure 5.2: DAESH Prisoners in Cage Being Lowered into Pool
Source: Hall, 2015; Numbering: Author rendered

After visually analyzing the image, the next step was to conduct a reverse image search of Figure 5.1. This type of search would have allowed for the collection of links that showcased this similar image which could furthermore assist with the validation of the date in which the incident occurred, possible locations, validation that this incident was conducted by DAESH, and other similar incidents. Unfortunately, the reverse image search did not prove very fruitful. Instead, a Google search for "ISIS Pool Iraq" returned some images which corroborated the reporting of the incident. While some inconsistencies were noted - date of execution, number of civilians executed - most of the websites utilised the same images from the initial website used in this investigation (see images from Hall, 2015). Noting these inconsistencies, the investigation continued forward with an Iraqi news website that listed the location of the swimming pool to be in the al-Faisaliya area of Nineveh Province, Iraq.

After learning that the al Faisaliya area was located in the vicinity of the city of Mosul, another search was conducted in reference to possible locations that may have outdoor pools (palaces, hotels). This search proved to be pivotal as it returned images of American soldiers taking part in a New Year's celebration at FOB XXXX in 2006. Observing Figures 5.3, 5.4, and 5.5, it is highly likely that this pool is a match to the execution site. Additionally, these images provided a clearer detail and

views of alternative angles of the pool and surrounding structures which proved fruitful in the final confirmation of the location.

Using the same numbering system as above, the three images depict similarities to the initial identifiers found in Figure 5.2:

1. Pool has curved edges (unique shape)
2. Decorative vertical bricks surrounding the perimeter of the pool
3. Possible letter or decoration on wall of building
4. A brick type building with an overhanging roof
5. Trees and other shrubs in location
6. A brick wall perpendicular to the building with an overhanging roof
7. Pool ladder

Figure 5.3: Soldiers Ring in the New Year with a Splash
Source: Tamburello & Gravelle, 2006a; Numbering: Author rendered

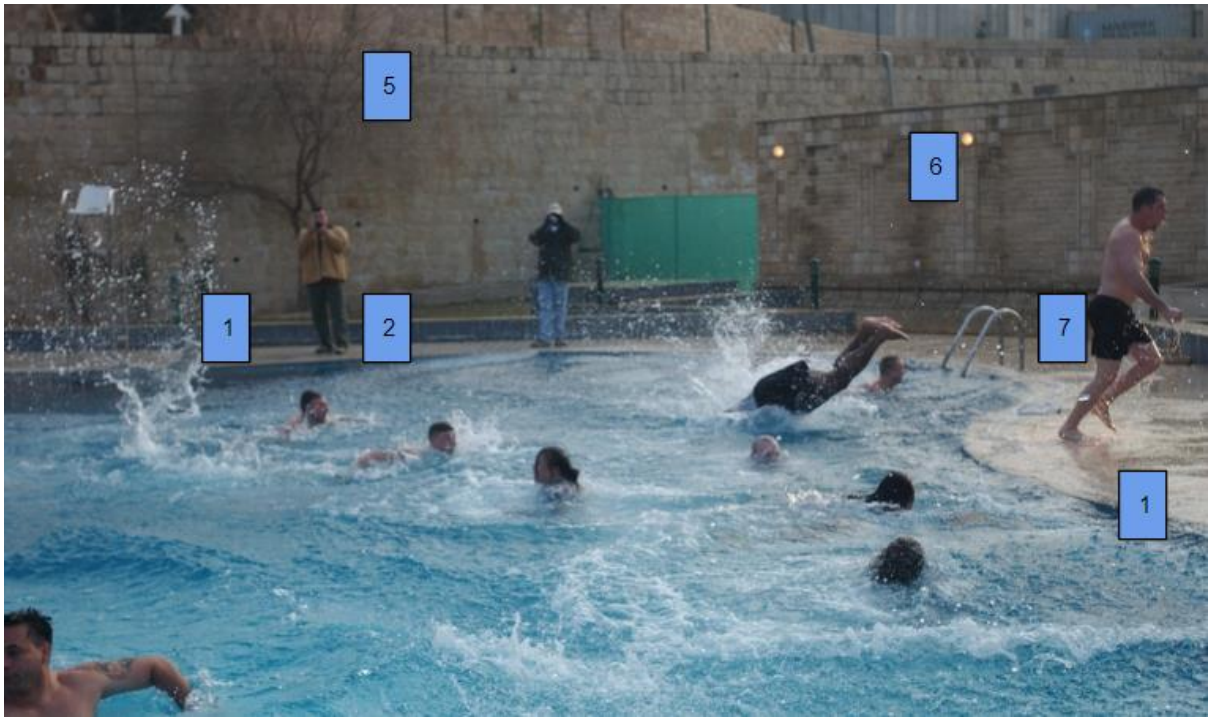

Figure 5.4: Soldiers Ring in the New Year with a Splash
Source: Tamburello & Gravelle, 2006b; Numbering: Author rendered

Figure 5.5: Soldiers Ring in the New Year with a Splash
Source: Tamburello & Gravelle, 2006c; Numbering: Author rendered

Through the analysis of these photos, it was determined that the location of the 2015 drowning execution was the former US Forward Operating Base (FOB) XXXX – also referred to as CAMP XXXX – which was furthermore identified by Tamburello and Gravelle to be a former palace of Saddam Hussein (2006c). From here, the next step was to conduct research on the locations of any palaces within the Mosul area, especially in reference to possible locations of FOB XXXX. The results from this indicated that FOB XXXX was located north west of Mosul, off the Dohuk highway, and within the Main Palace Compound[109] (Fig 5.6) in Mosul, and in May 2006, was being restored and returned back to the Iraqi government (Campbell, 2006) - prior to its occupation by DAESH.

Following a fair amount of research on palaces and US military bases in Mosul, Nineveh, one particular location had proven to be a possible match in terms of location and descriptions found online. Using Google Maps and Google Earth Pro, a terrain examination had been conducted and proved inconclusive. As a result of this, additional terrain analysis was conducted using historical imagery of the point of interest – the terrain of the Presidential Palace Complex in the vicinity of Mosul, Iraq. Of particular interest was the Independent High Electoral Commission (IHEC) Government office, located in the vicinity of the palace compound. This search had returned images that support the possibility of the pool's location within the IHEC complex (see Fig 5.6). Additional imagery analysis in Figures 5.7, 5.8, and 5.9 confirm the location of the pool. Further imagery analysis confirms the pool was filled with dirt in 2017 (Fig 5.8).

[109] A search for “Main Palace Complex” points directly to the Presidential Palaces Complex in Google Maps.

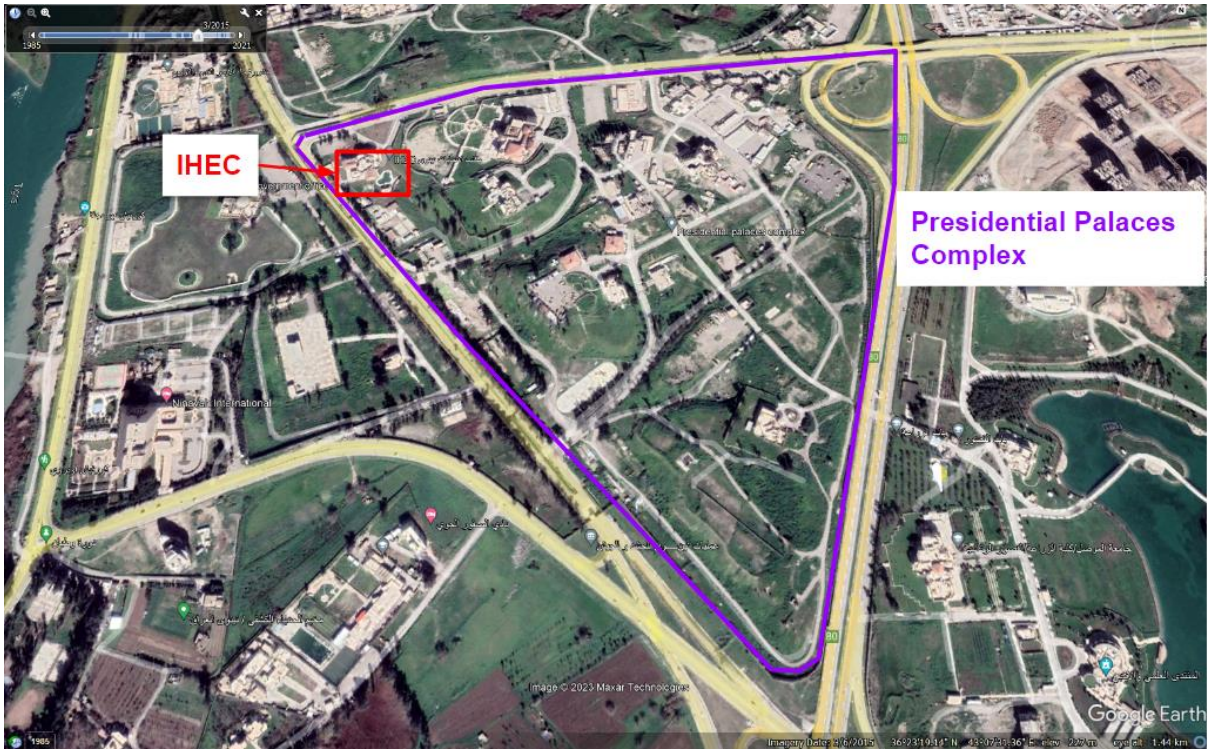

Figure 5.6: IHEC Government Office within the Presidential Palaces Complex, March 2015
Source: Google Earth Pro, March 2015a; Shape highlights: Author rendered
Note: The IHEC office was formerly a palace within the presidential palaces complex prior to US FOB XXXX and DAESH occupation.

Figure 5.7: IHEC Government Office, March 2015
Source: Google Earth Pro, March 2015b; Shape highlights: Author rendered

Figure 5.8: IHEC Government Office, June 2017
Source: Google Earth Pro, June 2017; Shape highlights: Author rendered

Figure 5.9: IHEC Government Office, August 2015
Source: Google Earth Pro, August 2015; Shape highlights: Author rendered

Figure 5.10: DAESH Prisoners in Cage Being Lowered into Pool
Source: Hall, 2015; Shape highlights: Author rendered
Assessment: Given the similarities provided in Figures 5.1 (prisoners) and 5.3, 5.4, and 5.5 (3x US pool photos), it is highly likely that both images were taken at the same location in the vicinity of Mosul, Iraq. Comparatively, given the similarities of Figures 5.1 (prisoners) and 5.3, 5.4, and 5.5 (3x US pool photos) to Figure 5.9 (2015 pool photo), it is highly likely that the images are each a depiction of the same area.

Ethical Issues

While this case study began as a simple investigation into a DAESH execution site, it developed into one that required ethical considerations in relation to revealing the location of former US operating bases within Iraq. As a practice of operational security, the location of current and former forward operating bases are generally not provided to the public as it may result in security issues for the local population, expose the location(s) of allied partners, and potentially impact future operations. Similarly, while it is not an issue at this time given the lapse in time between 2015 and the publication of this work (2023), it is also important to consider the sensitivities of reporting on the victims of DAESH executions (e.g. names, occupations, sexual orientation, purpose of execution, etc) as such sensitivities can possibly provoke social and cultural issues.

Separately, this case study showcased the significance of using historical imagery analysis tools (e.g. Google Earth Pro Historical Imagery) within open source investigations. As landscapes change through unnatural or natural events, it is critical to consider imagery from the timeframe of the

incident in order to better formulate an assessment. If not utilised in this investigation, an assessment of the location of the DAESH execution site could not have been made with the confidence level provided. Moreover, this case study employed reverse image search, search engine queries, web (Google maps) and application (Google Earth Pro) mapping tools in order to deduce the location of an image of an execution site found online. Although the location was determined with a high degree of confidence, ethical concerns about publishing the name of the location illustrated ways in which open source researchers/investigators must consider ethical implications prior to publishing their work.

5.3.3. Russian Convoy to Ukraine

This case study utilised an image from a capture the flag challenge[110] called "Borderline Crazy" taken from MilOsintCTF.com, a website that runs military themed CTF challenges. The scenario and question are as follows:

> "An asset has advised us of a military convoy approaching Ukraine from Russia. The asset has managed to retrieve an image from the Dash Cam of the convoy's lead vehicle, which they have sent to us.
> We have reason to believe the convoy will move from this position and onto the Aviation Museum in Luhansk. This would allow then [sic] to secure the Airfield and receive subsequent military aircraft.
> To the nearest 10 km's, how far by road is it from where this image was taken, to the airfield? Flag{###} km's" (MilOsintCTF, 2023).

Figure 5.11: Dash Cam Image Titled "CarCam1"
Source: MilOsintCTF, 2023

[110] Capture the flag (CTF) scenarios are often used as a means of training or to keep current on open source investigations. For an example of additional CTF scenarios, see. https://milosintctf.com/challenges

Being a geolocation challenge, it was important to make note of various portions of the image (most importantly, those that will not change or move). In the image provided, one can make note of traffic signs, tree lines, the T- shaped intersection, the markings on the road, the electrical distribution lines, and the culvert. Although these will be important to identifying the location, the first step here would be to identify an approximate location. Within this image, we can see a blue placard that is likely to be a city marker, given its location at the T-junction, a directional arrow and numerical distance indicator. After zooming in to read the sign, it is evident that it is in Cyrillic (Russian).

Figure 5.12: CarCam1
Source: MilOsintCTF, 2023

Since I have a limited knowledge of this language, I conducted a Google search for the Cyrillic alphabet, which returned an image that allowed me to compare the Cyrillic letters in the blue placard to the English alphabet (Figure 5.13). Using the figure below, I was able to determine that черников translates to Chernikov.

А а (A)	Р р (R)
Б б (B)	С с (S)
В в (V)	Т т (T)
Г г (G)	У у (U)
Д д (D)	Ф ф (F)
Е е (E)	Х х (KH)
Ё ё (YO)	Ц ц (TS)
Ж ж (ZH)	Ч ч (CH)
З з (Z)	Ш ш (SH)
И и (I)	Щ щ (SHCH)
Й й (Y)	ъ (–)
К к (K)	ы (Y)
Л л (L)	ь (')
М м (M)	Э э (E)
Н н (N)	Ю ю (YU or IU)
О о (O)	Я я (YA or IA)
П п (P)	

Figure 5.13: Cyrillic/ English Alphabet
Source: PBS, 1998

After learning the name of the city on the blue placard, a search for “Chernikov” was conducted on Google Maps, which returned two city options: Chernikov, Krasnodar Krai, Russia or Chernikov, Rostov Oblast, Russia. Alternatively, when searching specifically for “Chernikov, Russia”, the only result that appeared was Chernikov, Rostov Oblast, Russia. Since the latter result appeared to be the most popular in search, I decided to initiate a search from Chernikov, Rostov Oblast, Russia to the Aviation Museum in Luhansk (Ukraine) which provided the following suggested route (blue) and alternate route (grey):

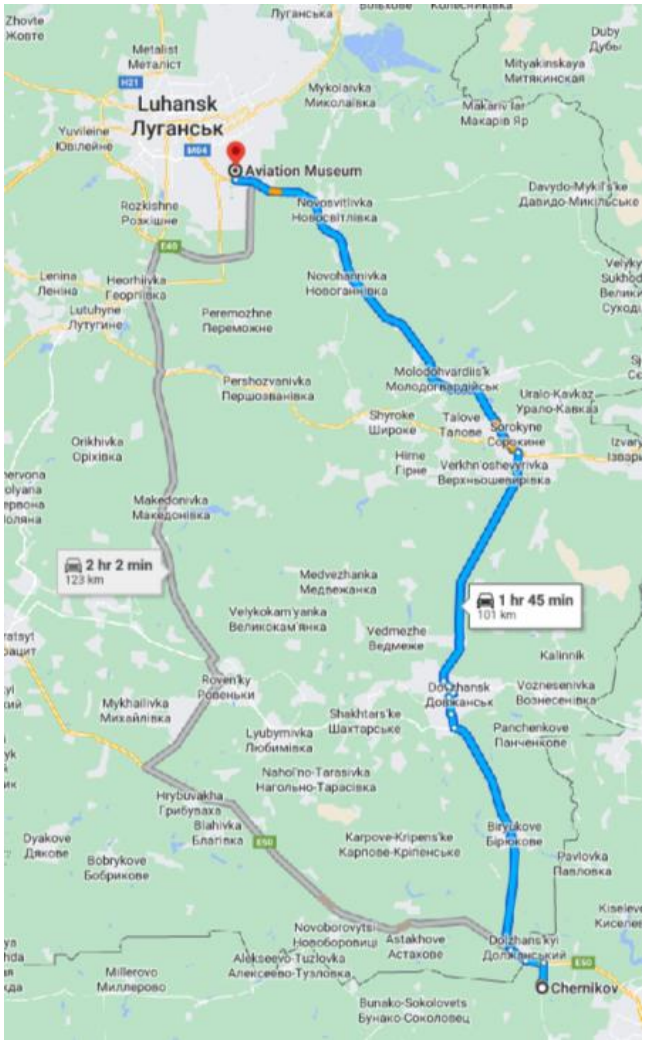

Figure 5.14: Directions from Chernikov, Rostov Oblast, Russia to the Aviation Museum in Luhansk, Ukraine
Source: Google Maps, 2023

As the scenario indicated that the convoy was travelling *from* Russia to Ukraine, the next step was to look for a T-junction approximately 2 km from Chernikov, Russia. Since there was only one major intersection between Chernikov and the E-50 highway, this was very simple to locate. However, it is important to note the importance of conducting a scan of the area to decrease the probability of error.

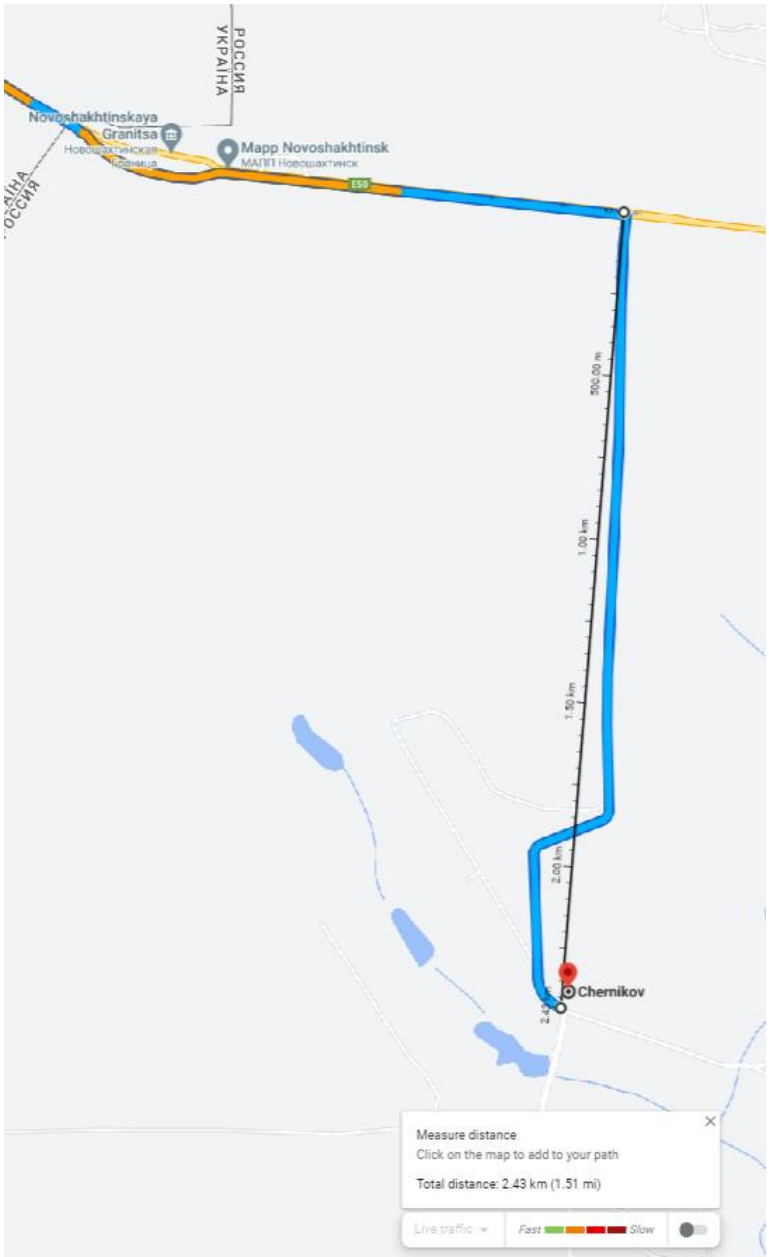

Figure 5.15: T-Junction measurement from E-50 to Chernikov, Russia is 2.43 km (1.51 mi)
Source: Google Maps, 2023

In order to confirm the location, the next step was to conduct a street view of the intersection and compare it to the original image provided from the dash cam. As depicted in Figures 5.16 and 5.17, there are at least seven key identifiers that can confirm that the location found matches the location provided, which is the E50, in Rostov Oblast, Russia (47°50'38.9"N 39°46'45.3"E). In this particular case, the convoy is travelling West on Highway E50 in Russia, towards Ukraine.

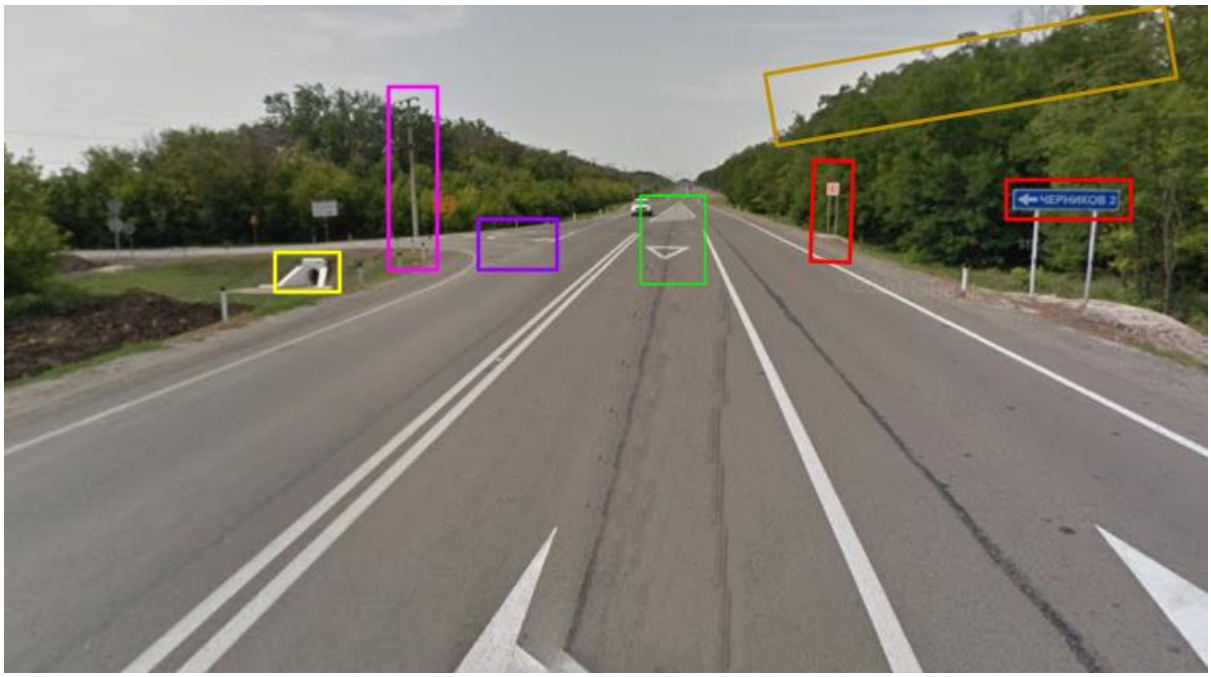

Figure 5.16: CarCam1
Source: MilOsintCTF, 2023

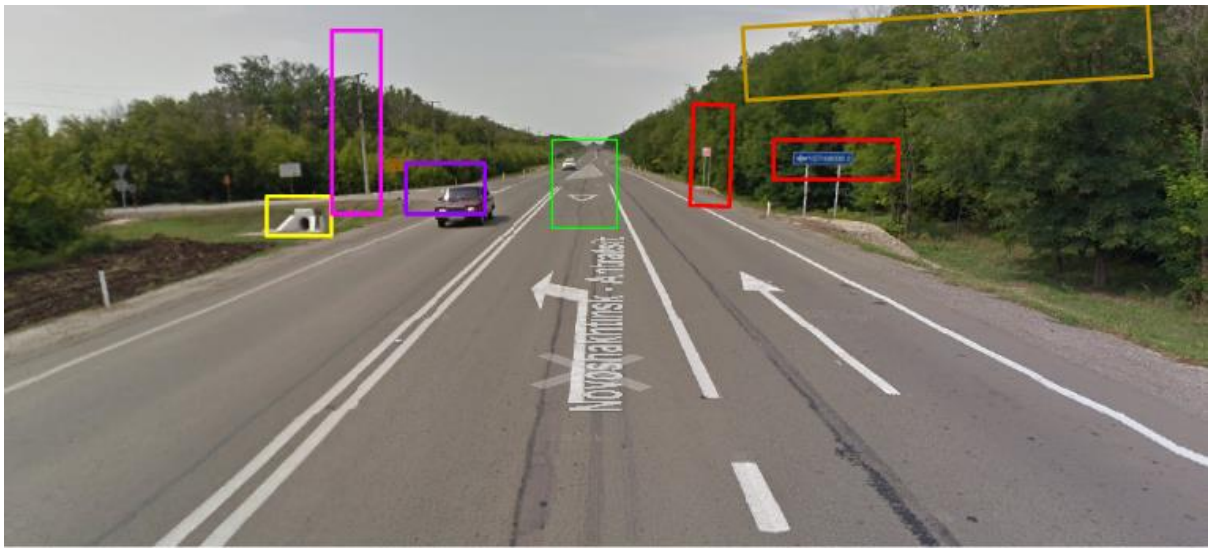

Figure 5.17: E-50 Highway, Westbound
Source: Google Maps Street View, 2023

Taking all of this into consideration to the original question, the convoy is approximately 98.7 km away from the Aviation museum, which rounded to the nearest tenth, would be 100km[111] or approximately 1 hour and 45 minutes if the convoy is travelling at the posted speed limits. For additional intelligence value, knowing the type(s) of vehicle(s) in transportation would further assist in assessing the travel time from one location to the other. For example, knowing the type of vehicle, one could assess the speed it is travelling at, its fuel capacity, and its capability/purpose, all of which could lend a hand in also determining the intent of travel (e.g. reconnaissance, electronic warfare, defensive/offensive action, troop transport, et cetera).

Ethical Issues

This case study utilised search engine queries, translation services, and web mapping tools for the analysis and geolocation of the route in question. Since this is a "capture the flag" scenario, there are no ethical issues at stake as the images had been taken directly from Google Maps, and there is no additional metadata associated with the image. However, if the dashcam footage was in fact real and was somehow leaked to the general public, knowledge of this could potentially alter the course of the convoy, which could ultimately impact Ukrainian operations and resources if they were planning any defensive operations against the Convoy - provided Russian counterintelligence was aware of the information leak. Despite this, a number of Russian locations are regularly shared on social media platforms such as Twitter, likely as a result of the amount of civilian support to the Ukrainian military, the ongoing information campaign against Russia (as the oppressor), and because at this time it can be argued that Russian information and operational security (INFOSEC and OPSEC) tactics are not on par with Ukraine.

All in all, this CTF scenario has provided a means by which to understand how various techniques can be used to geolocate a moving convoy. However, it also sheds light on ethical considerations that impact information security, operational security, and personnel security (INFOSEC, OPSEC, and PERSEC) by means of considering the impact of counterintelligence strategies that may be in use by an adversary. As people and societies become greatly intertwined in

[111] This answer was confirmed to be correct from the MilOsintChallenges website.

the online environment, so to will an array of adversaries find more purpose and effectiveness in this space. Whether in the form of intelligence collection, counter intelligence, information, psychological or cyber operations, digital security and awareness is just as critical - if not more - than physical security.

5.4. Issues: Verification

Case Study: Verification
Identify methods of verification that can be used when conducting open source intelligence.

The fabrication of information is not a phenomenon of modern times. Whether in the form of misinformation, disinformation, or propaganda campaigns, information fabrication can be said to be as old as speech itself. However, what differentiates information fabrication in the twenty-first century is the technology that enables its efficient and effective dissemination. Such fabrication can occur in varying mediums including, misinformation, disinformation or propaganda published by a news agency or government media,[112] a social media account posting manipulative content,[113] re-purposed or manipulated photos or imagery to influence a response, or a deepfake or staged video[114] leaked to benefit a particular group or cause. Undoubtedly, as advances in technology make it increasingly difficult to determine fact from fiction, forensic analysis and verification must adapt alongside such advances, especially as information becomes inherently easier to disseminate on a global scale.

As an OSINT researcher or analyst, it is fundamental to have and employ a critical thinking criterion to all investigations, especially in relation to the reliability of the source, the credibility of information, and how that information has been corroborated. Similarly, it is also important to

---

112 In April 2022, Ukraine's Foreign Ministry had released a statement condemning Russia for stealing grain and blocking shipments from Ukrainian ports. President Putin, and other Russian officials including Russia's ambassador to the U.S. Anatoly Antonov denied the accusation. In June 2022, a BBC investigation found significant evidence supporting Ukraine's claim, and satellite imagery from May 2022 published by Maxar Technologies (U.S. space technology company) had shown Russian-flagged ships at the Crimean port in Sevastopol being loaded with grain. Days later, satellite imagery of the same ship was observed in Syria - a known Russian ally. Analysis of the incident assessed Russia to be involved in laundering allegedly stolen grain from Ukraine to Syria and Turkey (Roache, et al, 2023; Maxar Technologies 2022; Windward, 2022).

113 On 24 February 2022, a threat on Twitter posted by @WarClandestine with the hashtag #USBiolabs suggested that Russia's invasion of Ukraine was meant to target U.S. bioweapons facilities located in Ukraine. The claims are a result of a misrepresentation of the Biological Threat Reduction Program, which sees the U.S. Department of Defense's collaboration with partner countries to secure dangerous pathogens, and detect outbreaks (Roache, et al, 2023).

114 In March 2023, a video circulated on social media by pro-Russian figures and supporters, showed what appeared to be dashcam footage of Ukrainian Soldiers Harassing a Woman. When online communities attempted to verify the authenticity of the video, it was geolocated to be filmed approximately 30 km from the Ukraine border, in Russia (Higgins, 2023).

consider agency bias, and what is implicated by the release of information at any given time. Some questions the researcher can ask when verifying information include, but are not limited to the following:

- How reliable is the source of information?
    - Have they been reliable in the past?
    - Are there any issues with authenticity, trust, or competency?
    - How is the source able to acquire this information?
    - What are the motives of the source?
- How credible is the information?
    - Did the source provide references for the information obtained?
        - Do the references have motives to deceive the source?
    - Has the information been confirmed by other sources?
        - Have you checked that those sources have stated what is claimed to have been stated?
    - What type of access is the information based upon?
- Has the information been corroborated?
    - By who?
    - What is their reliability?
    - What is their credibility?
- What is implicated by the release of information?
    - Was this information released at a time that would benefit a particular party?
- Is there any presence of bias or fallacy in the information?
    - Is there a bias toward a particular party?
    - Is there a bias resulting from professional or personal experiences?

In addition to the above questions, a researcher should keep a bookmark in their browser of multiple verification websites and tools to assist their verification process. These can include the following:

Ethical Guidelines for Evaluation

- UNEG Ethical Guidelines for Evaluation: https://www.alnap.org/system/files/content/resource/files/summary/UNEG_Ethical_Guidelines_for_Evaluation_2020.pdf
- Berkley Protocol on Digital Open Source Investigations: https://www.ohchr.org/sites/default/files/2022-04/OHCHR_BerkeleyProtocol.pdf

Fallacies and Biases

- List of Fallacies and Biases: https://philosophy.hku.hk/think/fallacy/
- Cognitive Biases: https://www.visualcapitalist.com/18-cognitive-bias-examples-mental-mistakes/
- Every Single Cognitive Bias in One Infographic: https://www.visualcapitalist.com/every-single-cognitive-bias/

Information Verification:

- Disinfo Database: https://euvsdisinfo.eu/disinformation-cases/
- Bot Sentinel: https://botsentinel.com/

- BotSlayer: https://osome.iu.edu/tools/botslayer
- Botometer: https://botometer.osome.iu.edu/
- Fake News Detector: http://cbcl-web.mit.edu/FakeNewsApp-public/index.html
- CaptainFact: https://captainfact.io/
- The Interactive Media Bias Chart: https://adfontesmedia.com/interactive-media-bias-chart/
- Africa Check: https://africacheck.org/
- Snopes: https://www.snopes.com/
- FactCheck.org https://www.factcheck.org/
- PolitiFact: https://www.politifact.com/
- Google Fact Check Explorer: https://toolbox.google.com/factcheck/explorer
- AllSides Media Bias: https://www.allsides.com/media-bias
- RMIT ABC News Fact Check: https://www.abc.net.au/news/factcheck
- Reuters Fact Check: https://www.reuters.com/fact-check
- Media Bias / Fact Check: https://mediabiasfactcheck.com/

Image Verification

- Forensically: https://29a.ch/photo-forensics/#jpeg-data
    - Cone detection, error level analysis, noise level analysis, and others
- Google Lens / Google Reverse Image Search: https://lens.google/#shopping
- TinEye: https://tineye.com/
    - Reverse Image Search
- Foto Forensics: https://fotoforensics.com/
    - Uses error level analysis to find EXIF data and portions of a photo that were edited

Deepfake Verification

- Deepware: https://deepware.ai/
    - Scan a video to determine if it has been manipulated

Additional Toolsets:

- RAND Disinformation Toolsets: https://www.rand.org/research/projects/truth-decay/fighting-disinformation/search.html
- College of Staten Island Fact Checking Websites: https://library.csi.cuny.edu/c.php?g=619342&p=4310783
- Alberta Teachers Association: https://teachers-ab.libguides.com/criticalthinking/factchecking

A list of tools such as this should be updated as necessary as new tools become available or as they become outdated. Not only will this provide analysts with a baseline start point, but it will also assist in any positional handovers and / or new hire positions. In relation to the above list, this section will now explore the forensic image analysis tool, Forensically.

Tool Exploration: Forensically - Image Analysis Verification

Using Forensically, one is able to examine parts of an image that cannot be seen otherwise. Take this image, which serves as the tool's sample photo that aids in showcasing its functionality:

Figure 5.18: Forensically Sample Image - Original Image
Source: Forensically, 2023

Within Forensically, one can utilise tools such as an image magnifier, clone detection, error level analysis, noise analysis, level sweep, luminance gradient, principal component analysis, meta data, geo tags, thumbnail analysis, JPEG analysis, and string extraction. We will explore these functions below.

Using the magnifier and enhancement features, one is able to magnify the pixels and contrast of the image. This is useful when one requires additional details of one of the objects within the image. While it can prove useful on its own, this tool is most valuable when used in combination with other features within the application.

Figure 5.19: Forensically Sample Image - Magnifier Tool
Source: Forensically, 2023

This exploration will begin with Foreniscally's clone detection tool, which is able to determine parts of an image that appear to be identical to each other. In the image below, the tool has determined similarities between several clouds, which can furthermore be confirmed using the magnifier tool. The clone detection tool also allows the user to scale the level of similarity, detail, cluster size, and block size as a means to assess similarity at varying levels.

Figure 5.20: Forensically Sample Image - Clone Detection Tool
Source: Forensically, 2023

In regards to the error level analysis tool. This tool is particularly interesting because it allows the user to detect edited artifacts within a JPEG image. In the image below, we are able to see portions of the image that appear in white and blue noise. For example, the blue noise represented by the floating disk not only occurs as a result of the variation in brightness of the disk itself, but is also an indication of the presence of an artifact that does not originate from the original content (Image Engineering, 2023). So, as a result of the noise within the artifact, the JPEG format will have difficulty compressing it, and will appear bright with noise amongst the original content (Wagner, 2015). Conveniently, this tool allows the user to adjust the JPEG quality, the error scale, and the opacity of the original photo. So, as one adjusts the JPEG quality to a lower level, additional noise appears in the image.

Reference the white noise below the blue noise of the floating disk in particular, we can see from the original image that this is the shadow of the disk in question. However, what is particularly intriguing is that the original photo also shows shadows cast among other parts of the area, which do not appear as noise using the error level analysis tool, thereby indicating that the shadow of the disk is not part of the original content of the image. Additionally, white noise is observed in several portions of the sky where the clouds are located in the original image. Recall from using the clone detection tool in Figure 5.20 that the tool had identified the clone of several clouds. Using this tool, the user is able to visually see that the clouds on the left have been copied from the clouds on the right, and were manipulated into the photo. Similarly, with the parachute on the right, which the error analysis tool is identifying as further image manipulation as the JPEG was not able to compress the content as well as the original content. Used in conjunction with the noise amplitude tool shown in Figure 5.22, these tools can quickly identify manipulation within JPEG images.

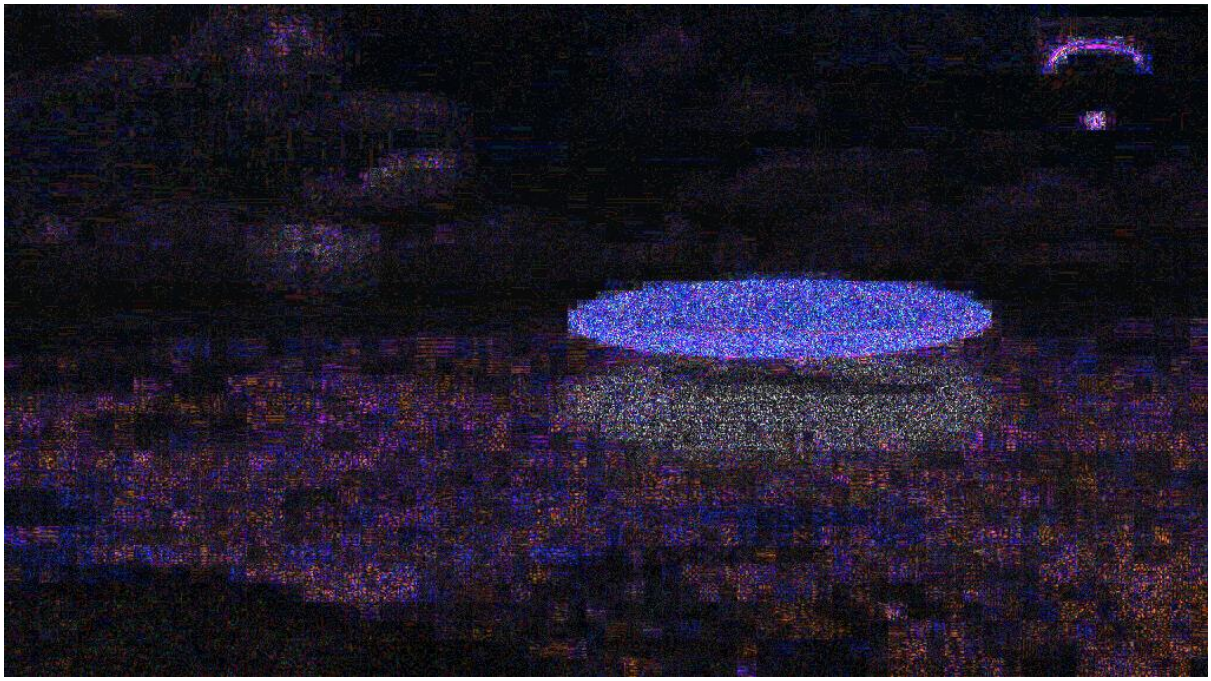

Figure 5.21: Forensically Sample Image - Error Level Analysis Tool
Source: Forensically, 2023

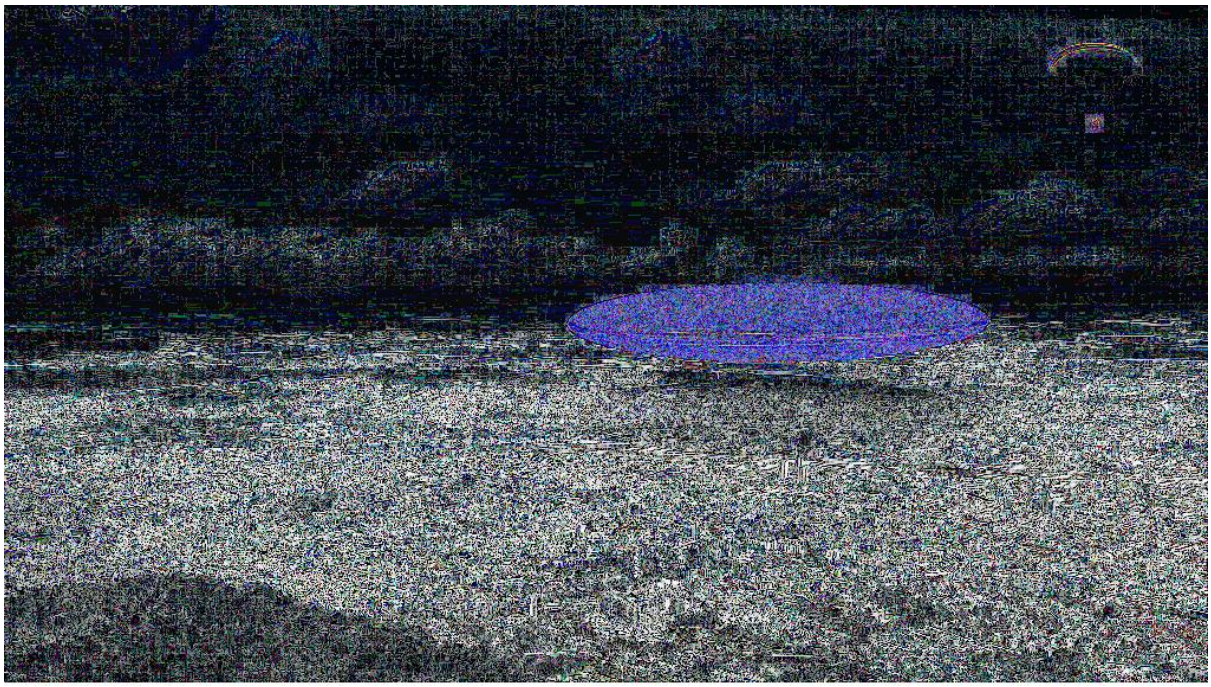

Figure 5.22: Forensically Sample Image - Noise Amplitude
Source: Forensically, 2023

Perhaps best used in conjunction with the clone detection tool, the level sweep tool below appears to increase the contrast levels of light within the image, thereby further allowing the user to identify areas which may have been copy and pasted within the image. This tool can therefore provide the user with an alternate means of identifying image manipulation.

Figure 5.23: Forensically Sample Image - Level Sweep
Source: Forensically, 2023

As depicted below, the luminance gradient tool "analyses the changes in brightness along the x and y axis of the image" (Forensically, 2023). This essentially allows the user to observe anomalies by looking at various portions, angles and edges of illumination within the image. So, when looking at an unmanipulated image, "parts of the image which are at a similar angle (to the light source) and

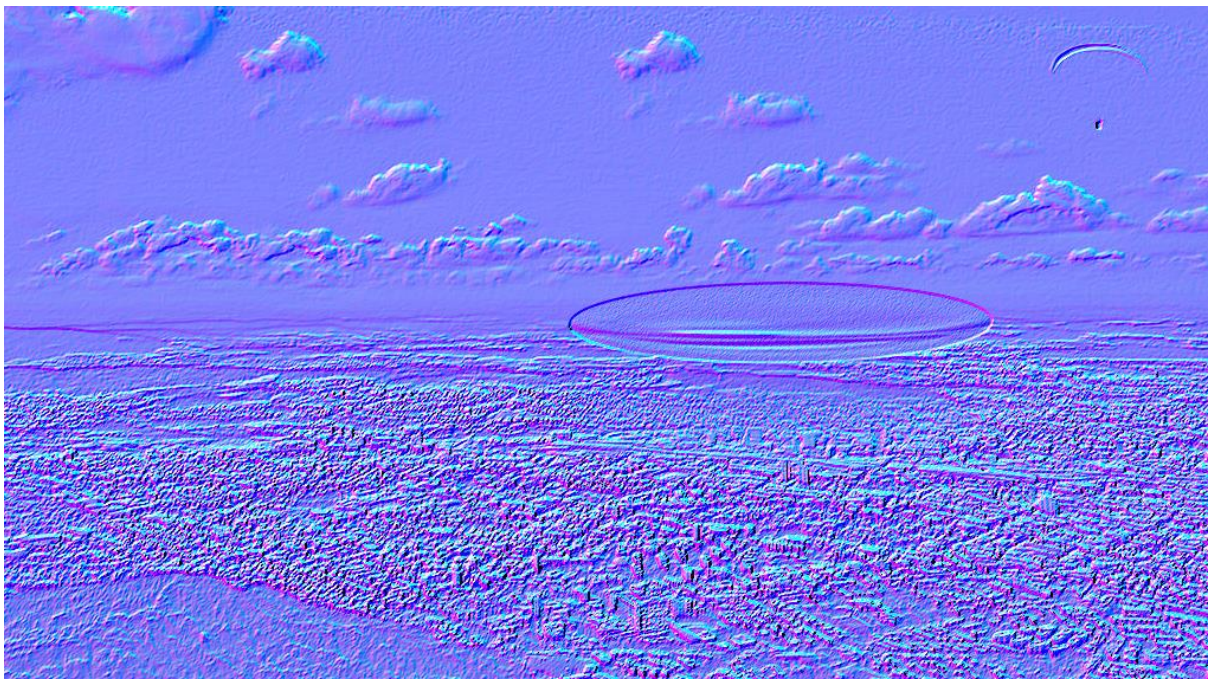

Figure 5.24: Forensically Sample Image - Luminance Gradient
Source: Forensically, 2023

under similar illumination should have a similar colour;" and the edges should have similar gradients (e.g. sharp gradients are indicated of manipulation) (Forensically, 2023).

Moving on to principal component analysis,[115] this tool allows the user to view the image from an alternative perspective. The user can view the image in three separate levels of components, which can assist in revealing hidden details within the image. The image below is taken at component level two, which has a medium variance. As this tool may show different perspectives depending on the component level, it is also best suited to be used in conjunction with other tools such as error level and noise analysis.

Figure 5.25: Forensically Sample Image - Principal Component Analysis
Source: Forensically, 2023

The metadata and geo tag tools depicted in Figures 5.26 and 5.27 pull any available EXIF data from a JPEG. For instance, the make and model of the camera that captured the image, the orientation of the image (portrait / landscape), the original date-time group, the modified date-time group, GPS coordinates, exposure time, and even the software that was used to edit the image - in this

115 "Large datasets are increasingly common and are often difficult to interpret. Principal component analysis (PCA) is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize variance. Finding such new variables, the principal components, reduces to solving an eigenvalue/eigenvector problem, and the new variables are defined by the dataset at hand, not *a priori*, hence making PCA an adaptive data analysis technique. It is adaptive in another sense too, since variants of the technique have been developed that are tailored to various different data types and structures" (Jolliffe & Cadima, 2016).

case darktable 1.6.6. Important to note here is that many social networking platforms will strip the metadata off any uploaded images. So, if a user were to download a photo from Twitter, for example, all the metadata associated with the image would be stripped away and not exploitable to the researcher or analyst.

This is also the case for the geo tag tool as it also relies on metadata. In this case however, the geo tag tool has provided the similar GPS coordinates as seen in the metadata tool; however, here it is also projected onto a world map. Nevertheless, metadata and geo tags may not always be trustworthy as they too can be manipulated. Unless corroborated with other information, metadata should always be critically assessed for possible manipulation. Unfortunately, at this time there is no tool available to assist in the assessment of manipulated EXIF data / metadata. The best method an analyst can undertake is to consider it in whole with other evidence (e.g. does the image match the indicated location and time of year it was taken in terms of weather and geography). Additionally, the user can also view the image in OpenStreetMap, Google Maps, or view other images around the provided coordinates on Flickr, which can moreover assist in determining whether or not the location matches the image and date posted in the metadata. Additionally, this method can also be used to determine if others have posted similar images, which can assist in corroborating the image. Again, depending on

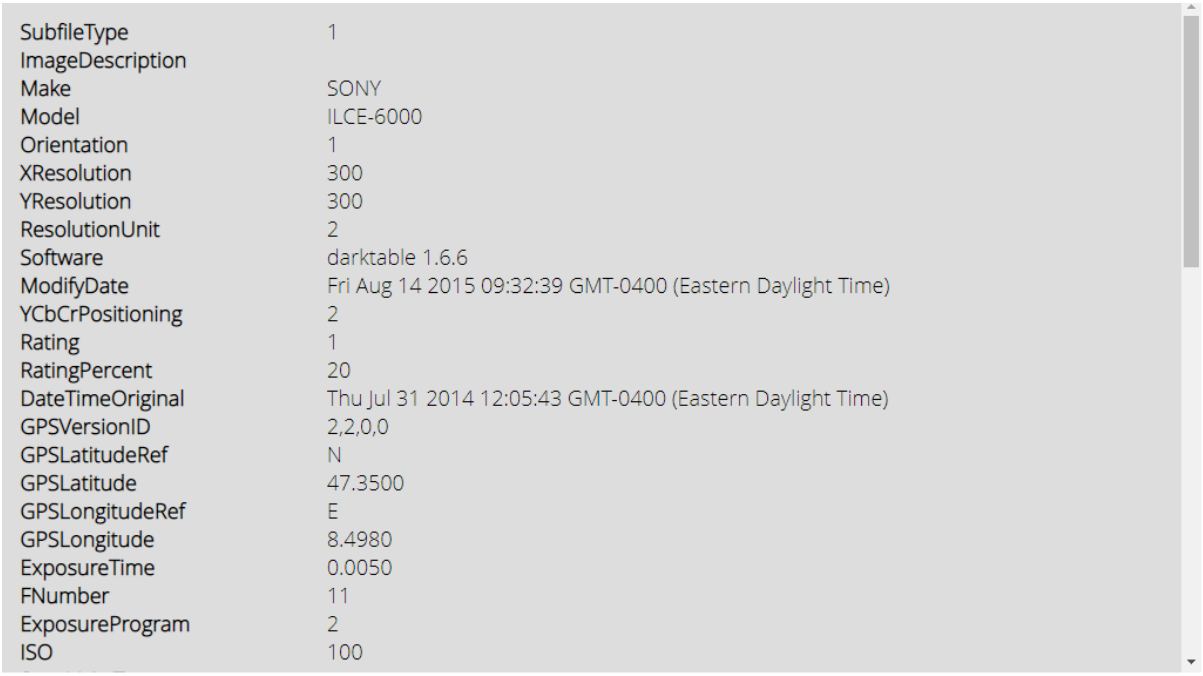

Figure 5.26: Forensically Sample Image - Metadata
Source: Forensically, 2023

the level of expertise, this may have been a factor taken into consideration by the editor, so it is important to always be critical about these results.

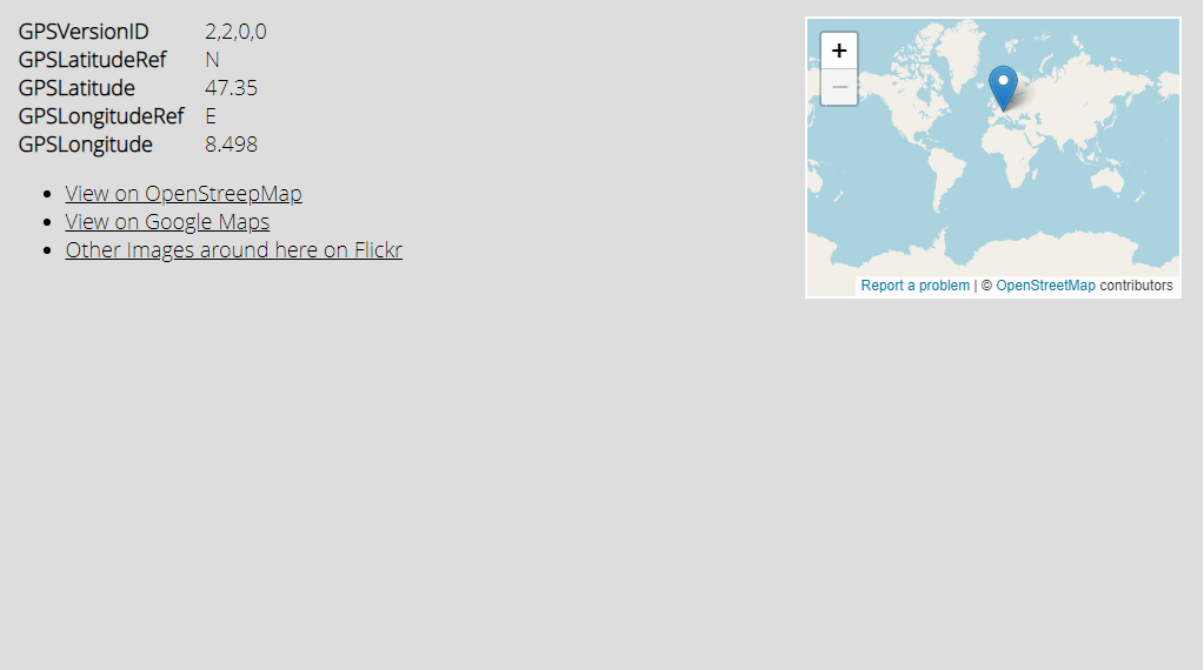

Figure 5.27: Forensically Sample Image - Geotags
Source: Forensically, 2023

Thumbnail analysis is another tool which can prove very useful, especially when comparing it to the original image. This is particularly helpful because the tool reveals the concealed image within the original image (Forensically, 2023). This is of course provided that one is available as not all cameras or software store a smaller version of the image at this time. In the case below, one can observe the original image prior to manipulation. Note the location of the paraglider, the absence of the floating disk, and the location of the clouds. In addition to the forensic analysis of the manipulated image, this thumbnail can be used side by side as further evidence of modification. If relevant, this image can also be used in a reverse image search to determine additional information of value to the investigation.

Figure 5.28: Forensically Sample Image - Thumbnail Analysis
Source: Forensically, 2023

More technical than previous tools, Forensically also employs JPEG analysis and string extraction within its toolset. In regards to JPEG analysis, this tool will capture data comments stored by applications, and will also display quantization tables, and information about the structure of sequence markers in a JPEG file (Forensically, 2023). According to Forensically (2023), "the quantization matrices used to compress a JPEG file reveals information about what software was last used to save the file in question…Most software and internet services save their files using the quantization matrices defined by the standard. The exception to this rule are Adobe products…" (np). In consideration to this, verifying or identifying different quantization matrices can be used as an indicator that the JPEG in question has either been edited, or at the very least, if it is not the original image. Regarding the structure of sequence markers in a JPEG file, Forensically (2023) indicates that JPEG images captured with a camera that has the same settings should produce the same sequence. If not, this is another indicator that the JPEG in question has been manipulated.

Figure 5.29: Forensically Sample Image - JPEG Analysis
Source: Forensically, 2023

In regards to the string extraction tool, this tool "scans for binary contents of the image looking for sequences of ASCII[116] characters" (Forensically, 2023). According to Forensically (2023), this tool functions as a "fallback" to assist in viewing metadata that is hidden, or not in a format that it currently understands. Moreover, in order to benefit most from the JPEG analysis and the string extraction tools, it is highly recommended that the researcher or analyst work together with an image analyst who specializes in these technical fields.

[116] American Standard Code for Information Interchange. ASCII codes are used to represent text within electronic communications equipment. It contains control characters, punctuation marks, digits, and the upper and lowercase English alphabet (IBM, 2023).

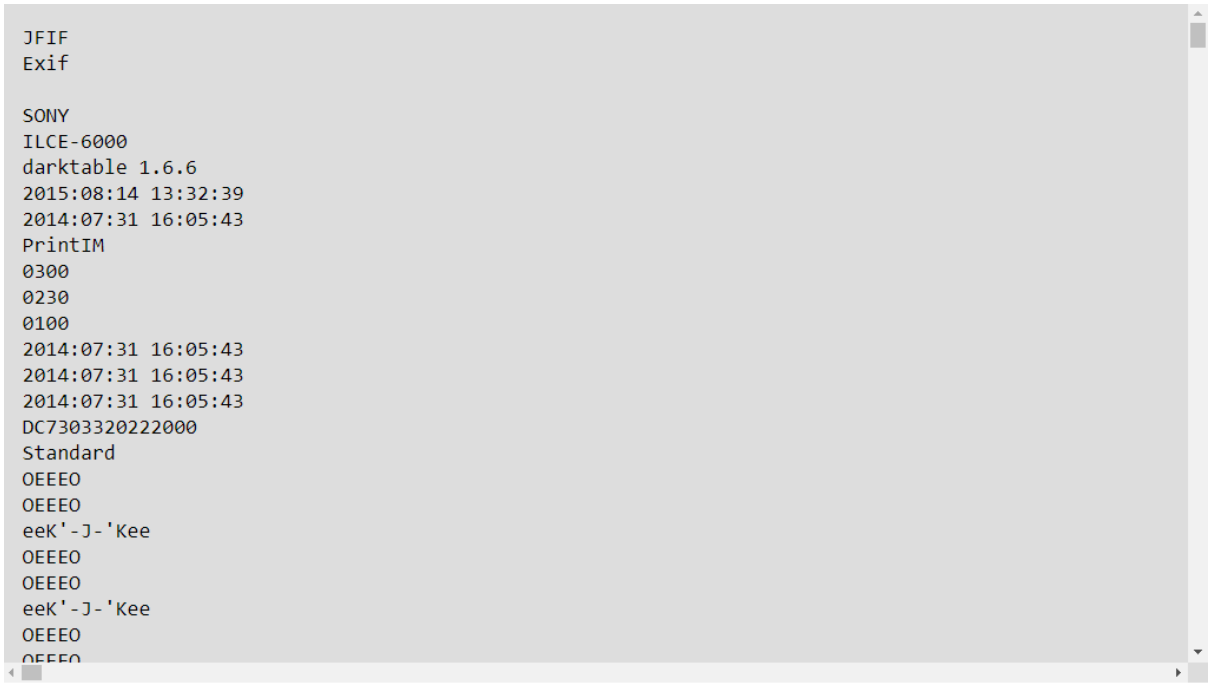

Figure 5.30: Forensically Sample Image - String Extraction
Source: Forensically, 2023

Utilised together, a critical thinking criterion and tools will equip any OSINT analyst with the ability to verify information effectively. However, it is important to note that these tools may not always be available or fully capable of detecting all false information or altered media as a result of a lack of available data or information. Additionally, in some cases, detection may not be on par with advances in media fabrication.

Earlier chapters of this work have explained the differences and similarities of the Internet and the World Wide Web, and later on, network theory and social network analysis concepts were introduced on small world theory (e.g. six degrees of separation). Due to the very nature of the Internet and the World Wide Web (herein, the Web), their structures can be said to be likened to a small world network. In experiments conducted by Barabasi and their team in 2003, it was determined that, “any document is on average only nineteen clicks away from any other” (Barabasi, 2003, 24). In consideration to the expansion of the Internet and the Web since then - including social networks, new websites, et cetera - the clustering coefficient (information available) has greatly increased while path lengths (amount of clicks to get to a page) continue to decrease.[117] This is significant when we

[117] For example, Facebook conducted a study of its 1.59 billion active users in 2016, and determined that each person is connected to every other Facebook user by an average of 3.57 other people (Edunov, et al, 2016). In 2011, a similar study from Cornell found the collective degrees of separation to be 3.74 amongst 721 million active users (Backstrom, et al, 2012; Ugander, et al, 2011).

consider the viral spread of fabricated information on the web, and furthermore demonstrates the significance of verification of information online. Therefore, just as one would explore a network of data or information, one should also conduct verification in the same manner; through the use of data and information hubs that can corroborate, contradict, or challenge a hypothesis.

5.5. OSINT and Social Network Analysis

As previously discussed in Chapter 3, network theory - specifically social network analysis - can be beneficial to an analyst as they are exploring and interpreting meaning from datasets. In a number of ways, networks can act as information maps to analysts as they develop an understanding of how a group communicates and operates. However, in some instances the use of social network analysis can actually deter the development of meaningful results. A critical consideration for social network analysis then is that an analyst must have an understanding of the type of data they are using, and how they plan on using the results of the analysis. This will ensure that the analyst inputs the correct type of data that is properly suited to the intent of their work. Thus far, this chapter as a whole has discussed ways in which OSINT may be used to develop situational awareness, to geolocate a person, group, or location, and how to conduct verification of information. At this time, this section will discuss the ways in which social network analysis may aid in the analytical process of conducting an OSINT investigation, thereby illuminating the interpretation of results that an analyst can deliver to their team.

While often associated with social media, social network analysis is not limited to social networking websites. However, given the type of data found on such websites, data retrieved from these sources is quite valuable to social network analysis especially since this type of analysis is largely focused on investigating social structures. While investigating social structures, an analyst may come across information that may reveal lines of communication (e.g. how information is passed in a network), communication hubs (e.g. a key actor(s), central actor(s) that link others together), hierarchical relationships (e.g. leaders, couriers, operators), possible recruitment zones, and a group's communication limit (e.g. local, regional, international, et cetera), to name a few. Combined together,

this information can provide valuable insight about a particular group that could then be leveraged to determine vulnerabilities to exploit within an investigation.

In combination with data that may be obtained from social media, other sources of data that also contribute to maximizing the output of social network analysis include metadata, mobile phone data, and/or computer forensics (e.g. contact lists, messages, application way points, et cetera). Together, these multiple data sources serve to better inform and provide context to datasets an analyst may already have. Coupled with other analytical practices such as geographic information systems (GIS)[118] and Intelligence Preparation of the Operational Environment (IPOE) or Intelligence Preparation for the Battlefield (IPB),[119] analysts can generate rather strong assessments using such methods of predictive analysis. These assessments can be further supported by creating social network analysis (SNA) simulations that showcase how a network could possibly react if a particular node was removed, which can moreover be used to identify pivot points in an investigation, refine targeting efforts, plan psychological operations, exploit gaps in the network, plan collection efforts, et cetera. Figures 5.31, 5.32, 5.33, 5.34, 5.35, 5.36, 5.37, and 5.38 depict a faux small insurgent cell's communication structure as an example of conducting social network analysis simulations where nodes are removed from the network to determine the impact on the flow of information to each insurgent group (represented by different colours).

As depicted in the figures below, removing a particular hub[120] from a network can have varying results. In this simulation, Figure 5.31 represents the baseline, non-fragmented network. As depicted in Figure 5.32, removing the Blue hub shattered the main (central) point of information output, limiting the hubs to communicate only with their adjacent neighbouring hub. In Figure 5.33, removing the Green hub isolates the rest of the Green nodes from obtaining any information. Given

---

[118] For example, in combination with GIS, social network analysis can be used to determine communication, temporal, location, and group patterns (among others). An example of this could be trying to understand an insurgent group's hierarchical structure, lines of communication, capabilities, and attack patterns.

[119] The overall process of IPOE and IPB (see footnote 81) are the same: Step 1: Define the battlespace (significant characteristics of the environment); Step 2: Describe the battlespace (terrain analysis and effects on operations); Step 3: Evaluate the Threat (assess the threats, identify targets); Step 4: Determine threat Courses of Action - most likely and most dangerous) (Department of National Defence, 2001). However, the main difference between them lies in the overall scope of their analysis. IPB primarily focuses on geospatial analysis (e.g. weather, enemy, terrain) with some consideration to socio-economic and cultural factors. In contrast, IPOE consists of both a geospatial and systems analysis of the operating environment, and includes a complete assessment on political, economic, military, social, information, and infrastructure (PMESII) in relation to area, structures, capabilities, organization, people, and events (ASCOPE) (La Pierre, 2020). Also see Chapter 3, footnote 81.

[120] Hub: A node that has a large degree (many links) (Nykamp, 2022).

the current one-way relationship between the Green and Purple networks, a power struggle may occur or the remaining Green network may simply join or establish new relations with the Purple network in order to obtain information from the Blue node. With the removal of the Purple hub in Figure 5.34, one can see that the two of the three Purple nodes are isolated from the rest of the network, with only one able to reach into the rest of the network via the Green hub. The removal of the Red hub in Figure 5.35 slightly fragments the rest of the Red network, however, they are still able to obtain information from the Blue hub via the Pink hub. While the removal of the Pink, Yellow, and Orange hubs are not as impactful as removing the Purple, Green, and Red hubs, (see Figures 5.36, 5.37, and 5.38), they still fragment the network by limiting the access to neighbouring hubs. Depending on the overall intent of the simulation and the desired effect the analyst wishes to achieve, any one of these simulations can be used to illustrate various assessments, including but not limited to: the fragmentation of a communication hub, the spread of information, a possible hierarchical relationship, or even the assessed impact of a targeting effort within an open source investigation.

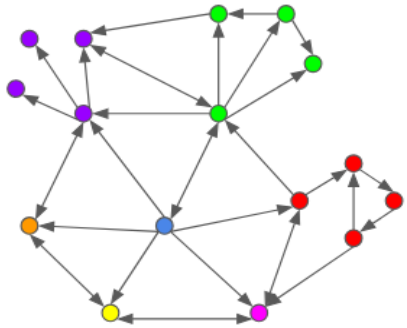

Figure 5.31: SNA Simulation of Communication Flow from Blue Hub
Source: Author Rendered

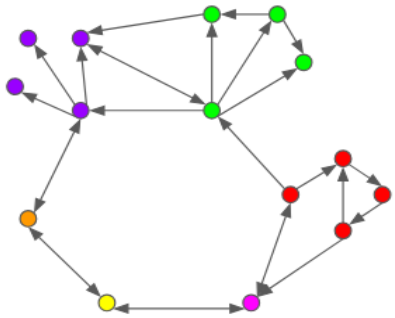

Figure 5.32: SNA Simulation of Communication Flow from Blue Hub - Blue Hub Remove
Source: Author Rendered

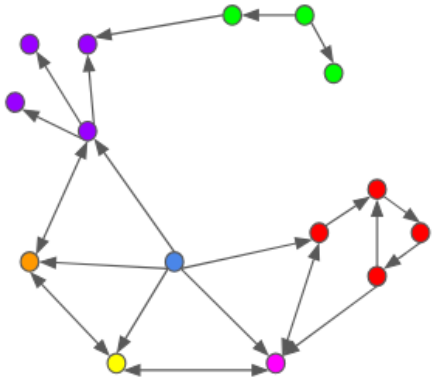

Figure 5.33: SNA Simulation of Communication Flow from Blue Hub - Green Hub Removed
Source: Author Rendered

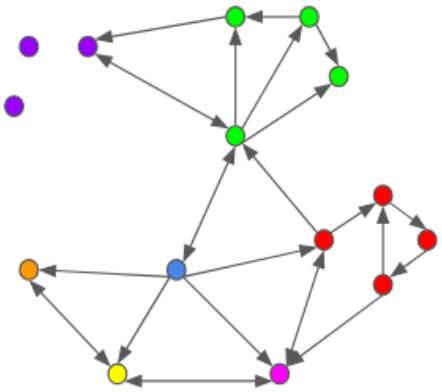

Figure 5.34: SNA Simulation of Communication Flow from Blue Hub - Purple Hub Removed
Source: Author Rendered

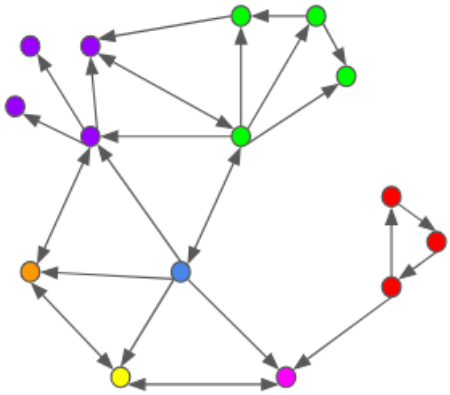

Figure 5.35: SNA Simulation of Communication Flow from Blue Hub - Red Hub Removed
Source: Author Rendered

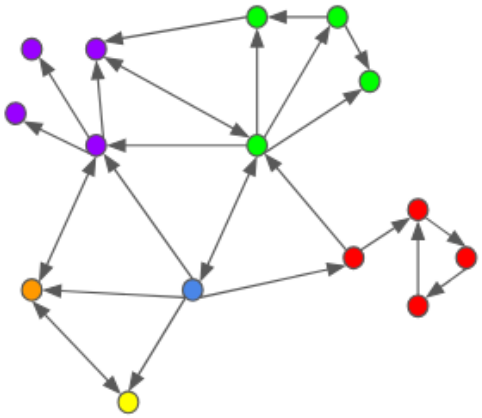

Figure 5.36: SNA Simulation of Communication Flow from Blue Hub - Pink Hub Removed
Source: Author Rendered

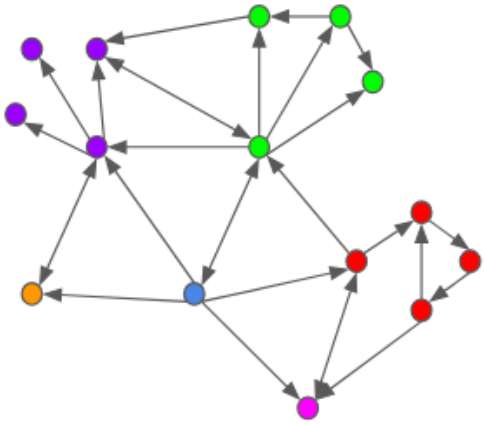

Figure 5.37: SNA Simulation of Communication Flow from Blue Hub - Yellow Hub Removed
Source: Author Rendered

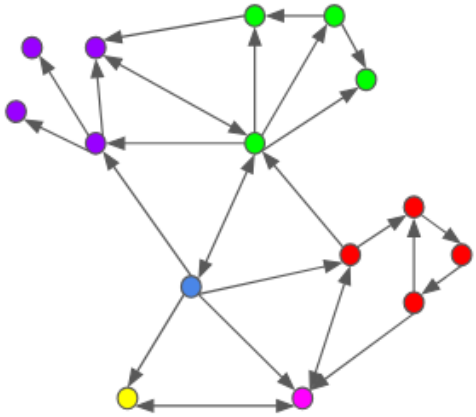

Figure 5.38: SNA Simulation of Communication Flow from Blue Hub - Orange Hub Removed
Source: Author Rendered

Despite the wide ranging applicability of social network analysis within OSINT, it also has some disadvantages. For instance, while social network analysis can be done using small datasets, it is likely most useful with larger datasets of *relevant* information in order to ascertain the big picture. This can be problematic as datasets often need to be purchased from companies that collect such data. These purchases often come with caveats as to how the data can be used, and how it can or cannot be shared. Additionally, social network analysis can require technical skill sets and be fairly time consuming to develop models and simulations that best fit what the analyst is attempting to assess. While large datasets may be preferred in order to develop a clear understanding of what is happening, too much data may degrade the entire analytical process. Therefore, it is important that the analyst ensures the data is clean and specific to what they require so their analysis is not distracted by background noise.

Overall, social network analysis maintains both advantages and disadvantages to its use. While it has a wide range of applicability for use in OSINT, an analyst must ensure they are using the data judiciously, and that they have the skillset to ensure the data is represented appropriately. Used in this way, social network analysis can illuminate data that may otherwise be overlooked or unseen, thereby informing an analyst's predictive analysis process. Additionally, social network analysis can also provide an analyst with an alternative view to understanding how a network functions, what their vulnerabilities may be (or where they can be created), and where they should focus their research. With this, an analyst can then take this information and conduct specific investigations online that seek to exploit their person, group, or area of interest. It is to this extent that social network analysis and networks as a whole serve as information maps to OSINT analysts.

5.6. ATIP Responses

Case Study: ATIP
Submit an access to information and privacy (ATIP) request to identify OSINT training resources and methods to the following government agencies: the Department of National Defence (DND), Transport Canada (TC), Financial Transactions and Reports Analysis Centre (FINTRAC), Global Affairs Canada (GAC), Canadian Border Services Agency (CBSA), Royal Canadian Mounted Police (RCMP), Canadian Security Intelligence Service (CSIS), and Public Safety Canada (PSC).

Access to Information and Privacy (ATIP) Online Request:

- I am seeking all records regarding Open Source Intelligence, specifically mentioning training, policies, and risks associated, from 2017 to present (2022).[121]

On 9 and 10 November 2022, the above request was submitted to eight federal agencies in an attempt to identify and assess the methods and resources allocated to OSINT training. Over the course of six months, a number of responses to the requests were made available. For clarity, this section will reference each government agency's response separately.

5.6.1. Department of National Defence
Response dated 28 November 2022 - Tracking Number A-2022-01333, File 2700469

The Department of National Defence (DND) had provided two documents in response to the ATIP request: (1) Briefing Note for the Deputy Minister and Chief of the Defence Staff: Open Source Intelligence Activities on the Internet, and (2) the Chief of Defence Intelligence (CDI) Functional Directive: Conduct of Open Source Intelligence Activities on the Internet.

While a significant portion of each document is redacted, there are some key pieces of information to note. For instance, the background section of the briefing note indicates that "CDI is in the process of revising a functional directive on the conduct of Open Source Intelligence Activities on the Internet (OAI). The previous directive, issued in February 2017, is outdated" (Chief of Defence Intelligence, 2020, 1). The revised document that this briefing note references is the second document

[121] Note: Some departments required additional clarification and specification from the original request. As a result, some departments received a question in different wording in order to reference specific files. However, the request provided here is a representation of how the question originated.

included in the request, with an effective date of 1 November 2019. The background statement thereby indicates that a revision of policy has occurred just over two years after the initial directive - Chief of Defence Intelligence Functional Directive: Framework for the Conduct of Intelligence Activities Using the Internet (IAI), February 2017[122] - was issued.

What is significant about the OAI directive is that despite its approval to take immediate effect, it remains in draft form without the signature of the Chief of Defence Staff (as of the ATIP request time of release, 28 November 2022). While not inherently stated in the released information, it appears as though the functional directive has been in a state of revision for over four years since its effective date of 1 November 2019. Unfortunately, it is not abruptly clear as to why this particular document has not been formally approved. However, some theories can be made.

First, is the acknowledgment of the effects that COVID has had on public and private institutions. Although the pandemic has had a significant effect on all government departments, this particular reason does not justify a four-year revision cycle, especially since this document was disseminated as one which superseded the former IAI document, and was approved for use. Second, is the acknowledgement that this document was drafted during the period that Bill C-11: Digital Charter Implementation Act, 2020 (2020-2021) and Bill C-27: Digital Charter Implementation Act, 2022 (2022-present) were both consecutively introduced to the House of Commons. While the implementation of Bill C-27 is likely a factor to consider in the formal approval of the OAI Functional Directive, it too should not be the reason that such directives are not officially approved. This is not to say that Bill C-27 will not have an effect on DND's policies and authorities on using open source data and information for intelligence purposes. Rather, it is implying that analysts employing such a capability require a solid foundation of authorities and policies to reference as they are conducting their work. Disseminating an unsigned draft and indicating it supersedes other directives can create confusion and division among units and departments with a mandate and capability to conduct OSINT, which can furthermore impede operations, limit the use of the capability, and / or create legal issues for collection later on.

[122] The Chief of Defence Intelligence Functional Directive: Framework for the Conduct of Intelligence Activities Using the Internet, February 2017 was not included in the release of information as the original request had a timeline of 2017-Present (November 2022).

On a separate note, the table of contents provides a list of subjects that the directive discusses. Of particular interest are sections (4) Legal Authority, (5) Policy Direction, (6) Information Management, (7) Gender Perspectives, (8) Roles and Responsibilities, and (9) Approval (Chief of Defence Intelligence and Director General Intelligence Policy and Partnerships, 2019). Although each section is redacted, the subject titles themselves are indicative of an ethical approach concerning open source collection. However, this cannot be confirmed without access to the subject matter therein. Similarly, an assessment of how training is conducted or how risk is assessed cannot be completed as a result of the redacted content. Although training and risk are not distinctly mentioned as subject matters, they may be discussed in general under subjects eight and nine of the directive. However, it is also possible that the document does not reference these matters specifically, and instead relies on individual mandated units to develop their own training plan and means of assessing risk. While this allows individual units and departments to customize training and assess risk as it relates to their level of operation (e.g. tactical, operational, strategic), it also creates an unestablished baseline from which to grow.

5.6.2. Transport Canada

Response dated 17 February 2023 - Tracking Number A-2022-00381

In a number of ways, the response to the ATIP request provided by Transport Canada was vastly different than that of the Department of National Defence. Where DND referenced policy and authority, Transport Canada simply provided a list of intelligence related courses it provides to its analysts. The courses themselves take place online, and are offered through the Intelligence Analyst Learning Program (IALP) - a joint effort between members of Canada's intelligence community, and are administered by the Privy Council Office (Privy Council Office, 2023). The list of courses provided is as follows:

- Open Source Intelligence Level 1 - 5 days / 1:30-2:30 + 3:00-4:00 / Total:10 hours
- Open Source Intelligence Level 2 - 5 days / 1:30-2:30 + 3:00-4:00 / Total:10 hours
- Structured Analytical Techniques Learning Series - 5 days / 10x 1 hour sessions / Total: 10 hours
- Analytic and Critical Thinking for Intelligence Analysts - 4 days / 8x 1 hour sessions / Total: 8 hours
- Introduction to Canadian Intelligence - 2 days / 4x 1 hour sessions / Total: 4 hours
- Effective Writing Course - 5 days / 10x 1 hour sessions / Total: 10 hours

- Argument Mapping for Intelligence Analysts - 4 days / 8x 1 hour sessions / Total: 8 hours
- Visual Thinking - 2 days / 4x 1 hour sessions / Total: 4 hours

As indicated from the above list, Transport Canada provides its analysts with two OSINT courses - an introductory level course designed to teach analysts the "strategies and tactics for collecting and exploiting open source information," and a second level course that teaches students "to plan and use efficient search strategies" (Privy Council Office, 2023). According to the course description, the students also learn methods of evaluating online information sources and how to maintain operational security, which is critical to identifying misinformation, disinformation, and protecting the individual and organization. Additionally, the second level course indicates advanced security, ethics, and case management (e.g. chain of evidence) as a key course topic, which is crucial to ensuring that open source collection occurs in a legal, proportionate, and purposeful manner. Unfortunately, the extent to which each of these topics is taught is unclear. Given the length of the courses, it is unlikely that any in-depth discussions occur on these matters, thereby putting the onus on the individual to pursue knowledge on their own.

Although individual pursuit of knowledge is not necessarily a bad thing, it can be problematic for setting standards within an organization. This can result in differing approaches to ethical interpretations, unstandardized case management, and varying acceptable levels of risk within an organization. Either of these can pose issues when such evidence is used in court, during an interrogation, or for targeting purposes. However, an individual pursuit of knowledge can also be greatly beneficial to an organization. This is especially true when it is in the form of an added skillset (e.g. tool or technique specialization). But, the values that an organization depends on to maintain its capability should not occur over an hour long serial or as an individual learning pursuit.

5.6.3. Financial Transactions and Reports Analysis Centre

Response dated 11 January 2023 - Tracking Number A-2022-00065

At the time of response, the Financial Transactions and Reports Analysis Centre (FINTRAC) had indicated via email correspondence that the requested records were non-existent (Widdis, 2023). However, pursuant to subsection 4(3)[123] of the Access to Information Act, FINTRAC determined it reasonable to produce the information in order to respond to the request.

With respect to OSINT training, FINTRAC had indicated that they leverage both internal and external training sources as a means to develop their strategic intelligence analysts. While internal training is largely developed and conducted by senior strategic intelligence analysts, FINTRAC also turns to subject matter experts within the private sector to obtain formal certification for open source intelligence courses (Widdis, 2023). Unfortunately, sample training material and course/company names were not included in the response, which limits the evaluation of content at this time.

In addition to this, a total of $43, 750.67 was forecasted as the budget allocated to the cost of training courses for the strategic intelligence and research team for the fiscal year 2022-2023 - of which, $2,750 had been allocated to introductory and intermediate OSINT courses (Widdis, 2023). As a result of not knowing the cost of each course or the amount of analysts this funding is directed toward, it is not feasible to determine if this is a sufficient amount. However, given the average cost[124] of OSINT training, this number does seem relatively low, and it is therefore likely that this figure is directed toward 1-5 analysts.

On the matter of legislation that governs the collection of publicly available information (PAI), FINTRAC has cited section 54(1)(b)(i) of the Proceeds of Crime (Money Laundering) and Terrorist Financing Act, 2000, which states the following:

[123] Records produced from machine readable records (3) "For the purposes of this Part, any record requested under this Part that does not exist but can, subject to such limitations as may be prescribed by regulation, be produced from a machine readable record under the control of a government institution using computer hardware and software and technical expertise normally used by the government institution shall be deemed to be a record under the control of the government institution" (*Access to Information Act*, 1985, 4-5).

[124] For example, these companies charge the following for their courses, per person: OSINTTechniques.net: $649-$949 USD; SANS $8,275 USD; Toddington $499 CAD, Canadian Police College $2,950 CAD; McAfee Institute $497-$997 USD; McMaster University $785-$1,049 CAD; OSINT Combine: $599-$2,500 USD.

"54 (1) The Centre[125]
(b) may collect information that the Centre considers relevant to money laundering activities or the financing of terrorist activities and that
(i) is publicly available, including in a commercially available database" (*Proceeds of Crime (Money Laundering) and Terrorist Financing Act*, 2000, 51-52).

While this Act indicates that the Centre may collect publicly available information, it is quite vague in nature, lacking any reference to ethics, limitations, roles, responsibilities, accountability, operational security, or personnel security. Additionally, it also fails to define key terms such as publicly available information, and personal identifying information. The lack of reference to such terms is likely because the information is not oriented specifically toward OSINT, but instead toward financial crime and terrorist financing in general.

Also interesting to note are sections 54(1)(d) and (e) of the Act which discuss information retention and destruction, and section 72(2) which discusses privacy and information protection. They indicate the following:

"54 (1) The Centre
…(d) subject to section 6 of the Privacy Act, shall retain each report referred to in paragraph (a) and all information referred to in paragraph (a) or (b) for 10 years beginning on the day on which the report is received or the information is received or collected; and
(e) despite the Library and Archives of Canada Act,[126] shall destroy, 15 years after the day on which a report referred to in paragraph (a) is received, any identifying information contained in the report if the report was not disclosed under subsection 55(3), 55.1(1) or 56.1(1) or (2)" (*Proceeds of Crime (Money Laundering) and Terrorist Financing Act*, 2000, 51-52).

"72 Review by Privacy Commissioner
(2) Every two years beginning on the day on which this section comes into force, the Privacy Commissioner, appointed under section 53 of the *Privacy Act*, shall review the measures taken by the Centre to protect information it receives or collects under this Act and shall, within three months after the review, submit a report on those measures to the Speaker of the Senate and the Speaker of the House of Commons, who shall each table the report in the House over which he or she presides without delay after receiving it or, if that House is not then sitting, on any of the first 15 days on which the House is sitting after the Speaker receives it" (*Proceeds of Crime (Money Laundering) and Terrorist Financing Act*, 2000, 89).

[125] The Proceeds of Crime (Money Laundering) and Terrorist Financing Act, 2000, refers to FINTRAC as "the Centre" throughout the Act.

[126] Section 12(1) of the Library and Archives of Canada Act states "No government or ministerial record, whether or not it is surplus property of a government institution, shall be disposed of, including by being destroyed, without the written consent of the Librarian and Archivist or of a person to whom the Librarian and Archivist has, in writing, delegated the power to give such consents" (*Library and Archives of Canada Act*, 2004, 6).

Together, these two sections complement section 54(1)(b) by providing some direction on the retention, destruction, and protection of information. However, they fail to provide direction with regards to how this is to occur, which leaves the process open to interpretation. This is problematic especially as it relates to data breaches and leaks. Although each of these sections provides relevant information on the issues of information collection, retention, destruction, and protection, they could benefit from additional detail catered specifically toward OSINT and PAI. However, this type of information is better suited to appear in the form of an OSINT directive as it can provide the details required in a form that can be more easily refined and updated than an Act. For example, DND's *Chief of Defence Intelligence (CDI) Functional Directive: Conduct of Open Source Intelligence Activities on the Internet*, is a positive start point in providing a functional directive for all OSINT analysts as it is plainly written specifically for the capability. This is not to say that CDI's functional directive as a whole meets the requirements of an effective OSINT directive - the document is too redacted for evaluation - but rather that it demonstrates a suitable approach of direction for OSINT analysts that is separate from but complementary to, formal Acts and legislation.

5.6.4. Public Safety Canada

Response dated 8 May 2023 -Tracking Number A-2022-00241 / JKB

On 8 May 2023, Public Safety Canada sent a response to the ATIP request via the ATIP Online Request Service Portal. The response contained introductory information on Canada's involvement with the European Centre for Countering Hybrid Warfare,[127] Helsinki. The only portion of the document that made any reference to open source intelligence was specific to the European Centre's identification of external capabilities provided by participating states, for example, open source (OSINT) training from the United Kingdom (Public Safety Canada, 2023, 7).[128] Outside of this, Public Safety Canada references the same two Intelligence Analyst Learning Program (IALP) courses that Transport Canada had listed in their response - Open Source Intelligence Level 1 and

[127] There are five major interpretations of hybrid warfare: (1) "the employment of synergistic fusion of conventional weapons, irregular tactics, terrorism, and criminal activities in the same battlespace; (2) the combined use of regular and irregular forces under a unified direction; (3) the use of various military and non-military means to menace an enemy; (4) sub-threshold activities involving any mix of violent and non-violent means; and (5) a way of achieving political goals by using non-violent subversive activities" (Solmaz, 2022, 1).

[128] The documents received consisted of approximately 100 pages, most of which were redacted, out of scope, or irrelevant to this work. The relevant pages have been included in Appendix 14.

Level 2. A discussion on this matter would be redundant to section 5.5.2 as no new information has been made available. Therefore, for a discussion on these courses, please refer to section 5.5.2.

In regards to legislation, this work looked at the *Department of Public Safety and Emergency Preparedness Act*, 2005. At this time, the Act did not contain any information pertaining to the collection of information or data. The most relevant information in the legislation was under the heading titled, "Powers, Duties and Functions of the Minister" which reads:

> "Portfolio Coordination and Leadership
> 5. The Minister shall coordinate the activities of the entities for which the Minister is responsible, including the Royal Canadian Mounted Police, the Canadian Security Intelligence Service, the Canada Border Services Agency, the Canadian Firearms Centre, the Correctional Service of Canada and the National Parole Board, and establish strategic priorities for those entities relating to public safety and emergency preparedness" (*Department of Public Safety and Emergency Preparedness Act*, 2005, 2).
>
> "Functions
> 6 (1) In exercising his or her powers and in performing his or her duties and functions and with due regard to the powers conferred on the provinces and territories, the Minister may
> (a) initiate, recommend, coordinate, implement or promote policies, programs or projects relating to public safety and emergency preparedness;
> (b) cooperate with any province, foreign state, international organization or any other entity;
> (c) make grants or contributions; and
> (d) facilitate the sharing of information, where authorized, to promote public safety objectives" (*Department of Public Safety and Emergency Preparedness Act*, 2005, 2).

In consideration to the above, it appears as though any collection that is required by Public Safety and Emergency Preparedness is likely coordinated through the applicable government agencies mentioned within the section 5 of the Act. As such, in order to prevent duplication of work, this section will defer to those relevant Acts for analysis, which in this case will occur in sections 5.6.6, 5.6.7, and 5.6.8 respectively.

5.6.5. Global Affairs Canada (Department of Foreign Affairs, Trade and Development)

Tracking Number A-2022-02412

On 1 May 2023, the Access to Information and Privacy Protection Division of Global Affairs Canada (GAC) had indicated that it was awaiting a response from the Office of primary interest, and

would process the request once it had been received. As of 16 May 2023, there has been no additional information on this request.

Current to 20 April 2023, the *Department of Foreign Affairs, Trade and Development Act*, 2013 does not contain any specific details on information collection directly related to publicly available information. However, section 10(2) outlines the powers, duties, and functions of the Minister under this Act. Specifically relevant is section 10(2)(f), which indicates the following:

> "(2) In exercising and performing his or her powers, duties and functions under this Act, the Minister is to
>
> (f) foster sustainable international development and poverty reduction in developing countries and provide humanitarian assistance during crises" (*Department of Foreign Affairs, Trade and Development Act*, 2013, 3).

In particular, subsection (f)'s mention of providing humanitarian assistance during crises can be read to imply a variety in types of assistance, possibly including the use of OSINT investigative techniques. For example, in humanitarian response efforts, OSINT can be used to monitor social media for signs of unrest, assess the impacts of an incident, track movement (e.g. hostages, refugees), and relay or verify information, thereby mitigating against any additional potential damages.

Separate, but relevant to this legislation is Section 16 of the *Canadian Security Intelligence Services Act,* 1985, which states the following:

> 16 (1) Subject to this section, the Service may, in relation to the defence of Canada or the conduct of the international affairs of Canada, assist the Minister of National Defence or the Minister of Foreign Affairs, within Canada, in the collection of information or intelligence relating to the capabilities, intentions or activities of
>
> (a) any foreign state or group of foreign states; or
>
> (b) any person other than
>
> (i) a Canadian citizen,
>
> (ii) a permanent resident within the meaning of subsection 2(1) of the Immigration and Refugee Protection Act, or
>
> (iii) a corporation incorporated by or under an Act of Parliament or of the legislature of a province (*Canadian Security Intelligence Services Act*, 1985, 26).

Therefore, on the basis of Section 2(f) of the *Department of Foreign Affairs, Trade and Development Act*, 2013, and in combination with Section 16(1) of *the Canadian Security Intelligence Services Act*, 1985, the Minister of Foreign Affairs, Trade and Development may request assistance from CSIS to

collect information on persons (with the above exclusions) or foreign state(s) so long as it has a nexus to the defence of Canada or the conduct of Canada's international affairs.

While this places the majority of responsibility on CSIS, GAC (the Department of Foreign Affairs, Trade and Development) may maintain some ability to utilise publicly available information through Section 2(f) of its Act, so long as the Federal Court is in agreement to this interpretation. The lack of presence to any mention of publicly available information within the Act is therefore not so much an indication that the department is unable to conduct open source investigations, but is more suggestive of a lack of detail in policies and procedures that govern the department's collection capability. In consideration to its mandate relating to the conduct of the external affairs of Canada, this deficiency can have profound impacts on the department's ability to effectively fulfil its mandate, especially if it must rely on the availability of personnel in departments external to its own.

5.6.6. Canadian Border Services Agency

Tracking Number A-2022-25958

On 5 May 2023, the Canadian Border Service Agency (CBSA) Access to Information and Privacy clerk had indicated that they were experiencing a surge in the volume of requests, and were therefore unable to meet the legislated timeline to respond. Although they indicated that they would process the request as soon as possible, they could not specify an estimated completion date.

Similar to the *Department of Foreign Affairs, Trade and Development Act*, 2013, the *Canada Border Services Act*, 2005 does not contain any specific details on information collection directly related to publicly available information. However, sections of the *Canada Border Services Act, 2005* which may allow such collection include sections 12(1), 12(2), and 13(2), which read as follows:

> "Powers of the Agency
> Exercise of powers conferred on Minister
> 12 (1) Subject to any direction given by the Minister, the Agency may exercise the powers, and shall perform the duties and functions, that relate to the program legislation and that are conferred on, or delegated, assigned or transferred to, the Minister under any Act or regulation" (*Canada Border Services Agency Act*, 2005, 5)

> "Officers and employees
> 12 (2) An officer or employee of the Agency may exercise any power or perform any duty or function referred to in subsection (1) if the officer or employee is appointed to serve in the Agency in a capacity appropriate to the exercise of the power or the performance of the duty or function, and, in so doing, shall comply with any general or special direction given by the Minister" (*Canada Border Services Agency Act*, 2005, 6).
>
> "Arrangements and agreements
> 13 (2) The Agency may, for the purposes of carrying out its mandate,
> (a) enter into an arrangement with a foreign state or an international organization; or
> (b) enter into an agreement or arrangement with the government of a province, a department or agency of the Government of Canada or any person or organization" (*Canada Border Services Agency Act*, 2005, 6).

As depicted above, the language of the act is rather vague, especially concerning the use of "...any power any duty or function referred to in subsection (1)" (*Canada Border Services Agency Act*, 2005, 6). While there is no mention of information collection or the use of publicly available information throughout the Act, the sections discussed above imply that such actions may occur so long as special direction is provided by the Minister. Although it is possible that the Minister may have provided a directive on the use of publicly available information relating to the CBSA's mandate, as of 17 May 2023, no supporting documentation has been received that can confirm its existence. Further to this, no supporting documentation has been received on the topic of OSINT training. As such, an evaluation of the CBSA's OSINT training plan cannot be made at this time.

5.6.7. Royal Canadian Mounted Police

Tracking Number A-2022-10021 / F281, File 2700454

On 3 May 2023, the Access to Information and Privacy Branch of the Royal Canadian Mounted Police (RCMP) stated that they were experiencing a large volume of requests, which had resulted in prolonged delays in processing. At the time, the Branch had indicated that the request was with the Disclosure Review Team, but that a timeframe as to when it could be completed was unavailable due to other requests in queue that might further impact the period in which the request could be completed. In consideration to this, this section will review the *Royal Canadian Mounted Police Act,* 1985 (herein *RCMP Act*), the RCMP Regulations, 2014, and the RCMP Internal Audit of Open Source Information, 2021.

Last amended on 12 July 2019, the *RCMP Act*, 1985 makes no specific mention of collection pertaining to the use of publicly available information. However, as a law enforcement agency, the RCMP does maintain a broad authority to collect information. As illustrated in section 18 of the *RCMP Act*, 1985, the duties and obligations of the RCMP include the preservation of the peace, the prevention of offences and criminal acts in violation of Canadian laws, through to the apprehension of criminals and offenders. Section 18 reads as follows:

> "Duties Obligations
> 18 It is the duty of members who are peace officers, subject to the orders of the Commissioner,
>
> (a) to perform all duties that are assigned to peace officers in relation to the preservation of the peace, the prevention of crime and of offences against the laws of Canada and the laws in force in any province in which they may be employed, and the apprehension of criminals and offenders and others who may be lawfully taken into custody;
> (b) to execute all warrants, and perform all duties and services in relation thereto, that may, under this Act or the laws of Canada or the laws in force in any province, be lawfully executed and performed by peace officers;
> (c) to perform all duties that may be lawfully performed by peace officers in relation to the escort and conveyance of convicts and other persons in custody to or from any courts, places of punishment or confinement, asylums or other places; and
> (d) to perform such other duties and functions as are prescribed by the Governor in Council or the Commissioner" (*Royal Canadian Mounted Police Act*, 1985, 10).

In order to fulfil the duties and obligations as depicted in section 18, the RCMP must collect information from a diversified range of sources. According to the Public Order Emergency Commission Institutional Report - Royal Canadian Mounted Police (2023), these sources may include, but not be limited to "police observations, statements from third parties (such as victims, witnesses or suspects), confidential informants, publicly available sources (social media, etc.), other law enforcement agencies, and judicial authorized searches and orders" (23). Section 18 of the *RCMP Act*, 1985 is furthermore expanded upon in section 14(1)(a) of the Royal Canadian Mounted Police Regulations, 2014 (an Annex to the *RCMP Act*, 1985), which states the following:

> "Duties
> 14 (1) In addition to the duties set out in the Act, it is the duty of members who are peace officers to
>
> (a) enforce all Acts of Parliament and regulations and render assistance to departments of the Government of Canada as the Minister directs" (Royal Canadian Mounted Police Regulations, 2014, 5).

The RCMP Regulations, 2014 therefore sets out the authority of the Force[129] to be a derivative of both common law and statutory law in relation to the collection, use, and disclosure of applicable information for the preservation of peace, protection of life, and for criminal investigative purposes (Public Order Emergency Commission, 2023). In this regard, these two sections of the *RCMP Act*, 1985, and the RCMP Regulations, 2014, legitimize the collection of PAI in accordance with the duties and obligations of the RCMP. Although further information on the collection and use of PAI was not available within the Act or Regulations, an internal audit of the RCMP's OSINT activities completed in 2021 referenced an existing policy document - O.M. (Operational Manual) 26.5, "Using the Internet for Open Source Intelligence and Criminal Investigations." However, at the time of writing, the manual was not accessible via the surface web, and so its specific contents cannot be independently evaluated here. Although it is reasonable to assume the manual - whether in whole or in part - would be included in the ATIP request, as of 19 May 2023, a response has yet to be received. Despite this, the RCMP's internal audit provides key information that satisfies the requirement of this section.

The objective of the RCMP's internal audit was, "to determine whether internet-related open source activities conducted across the organization were consistent and compliant with policy" (Royal Canadian Mounted Police, 2021, 6). As a means of doing so, the audit examined open source investigations that supported national and divisional levels of criminal investigations and intelligence gathering between 1 April 2018 to 31 March 2019 (Royal Canadian Mounted Police, 2021). The overall findings of the internal audit were indicative of a lack of awareness, training, and compliance on the use of publicly available information in open source investigations. A portion of the executive summary which highlighted some of the critical findings of the audit has been included here for clarity.

> "Overall, the audit determined that internet-related open source activities conducted across the organization were not consistent nor compliant with OM 26.5. The audit found that opportunities exist to develop a more robust governance framework and enhance national and divisional oversight of open source activities. Without clear roles and responsibilities, and adequate monitoring and oversight, the Force could be exposed to liability and reputational

[129] The Force, in reference to the RCMP.

risk, or criminal prosecutions could be jeopardized and the resulting case law could restrict the future use of open source.

There was limited consultation for the development of OM 26.5 in 2015, and none for the policy update in 2019. The audit found that many employees were not aware that an open source policy existed or that it was applicable to the open source activities that they performed. In addition, OM 26.5 does not provide adequate information to guide users on how to capture, store and retain OSI.[130] There is an opportunity to more broadly communicate OM 26.5 and educate employees on when it applies to their work. Supplementary guidance should be developed and shared across the Force, and should include: how to capture, store and retain OSI, and what constitutes acceptable Tier 2 infrastructure.[131]

Finally, training and information sharing are valuable investments that would ensure that OSI users have access to current and up-to-date information, including relevant case law. Providing introductory-level training to RCMP employees conducting open source activities at the Tier 2 level would facilitate information sharing across the organization" (Royal Canadian Mounted Police, 2021, 3-4).

As depicted above, the findings of the internal audit were quite bold in identifying multiple deficiencies throughout the open source intelligence capability. Despite this, the RCMP's Federal Policing (FP), Specialized Policing Services (SPS), and Contract and Indigenous Policing (C&IP) forces were all in agreement with the findings and recommendations of the audit, recognizing that formal training, support, and governance of this capability was critical to the Force's function (Royal Canadian Mounted Police, 2021). At the time of writing, this was the only publicly disclosed internal audit of a government agency's use of open source information. In consideration to other government departments and their current position in reference to the use of publicly available information for open source investigations - especially those mentioned in this section - it is highly likely that they too have dealt with similar issues, and would moreover benefit from the recommendations of this report. The recommendations of the report are as follows:

"Recommendations

1. The D/Commrs. FP,[132] SPS[133] and C&IP[134] should collaborate to determine how best to fulfill the oversight function for open source activities, recognizing that open source is used by multiple business lines across the Force.
2. Following the decision relating to recommendation #1, the identified business lead(s) should ensure that:

[130] Open source information.
[131] See Chapter 2.2 for a breakdown of the three tier framework in use by the RCMP.
[132] Federal Policing.
[133] Specialized Policing Services.
[134] Contract and Indigenous Policing.

a. National oversight mechanisms are in place to improve visibility over open source activities being conducted in support of investigations and intelligence-gathering activities.
b. The national policy is updated to reflect agreed-upon accountabilities.
c. Supplementary guidance is developed and published related to appropriate capture of OSI, recommended tools and best practices, relevant case law summaries, and advice on court testimony for open source practices.
d. Supplementary guidance is developed and published on the information management requirements for the storage and retention of OSI and OSINT. This guidance should clarify when OSI and OSINT would be considered a record of business value.
e. A communication strategy for OM 26.5 is developed and implemented to educate employees Force-wide about requirements for open source activities and when the policy would be relevant to their duties. This may include the further development and completion of the OSI introductory AGORA[135] course and consideration by the National Mandatory Training and Oversight Committee for making it mandatory for all Regular Members once it is available, as they are the main category of employee unknowingly conducting open source activities at the Tier 2 level.

3. The D/Commr. SPS, in collaboration with D/Commrs. FP and C&IP should consider the feasibility of developing an enterprise-wide solution to address the current gap in Tier 2 infrastructure. In the interim, guidance for technical requirements and/or any mitigating measures (e.g., virtual private network) for Tier 2 infrastructure should be developed and disseminated to all Tier 2 open source users" (Royal Canadian Mounted Police, 2021, 20-21).

Overall, it is unknown how the RCMP intend to follow through on the findings and recommendations of this audit, or within what timeframe it will occur. Regardless of this, the orchestration of the internal audit of its use of open source information was at least indicative of a step forward in the development of a more robust policy, infrastructure for the RCMP to utilise publicly available information. If the RCMP or other government agencies intend to continue to use publicly available information, internal audits similar to this should be conducted regularly. Acknowledging that this could be a time consuming task, it may be more realistic to internally establish teams whose sole purpose is to conduct audits on the use of capabilities that can have significant impact on the duties, obligations, and risks associated with the mandate of the respectful agency.

[135] No further information is available on AGORA. In consideration to the language in the document, AGORA is likely a learning platform hosting online courses. However, this cannot be confirmed at this time.

5.6.8. Canadian Security Intelligence Service
File 117-2022-802

On 19 December 2022, the Canadian Security Intelligence Service (CSIS)'s access to information and privacy coordinator had requested an extension of the 30-day time limit to complete the request by an additional 180 days. The rationalization provided by the coordinator had indicated that the original statutory time limit would unreasonably interfere with the operations of the Service. On 16 June, CSIS had responded to the ATIP request, stating that all the information requested had been exempted from disclosure by virtue of one or more sections of 15(1) (as it relates to the efforts of Canada towards detecting, preventing or suppressing subversive or hostile activities, 16(1)(c), 19(1) or 24(1) of the Act. As a result of this, this section will not include any evaluation of the Service's training or policy documentation. Instead, it will examine the *Canadian Security Intelligence Service Act*, 1985.

Last amended on 13 July 2019, *the Canadian Security Intelligence Service Act*, 1985 contains numerous references to the collection, use, retention, dissemination, and destruction of publicly available information regarding Canadian citizens, permanent residents, and foreign datasets. Additionally, the Act also clarifies prohibited conduct of the Service. However, similar to a number of Acts discussed above, this Act does not define what constitutes publicly available information.

As a comparison, the following page includes several portions of the *Canadian Security Intelligence Service Act*, 1985 (herein, *CSIS Ac*t). This comparison serves to show the inherent differences in the presence of legislation pertaining specifically to the collection and limitations of publicly available information as compared to other Acts discussed in this section.

> "Collection of datasets
> 11.05 (1) Subject to subsection (2), the Service may collect a dataset if it is satisfied that the dataset is relevant to the performance of its duties and functions under sections 12 to 16" (*Canadian Security Intelligence Service Act*, 1985, 10).
>
> "Limit
> 11.05 (2) The Service may collect a dataset only if it reasonably believes that the dataset
> (a) is a publicly available dataset;
> (b) belongs to an approved class; or
> (c) predominantly relates to non-Canadians who are outside Canada" (*Canadian Security Intelligence Service Act*, 1985, 10-11).

“End of evaluation period - Canadian datasets
11.09 (1) If a designated employee confirms that a dataset is a Canadian dataset, the Service shall make an application for judicial authorization under section 11.13, as soon as feasible but no later than the 90th day referred to in subsection 11.07(1)” (*Canadian Security Intelligence Service Act*, 1985, 13).

“Dataset publicly available
11.11 (1) For the purposes of sections 12 to 16, the Service may retain, query and exploit a publicly available dataset” (*Canadian Security Intelligence Service Act*, 1985, 14).

“Record keeping - publicly available datasets
11.24 (1) The Service shall, with respect to publicly available datasets,

(a) establish record keeping requirements for those datasets with respect to the rationale for their collection, the details of each exploitation, the statutory provision under which the result of a query or exploitation is retained and the results that were retained; and

(b) verify, periodically and on a random basis, if the results obtained from the querying and exploitation of those datasets were retained in accordance with subsection 11.11(2)” (*Canadian Security Intelligence Service Act*, 1985, 22).

“Collection, analysis and retention
12 (1) The Service shall collect, by investigation or otherwise, to the extent that it is strictly necessary, and analyse and retain information and intelligence respecting activities that may on reasonable grounds be suspected of constituting threats to the security of Canada and, in relation thereto, shall report to and advise the Government of Canada” (*Canadian Security Intelligence Service Act*, 1985, 23).

“No territorial limit
12 (2) For greater certainty, the Service may perform its duties and functions under subsection (1) within or outside Canada” (*Canadian Security Intelligence Service Act*, 1985, 23).

“Canadian Charter of Rights and Freedoms
12.1 (3.1) The Canadian Charter of Rights and Freedoms is part of the supreme law of Canada and all measures taken by the Service under subsection (1) shall comply with it” (*Canadian Security Intelligence Service Act*, 1985, 24).

“Warrant - Canadian Charter of Rights and Freedoms
12.1 (3.2) The Service may take measures under subsection (1) that would limit a right or freedom guaranteed by the Canadian Charter of Rights and Freedoms only if a judge, on an application made under section 21.1, issues a warrant authorizing the taking of those measures” (*Canadian Security Intelligence Service Act*, 1985, 24).

“Prohibited conduct
12.2 (1) In taking measures to reduce a threat to the security of Canada, the Service shall not

(a) cause, intentionally or by criminal negligence, death or bodily harm to an individual;

(b) wilfully attempt in any manner to obstruct, pervert or defeat the course of justice;

(c) violate the sexual integrity of an individual;

> (d) subject an individual to torture or cruel, inhuman or degrading treatment or punishment, within the meaning of the Convention Against Torture;
> (e) detain an individual; or
> (f) cause the loss of, or any serious damage to, any property if doing so would endanger the safety of an individual" (*Canadian Security Intelligence Service Act*, 1985, 24-25).

As depicted above, the level of detail within the *CSIS Act*, 1985 varies greatly compared to other Acts, and functional directives. Although not fully included here, the Act provides sufficient detail on how the Service is to conduct itself on the matter of collecting publicly available information. For example, in section 11, the Act outlines the circumstances in which the Service may collect PAI, as well as the limitations in which it may do so. Further to this, the section also outlines circumstances by which datasets may or may not be queried, how to obtain judicial authorization to evaluate datasets containing information on Canadians, where collection may occur, how long and when datasets may be retained, and when they must be destroyed. Legislation on this matter has thus far not been observed in any other Act.

Moving forward, section 12 of the *CSIS Act*, 1985 sheds light on the restrictions of the Act as they relate to Part I of the *Constitution Act*, 1982, the *Canadian Charter of Rights and Freedoms*. While not explicitly included here, the Act also references the applicability of the *Privacy Act*, 1985 as it relates to evaluating and deleting personal information that is not relevant to the performance of the Service's duties and functions, and, activity "that does not directly and immediately relate to activities that represent a threat to the security of Canada" (*Canadian Security Intelligence Service Act*, 1985, 9). Additionally, this section also establishes the actions of which the Service is prohibited from conducting. This is of particular interest since prohibited actions have rarely been defined in Acts thus far. Instead, they have often been implied by policy documents that largely outline actions that are permitted under a particular authority.

Taking the above into account, the *CSIS Act*, 1985 could benefit from several inclusions. For example, a section outlining ethical collection, especially in relation to confidentiality,[136] integrity,[137]

[136] Confidentiality, especially in relation to access or restrictions to collected information.
[137] Integrity, in relation to the process of verification that the information collected is accurate.

accountability,[138] and respect.[139] Further to this, the Act would also benefit greatly from defining what is meant by the terms publicly available information, personal identifiable information and what constitutes Canadian citizen (CANCIT) information. A lack of an agreeable definition on these terms can potentially cause issues with evidence used in court. For example, does the Service's interpretation of PAI include breach data? Considering that such data is generally made available as a result of a data breach, would this evidence be inadmissible in court? In what circumstances would this interfere with the *Privacy Act*? What is meant by personal identifiable information? Does this pertain to information that can be directly or indirectly connected to an individual, or both? What constitutes CANCIT information? How should data or information be evaluated to verify its origin? Is a social media profile with a Canadian phone number listed on its page constitute CANCIT information? What if that profile belongs to a foreign national?[140] While fairly simple questions, the answers can manifest complex responses that must take into account a number of considerations that can, overall, complicate an analyst's collection, evaluation, analysis, and dissemination process. These considerations, moreover, must be derived from how key terms are defined.

Overall, the *CSIS Act*, 1985 does shed light on important information that is highly relevant to the use of PAI in OSINT investigations. Although it can benefit from a number of amendments, the *Act* does enable the conduct of open source investigations to occur so long as they are in line with the mandate of the Service. At this time, it is unclear how CSIS trains its analysts to carry out OSINT investigations, or if the Service has any additional policy documents that further outline the use of PAI within OSINT investigations. However, it is clear that with ongoing and future changes in legislation (e.g. the Digital Charter) and new and/or improved technologies (e.g. AI), the Service - alongside other government agencies - must be ready to adapt its policy as its relevance changes with time.

---

[138] Accountability, in that the collected data or information will only be accessed by authorized individuals, thereby respecting the chain of custody of all evidence

[139] Respect in regards to an individual's well being, dignity, and personal agency.

[140] Custom phone numbers (or PYONs; pick your own number) are available through a number of voice over Internet Protocol (VoIP) providers such as OomaOffice, Vonage, Zoom Phone, or RingCentral. In these circumstances, individuals can select their own phone number to appear to be originating from a specific location.

5.7. Discussion and Concluding Chapter Remarks

This section will take each of the case studies into consideration in order to elaborate on the discussions that have already occurred within the individual results of the case studies. More specifically, it will explore interpretations, in respect to what the results mean; implications, in regards to why the results matter; limitations, in consideration to what the results cannot tell us; and recommendations, in terms of what practical actions or studies should follow thereafter.

5.7.1. Summary of Key Findings

To review, this work was oriented to the research question, "In what regards is OSINT an effective tool for emergency management, especially in relation to public safety?" In support of this, the four objectives of this work are indicated below.

> Objective 1: To examine OSINT from a national security perspective.
> Objective 2: To identify potential challenges and barriers that limits an analyst's use of OSINT techniques.
> Objective 3: To explore the changing nature of threats to national security, and identify how OSINT may provide a direct means of assisting with mitigation, prevention, preparation, response, and recovery tactics, techniques, and procedures (TTPs).
> Objective 4: To understand how analysts are training in OSINT collection and methodologies.

On account of this, the results indicate that overall, OSINT can be regarded as an effective tool in maintaining public safety. However, this is not a statement that can be taken without consideration to the accessories that must accompany an OSINT capability. In other words, in order for OSINT to remain an effective tool, the deficiencies identified in the results must be addressed. Although the case studies have examined various tools and how they can be used to collect relevant information that can be used to find people and locations - in addition to a means of verifying images and information - it also revealed the complexities and necessary considerations required in deploying OSINT. These were largely displayed through a lack of consistent and ethical collection policies, training, and infrastructure, which moreover positions a number of government agencies in situations that could result in risk to - among others - reputation, and/or operational and personnel security, which can furthermore potentially restrict access to more technical capabilities. The next four sections will deliver additional discussion on matters relating to this.

5.7.2. Interpretations and Implications

Summary of Interpretations:

- Overall interpretation: OSINT can be regarded as an effective tool in maintaining public safety. However, there are concerns arising from a number of issues, such as:
    1. Legislation & Policy
        a. Legislation, policy, and directives are unclear or absent
        b. Lack of clear oversight
    2. Training & Infrastructure
        a. Adequate training is not clearly outlined
        b. Technical infrastructure to conduct OSINT has not been outlined

Addressing issues relating to these matters will ensure that OSINT is not just an effective tool, but also an ethical one with limited liabilities.

The results in this work indicate that OSINT tools and techniques can be effectively employed by intelligence researchers and analysts in a way that can contribute to public safety. The means by which these tools and techniques can be used to track and/or locate persons, places, or subjects of interest can serve to complement other intelligence disciplines in the interest of public safety, national security, and therefore, the public good. However, the results were also indicative of concerns arising from a lack of policy to enforce training, ethical considerations, collection, use, retention, destruction, and dissemination of information, which can moreover outweigh the benefits of deploying an OSINT capability and manifest a number of liabilities to an organization.

In regards to tools, it is evident that new open source tools are being developed regularly and are shared online via OSINT communities on a variety of social networking sites. Counter to this, OSINT policies and privacy legislation have not been able to keep pace with these developments, which has resulted in a significant policy and legislative delta that has yet to be sufficiently addressed. Similar themes on this matter have been implied by several authors referenced in the literature review, including Cardenas, et al (2013), Pastor-Galindo, et al (2020), and Janjeva, et al (2022).[141] For example, Cardenas, et al (2013) cited the relevancy of privacy as it pertained to the sharing of data between industries and agencies, and the privacy principle of avoiding data reuse. Pastor-Galindo, et al (2020) identified the challenges in using OSINT techniques, especially as they relate to privacy

[141] See Chapter 2.4 and 2.5 of the literature review.

considerations and training associated with data and information extraction. Additionally, Janjeva, et al (2022) emphasized the institutional barriers that acted as deterrents to the progress of realizing an OSINT capability. In comparison with the results, it is evident that these issues remain present in current legislation and directives, which in turn met the expectation of this work. Briefly, the expectation of this work was that OSINT was likely to be observed as an effective capability to support public safety, especially if it were to be used in combination with other intelligence disciplines. But, if legislation were not in place to address issues relating to the collection, use, retention, and dissemination of intelligence derived from OSINT, then the capability would inevitably result in being a liability for government agencies and the people they intend to protect.

Further to the above, the results were indicative of gaps in legislation, policy, and directives associated with utilising an OSINT capability. It is expected, however, that if Bill C-27, *the Digital Charter Implementation Act, 2022* is in fact enacted, it will serve to assist in defining what OSINT policies should entail onwards. Of course, it is important to consider that an alternative explanation as to why it could appear as though there is a lack of key documentation to guide the use of OSINT is that it exists at a higher classification where the release of that information can potentially cause injury to the national interest. If this is in fact the situation, then it would be imperative for each individual agency to conduct an audit (e.g. similar to the RCMP's internal audit) so they too can assess if their OSINT activities are consistent and compliant with policy.

On the matter of implications, the results predominantly imply that the digital environment has and will continue to be a domain of interest relevant to maintaining public safety. As the world becomes ever more connected, this will only continue to increase in magnitude. Thus, it will remain crucial to bear in mind - and weigh appropriately - the principles of security, privacy, and liberty. It is difficult to attest to the idea of there being an absolute solution to this dilemma as technology will continue to evolve, and likely continue to surpass society's efforts to enact legislation to safeguard privacy rights. It is for this reason that ethical training is foundational to OSINT training. Nothing is absolute in theory or practice, but together, theory and practice can inform each other to make ethical decisions and take proportionate and justifiable action.

In closing, the results convey that OSINT as a capability can be beneficial to contributing to public safety. They also identify clear gaps in legislation and training that could result in serious liabilities. To counter this, clear guidelines in policy and legislation paired with an established training regimen can prove advantageous to evading negative predicaments that can place unnecessary risks onto an agency or individual. Moreover, this is significant in relation to the constant evolution of technology, paired alongside individual rights to privacy, liberty, and security. Therefore, it is imperative to address such shortcomings before they are institutionalized as common operating procedures.

5.7.3. Limitations

The reliability of the results in this chapter is impacted by the limitations in access to information. Given the sensitive nature of some collection procedures and the overstrained ATIP request portal, a number of documents were not available for evaluation and analysis. This was either a result of their classification level, security requirements,[142] or because of the time some departments require to make these documents available for public disclosure. Even with a legislated time frame in place to respond to requests, several departments were unable to meet this timespan, suggesting that there are not enough resources allocated to this function, and/or a number of government departments are receiving more than the expected number of requests. Taking this into account, these limitations have had some impact on each of the objectives of this work, but in particular have largely impacted objective four.

> Research Objectives
> Objective 1: To examine OSINT from a public safety (e.g. national and international security) perspective.
> Objective 2: To identify potential challenges and barriers that limit an analyst's use of OSINT techniques.
> Objective 3: To explore the changing nature of threats to national security, and identify how OSINT may provide a direct means of assisting with mitigation, prevention, preparation, response, and recovery tactics, techniques, and procedures (TTPs).
> Objective 4: To understand how analysts are training in OSINT collection and methodologies.

[142] For example, information and operational security: Discussing some policy directives could result in risks to individuals, agencies, and / or missions.

Despite this, the information, legislation and audits obtained served to provide a broad level of evaluation and analysis that functioned to provide an overview of how OSINT is deployed as a tool and technique to complement public safety.

With respect to objective four, the limitations of this work prevented a complete examination of how analysts are trained in OSINT collection. Although this work was able to obtain the necessary documentation from the Department of National Defence, Transport Canada, the Financial Transactions and Reports Analysis Center, and Public Safety Canada, it was unable to obtain specific documentation from Global Affairs Canada, the Canadian Border Service Agency, the Royal Canadian Mounted Police, and the Canadian Security Intelligence Service. However, in the circumstance of the latter agencies, the legislated Acts, regulations, and audits served to provide some perspective on the matter. In particular, evaluation of Global Affairs Canada, the Canadian Border Services Agency, and the Canadian Security Intelligence Service were most impacted by lack of available documentation.[143]

In regards to objectives one, two, and three, limited negative impact occurred here for reasons that could be attributed to the type of information required to satisfy these objectives. Specifically, the information required was found in the cited literature in chapter two, and moreover examined in the chapter five case studies. Thus, the limitations that arose from these objectives were not particularly significant to the overall aim[144] of this research.

Overall, the limitations of this research have been acknowledged, evaluated, and identified here as they relate to and impact the research question, aim, and objectives of this work. In consideration to the limitations referenced above, the results of this work can be noted as valid as they relate to answering the research question[145] based on the sufficient amount of collected information, paired with the evaluation and analysis of that information, in addition to the consideration of any applied biases.

---

[143] The Royal Canadian Mounted Police are not included in this list because their Audit of Open Source Information provided sufficient information to evaluate and analyse.

[144] **Research Aim**: To understand how OSINT may contribute to public safety.

[145] **Research Question**: In what regards is OSINT an effective tool for emergency management, especially in relation to public safety?

5.7.4. Recommendations

Summary of Recommendations:

- Overall recommendations: Address concerns arising from legislation, policy, training, and infrastructure. These include, but are not limited to:
    1. Legislation & Policy
        a. Clarify legislation, policy development, communication, and application;
        b. Establish a clear framework that addresses oversight (e.g. audits), roles, and responsibilities;
    2. Training & Infrastructure
        a. Identify how analysts should be trained when conducting various levels of OSINT, and develop a training plan. Additionally, develop specialty training plans for analysts to become better equipped in conducting OSINT;
        b. Establish technical infrastructure to support:
            i. Collection, Retention, Analysis, and Dissemination of OSINT relating to CANCIT collection and foreign nationals, in isolation of each other;
            ii. Different Tiers of OSINT Collection that address information, operation, and personnel security (e.g. attribution);
            iii. Collaboration between departments to develop best practices;
            iv. An outline of the technical infrastructure to safely conduct OSINT.

The results of this work have identified several gaps in legislation, policy, training, and infrastructure as they relate to OSINT. In response to this, recommendations for practical implementation and further research have been included here with a view to propose a means to address these gaps. For additional clarity, these recommendations are as follows:

Recommendation 1: Legislation & Policy - Clarify Legislation, Policy Development, Communication, and Oversight

In consideration to society's interconnectedness and interdependence on the Internet and World Wide Web, it will be necessary to ensure that legislation and policy reflect the constantly expanding relationship between privacy and the evolution of technologies that require personal data and information. Since this is a fairly complex affair, ongoing reviews of both private and public sector privacy laws will be necessary in order to be in tune with rapidly changing technology. This

may require regular audits of legislation, amendments to legislation, and tribunals to enforce the legislation.

Although Bill C-27, *the Digital Charter Implementation Act*, 2022 will institute some of these recommendations, additional clarity should be provided in the *Privacy Act*, in addition to the Acts which govern the respective departments that collect, use, and retain open source information, and the policy directives those departments issue to their employees. Since the collection of open source data and information has both private[146] and public[147] sector implications and considerations, updating the legislation and policy will necessitate greater transparency and accountability to agencies that collect and utilise OSINT. Recommendations on this matter include, but are not limited to:

- Clearly defining what constitutes publicly available information and personal identifiable information;
- Clarification of different tiers of open source collection, and what is and is not operationally permitted or authorized in accordance with the department's mandate;
- Clearly delineating the difference between Canadian and Foreign National data and information, and the authorities to collect on such individuals and organizations;
- Clarifying how agencies retain data and information on non-criminal incidents (e.g. live monitoring of a situation; collateral collection resulting from network development);
- Clarifying guidelines on the collection, use, and retention of collected data and information;
- Implementing internal and/or external audits to ensure the collection, use, and retention of open source data and information are in line with current legislation, policy, and functional directives;
- Where none exist, creating functional directives that clearly outline the policy and authorities for conducting open source intelligence investigations;
- Clarifying the mitigation, preparation, prevention, response, and recovery methods for handling circumstances where data and/or information may be mishandled, and/or policy and/or authorities were not followed appropriately.

Recommendation 2: Training & Technical Infrastructure

Training and technical infrastructure are paramount to conducting a thorough and ethical OSINT investigation. On this matter, further research is recommended in order to establish a proper training regimen and develop the necessary technical infrastructure that is suited to each department. With respect to this, the following recommendations address gaps in training and technical

[146] For example, obtaining data from private corporations for public use (e.g. agreements with social media platforms, data collection companies, etc.).
[147] For example, an OSINT analyst working for a government department, collecting general information and data online.

infrastructure that impact the ability of government departments and their personnel to conduct OSINT investigations that are in line with Recommendation 1.

- Establish different tier levels of OSINT investigations;
- Establish a mandatory training plan for various levels of OSINT operations (e.g. by tier level);
- Establish a process to evaluate and adjust training as necessary (e.g. Training exercises, professional development, Capture-the-Flag scenarios, missing person searches);
- Clearly assess the impact of attribution to OSINT investigations, and differentiate between the technical infrastructure that should be used when conducting specific tiers of OSINT investigation (e.g. risk assessment);

As a means of capturing these recommendations, Table 5.2 and the information thereafter will outline the necessary details relevant to these suggestions.

Establish Different Tier Levels of OSINT Investigations:

Tier Level	Tier 1 - Overt	Tier 2	Tier 3 - Covert	Tier 4 - Covert
Purpose	Day to Day	Minor operational queries	Operations	Dark Web and Specialized Operations
Authority Level	All personnel	Trained Personnel	Trained Personnel in support of operations	Select trained personnel in support of specific mission requirement
Attribution	Overt - Government IP	Discreet	Covert	Covert
Purpose	Training, and general queries	Exercises, minor operational queries	Operational Support + Targeting	Mission Specific + Specialized Targeting
Training	Basic safety on the Internet	Tier 1 + Verification training, Introductory OSINT Techniques	Tier 1 + Tier 2 + Social Engineering, Advanced OSINT Techniques, Mission Specific Training	Tier 1 + Tier 2 + Tier 3 + training relevant to specific person(s) targeted (e.g. Psy Ops, Info Ops, Tactical Questioning, etc)
Access / Internet Configuration	Regular Internet, VPN for some security	Mandatory VPN	Via Virtual Computer or VPN with Kill Switch	Via Virtual Computer or VPN with Kill Switch
Hardware	Overt Computer / Mobile Device	Computer or Mobile Device with VPN access	Air gapped computer with isolated internet connected (e.g. internet puck)	Virtual Computer or VPN with Kill Switch
Authorization	Section/Dept Head	Section/Dept Head	Ministerial approval, Director/Commanding Officer	PM, Ministerial approval, Director/Commanding Officer
Online Interaction / Simulated Content	None	None	Limited (e.g. games to show activity on social media account)	Yes
Use of Sock Puppet Account	No	No	Yes - for login purposes only. Limited interaction.	Yes
“Log In” Platforms (e.g. social media account)	No	No	Yes	Yes

Table 5.2: Establish Different Tier Levels of OSINT Investigations
Source: Author rendered, with partial input adopted from RCMP 3 Tier Framework (See Royal Canadian Mounted Police, 2021).

OSINT Training Plan Courses - Individual Lead (Commercial Off-The-Shelf Solution)

- Tier 1
    - Legislation, Policy and Ethics
        - Legislation as Appropriate (e.g. PIPEDA, the Privacy Act)
        - Policy and Authorization Documents relevant to specific the Department / Agency / organization
        - Ethical Documentation specific to the Department / Agency / organization
        - Reading: The Berkeley Protocol on Digital Open Source Investigations
            - Available here: https://www.ohchr.org/sites/default/files/2022-04/OHCHR_BerkeleyProtocol.pdf
    - Basic Safety on the Internet
        - MGT433 [LDR433] - Managing Human Risk ($4,950 USD):[148] "Security leaders realize that cybersecurity is no longer just a technical issue but also a human one. Their greatest challenge now has become how to most effectively manage their human risk, as people are involved in over 80% of all breaches. Many organizations attempt to address this by running security awareness programs, but far too often most programs are compliance focused, nothing more than mandatory annual training. As a result, not only is their workforce highly insecure, but most of their workforce has a very negative perception of cybersecurity. This course enables organizations to effectively manage and measure their human risk by changing people's behavior and building a strong security culture" (SANS, 2023).
    - Mental Health and Awareness
        - Crisis & Trauma Resource Institute - Vicarious Trauma: Strategies for Resilience ($129):[149] "Helpers regularly encounter stories and symptoms of trauma in their roles. There is growing evidence that the impact of directly supporting others through experiences of trauma goes beyond burnout or fatigue. The toll of witnessing intense human experiences and emotions can contribute to a negative transformation of a helper's own sense of safety, and of being competent and purposeful. This workshop will provide participants with the opportunity to examine their own experiences and become aware of the signs of both vicarious trauma and vicarious growth. Participants will have the opportunity to develop a personalized plan to repair negative effects as well as accelerate their resilience" (Crisis & Trauma Resource Institute, 2023).

- Tier 2
    - Prerequisite: Tier 1 Training Ethics, privacy and legal considerations when conducting online investigations
    - Legislation, Policy and Ethics
        - Review of documentation from Tier 1
    - Mental Health and Awareness
        - Review from Tier 1. Additional resources as required.

[148] For more information, see: https://www.sans.org/security-awareness-training/career-development/courses/?msc=ssa-career-development

[149] For more information, see: https://ctrinstitute.com/product/on-demand-workshop-vicarious-trauma/

- Verification:
  - Verification Handbook (Free)[150]
    - Separating Rumor from Fact
    - Verification Fundamentals
    - Verifying Images
    - Verifying Video
    - Verification and Fact Checking
    - Verification Tools
    - Creating a Verification Workflow
- OSINT
  - Inteltechniques.net ($649 US/year)[151]
    - OSINT Workflow
    - Capture and Preservation
    - Search Sites and Tools
    - Names, Email, Usernames, and Telephone Numbers
    - Social Media Investigations
    - Photos and Videos
    - Documents, Maps, and Corporations
    - Groups, Communities, and Events
    - Domains and IPs
    - Breaches and Leaks
    - Virtual Machines
    - Command Line and Scripts
    - Darknet and Criminal Markets
    - Security and Privacy
    - Policy
    - Using Virtual Machines and VPNs
- Individual Study
  - Social Network Analysis
    - Key Concepts
    - Common Patterns Amid Ties
    - Common Measures of Individual Actors
    - Common Measures to Describe Networks
- Data Analysis
  - Option 1: Data Camp ($25 USD/month)[152]
    - Data Analysis in Excel
    - Data Analysis in Spreadsheets
    - Introduction to Python
    - Data Analysis in Python
    - Cleaning Data in Python
    - Data Visualization
  - Option 2: WithYouWithMe (Cost Unknown)[153]
    - Data Pathway

[150] For more information, see: https://s3.eu-central-1.amazonaws.com/datajournalismcom/handbooks/Verification-Handbook-1.pdf

[151] For more information, see: https://www.inteltechniques.net/

[152] For more information, see: https://www.datacamp.com/courses-all

[153] For more information, see: https://withyouwithme.com/career-pathways/data-pathway/

- Tier 3
    - Prerequisite: Tier 1 and 2 Training
    - Legislation, Policy and Ethics
        - Review of documentation from Tier 1
    - Mental Health and Awareness
        - Review from Tier 1. Additional resources as required.
    - Verification
        - Review from Tier 2
    - OSINT
        - OSINT Combine: Advanced OSINT ($1,799 USD/year)[154]
            - Monitoring Areas of Interest
            - Region Assessment
            - Image and Video Analysis
            - Data Mining and Artificial Intelligence
            - Advanced Tactics, Techniques, and Procedures (paste sites and forums)
        - OSINT Combine: Mission Specific Training[155]
            - Custom OSINT Course (Instructor Lead, Cost Dependent on Configuration)
                - A tailored approach for specific OSINT projects and operations
                - Utilising a Sock Puppet Account
                - Collecting on Persons and Places of Interest
                - Special Considerations when Conducting OSINT
                - Digital Forensics
                - Introduction to using the Dark Web for OSINT Investigations
        - Optional: SEC 587: Advanced Open-Source Intelligence Gathering and Analysis ($8,275/4 months)[156]
            - Create Sock Puppet Accounts
            - Locating deleted information
            - Understanding Breach Data
            - Triage and Find Meaning in Large Datasets
            - Identifying Malicious Documents
    - Data Analysis
        - Data Camp ($25 USD/month)
            - Intermediate Python
            - Web Scraping in Python
            - Working with APIs
    - Optional Software Tools (Pending Departmental Usage Requirements)
        - Gephi (Free)
            - Data Analysis, Link Analysis, Social Network Analysis (Gephi, 2022).

[154] For more information, see: https://academy.osintcombine.com/p/advanced-open-source-intelligence-course1

[155] For more information, see: https://www.osintcombine.com/custom-osint-course

[156] For more information, see: https://www.sans.org/cyber-security-courses/practical-open-source-intelligence/

        - Palantir Foundry (Cost Dependent on Configuration)
            - E.g. Data integration, Analytics, Time Series Analysis, Audit Logging (Palantir, 2023).
        - IBM i2 Analyst's Notebook
            - Visual Analysis Capabilities: Connections, Patterns, and Key Intelligence (IBM, 2017).
        - NexusXplore (Cost Dependent on Configuration)
            - Collection and Analysis Tool
            - "Find & explore people, places & information across social media, surface, deep & dark web, observe geographical areas for situational awareness, find documents & use artificial intelligence for image analysis & translation to support your operations" (NexusXplore, 2021).

- Tier 4
    - Prerequisite: Tier 1, 2, and 3 Training
    - Legislation, Policy and Ethics
        - Review of documentation from Tier 1
    - Mental Health and Awareness
        - Review from Tier 1. Additional resources as required.
    - Verification
        - Review from Tier 2
    - OSINT
        - OSINT Combine: Illuminating the Dark Web ($499 USD/year)[157]
            - Understanding the TOR network, client, and network attribution
            - Investigating marketplaces, forums, and illicit services
            - Identifying Dark Web communication channels
            - Investigating Other Dark Nets: I2P, Freenet, and Lokinet
            - Tracking Cryptocurrency Transactions and Addresses
    - Cultural and Language Training
        - As defined by operational requirements and team composition
    - Social Engineering
        - SEC 467: Social Engineering for Security Professionals ($3,305 USD/4 months)[158]
            - Perform reconnaissance
            - Create and track phishing campaigns
            - Develop Media Payloads
    - Optional Specialized Training:
        - Information Operations
        - Psychological Operations
        - Questioning Techniques
        - Cyber Operations
    - Note: Additional training will be required, pending purpose and intent of investigation.

[157] For more information, see: https://www.osintcombine.com/darkweb

[158] For more information, see: https://www.sans.org/cyber-security-courses/social-engineering-security-professionals/

OSINT Training Plan Topics - Instructor Lead (Lecture based)

- Tier 1
    - Basic Safety on the Internet
        - Security Awareness:
            - Defining Online Risk and Vulnerability
            - Position/Role-based Risk and Vulnerability
            - Methods of Exploitation and User Vulnerabilities
            - Online Risk Management
            - Enabling, Managing and Maintaining a Culture of Online Security
    - Legislation, Policy and Ethics
        - Legislation as Appropriate (e.g. PIPEDA, the Privacy Act)
        - Policy and Authorization Documents relevant to specific the Department / Agency / organization
        - Ethical Documentation specific to the Department / Agency / organization
        - Reading: The Berkeley Protocol on Digital Open Source Investigations
            - Available here: https://www.ohchr.org/sites/default/files/2022-04/OHCHR_BerkeleyProtocol.pdf
    - Mental Health and Awareness
        - Understanding Vicarious Trauma and Maintaining Resilience:
            - Witnessing Traumatic Experiences
            - Understanding, Identifying, and Coping with Symptoms
            - Develop and Maintain a Resilience Plan

- Tier 2
    - Prerequisite: Tier 1 Training
    - Legislation, Policy and Ethics
        - Review of documentation from Tier 1
    - Verification:
        - Fundamentals of Verification
        - Verification of: User-generated content, images, video, geo-location, data, and information.
        - Verification Workflow
        - Vicarious Trauma
    - Mental Health and Awareness:
        - Review from Tier 1. Additional resources as required.
    - Introduction to OSINT:
        - Fundamentals of Open Source Intelligence
            - What is OSINT?
            - How to Use OSINT
            - OSINT and the Intelligence Cycle
            - Fundamentals of a Search: Boolean, Dorks, Verification, Correlation
            - Understanding Attribution: Operational Security, and Risk Assessments
            - Searching Social Media (No login)
            - Social Network Analysis
    - Special Lecture Topics
        - Social Network Analysis

- Key Concepts
- Common Patterns Amid Ties
- Common Measures of Individual Actors
- Common Measures to Describe Networks

- Tier 3
  - Prerequisite: Tier 1 and 2 Training
  - Legislation, Policy and Ethics
    - Review of documentation from Tier 1
  - Mental Health and Awareness:
    - Review from Tier 1. Additional resources as required.
  - Verification
    - Review of Previous Tier
    - Pending specific objectives: additional verification methods, including facial recognition verification, Fake Account/Bot Analysis and Verification
  - Advanced OSINT Techniques
    - Using an established Sock Puppet Account (no interaction)
      - Pending investigative requirements
    - Searching Relevant Social Media Platforms (Logged in)
      - Person of Interest Collection
    - Special Country/Region Considerations
    - Data Mining
    - Advanced Image and Video Analysis (e.g. Digital Forensics)
    - Artificial Intelligence and OSINT
    - Introduction to the Dark Web and Cryptocurrency
  - Social Engineering
    - Profiling People
    - Defences Against Social Engineering
  - Special Lecture Topics: Data
    - Data Analytics
    - Cleaning Data
    - Using Python in OSINT Investigations
    - Data Visualization
    - Data Analysis with Excel
  - Optional Software Tools (Pending Departmental Usage Requirements)
    - Gephi (Free)
      - Data Analysis, Link Analysis, Social Network Analysis (Gephi, 2022).
    - Palantir Foundry (Cost Dependent on Configuration)
      - E.g. Data integration, Analytics, Time Series Analysis, Audit Logging (Palantir, 2023).
    - IBM i2 Analyst's Notebook
      - Visual Analysis Capabilities: Connections, Patterns, and Key Intelligence (IBM, 2017).
    - NexusXplore (Cost Dependent on Configuration)
      - Collection and Analysis Tool
      - "Find & explore people, places & information across social media,

surface, deep & dark web, observe geographical areas for situational awareness, find documents & use artificial intelligence for image analysis & translation to support your operations" (NexusXplore, 2021).

- Tier 4
    - Prerequisite: Tier 1, 2, and 3 Training
    - Legislation, Policy and Ethics
        - Review of documentation from Tier 1
    - Mental Health and Awareness
        - Review from Tier 1. Additional resources as required.
    - Specific OSINT Training
        - Online Interaction
        - Questioning Techniques
        - Culture Awareness
        - Linguistics: Language, Slang and Age Appropriate Terminology
            - Working with a Language Assistant / Linguist / Interpreter, etc (if relevant)
        - Profile Awareness
        - Attribution
        - Creating or Using an Established Sock Puppet Account
        - Specific Country/Region OSINT Fundamentals
            - Country Specific: Social Media Platforms, Search Platforms, Messaging Applications, et cetera.
    - Social Engineering
        - Review of Previous Lessons
        - Offensive and Defensive Social Engineering
        - Pretexting (Fabricating a Story)
        - Building Rapport
        - Framing, Manipulation, Elicitation, Influence and Persuasion
        - Psychology and Social Engineering
        - Mitigation
    - Note: Additional training will be required, pending purpose and intent of investigation.

Process to Evaluate and Adjust Training as Necessary

The process of evaluating training can be conducted in a variety of ways. Perhaps most effective are exercises and professional development seminars where analysts attend, and at times take the lead on delivering training sessions or scenarios. The information provided below are some examples as to how an approach to evaluating training can be conducted. Depending on the outcome

of the results of each training scenario, training can be adjusted to accommodate the skill level of individuals or analytical teams, as required.

Training Exercise 5W+H

- **Who**: Team A will be delivering a Capture-the-Flag exercise to Team B.
- **What**: Capture-the-flag. **Scenario**: Rebels in Libya have posted photos online regarding the capture of a local town. At this time, the town name is unknown. Verify the information in the read-in package, and determine where the photo was taken. Provide as much detail as possible.
- **Where**: Virtual environment with in-person briefs in Team B's workspace.
- **When**: 1 Day - Thursday.
- **Why**: To exercise core Level 3 OSINT skills, in addition to leadership, teamwork, research, critical thinking, analysis, and briefing skills. To identify where OSINT skills and professional development can be adjusted and addressed.
- **How**: Team A will issue a capture-the-flag scenario to Team B. Team B will then split into 2 teams and individually attempt to solve the scenario. At the end of the day, both respective teams will brief their approach and findings to Team A, and to each other.

Scenario Exercise 5W+H

- **Who**: External OSINT team will be delivering a scenario to Teams A, B, and C who will each have, integral to them, a respective representative of a government department, a stakeholder, and other intelligence disciplines. **Scenario**: Based on the provided read-in package, Team B has been tasked to find the location of a group of individuals linked to a Mexican drug cartel, suspected to have kidnapped 2x Canadians in the vicinity of the US-Mexico border. After Action Review to follow closing remarks of exercise.
- **What**: Virtual and in-person meetings as required.
- **Where**: Virtual environment with in-person briefs in Conference Room A.
- **When**: 1 week (Monday - Friday).
- **Why**: To exercise core Level 3 OSINT skills, inter-governmental networking, leadership, research, analysis, critical thinking, independent workmanship, and briefing skills.
- **How**: Analysts will use OSINT investigative techniques up to and including Level 3 to support their analysis. Requests for information from other intelligence disciplines will be provided if they are within the scope of the exercise.

PD Session 5W+H

- **Who**: Monday - Analyst from Team A; Wednesday - Analyst from Team B; Friday - Analyst from Team C.
- **What**: Tool & Technique Exploration: Google Maps Pro, OSINT Framework, Google Dorks.
- **Where**: Monday - Classroom A; Wednesday - Classroom B; Friday - Classroom C.
- **When**: 3 Days: 2 hours/day; Monday, Wednesday, and Friday.
- **Why**: To assist in the continuous development of junior and senior analysts.
- **How**: Analysts will use OSINT investigative techniques up to and including Level 2 in order to gain familiarity with tools and techniques they may use in their day to day work or on operation. Teams will alternate in delivering and receiving professional development lessons.

Assess the Impact of Attribution to OSINT Investigations

Since some OSINT investigations can be sensitive in nature, it is important to consider the type of infrastructure and policies that should be used to protect the analyst, the institution and the target of interest. Therefore, it is recommended to conduct a risk assessment that considers the safety of these three entities, especially as they relate to attribution. Such a risk assessment could produce some guidance regarding how to manage attribution (e.g. VPN/Virtual Machine), sock puppet account usage, or suggested OSINT Tier level. Additionally, it can also discuss and rate the risk to personnel, institution, and/or target of interest. A simple illustration of how this can be depicted is provided below. Of course, pending the requirements, a risk assessment can be something as simple as the below, or may require additional categories to consider. Further specifications may also be required in order to complete the risk assessment, especially as they relate to rules of engagement, and/or computer and network configuration.

Risk to Personnel	Nil	Low	Medium	High	Severe
Risk to Institution	Nil	Low	Medium	High	Severe
Risk to Target of Interest	Nil	Low	Medium	High	Severe
Suggested VPN Configuration / Virtual Computer	No VPN required	VPN Required - Canadian IP	VPN Required -Recommended Country	VPN Required - Multiple Hops	TOR over VPN
Virtual Computer	Not Required		Required		
Use of Sock Puppet Account	Not Permitted	Permitted - No interaction	Permitted - With Simulated Content	Permitted - With Interaction	Permitted - With Simulated Content and Interaction
OSINT Tier Level	Tier 1	Tier 2	Tier 3	Tier 4	
Additional Training and/or Considerations	Recommended courses or exercises to prepare for a task.				
Assessment	Analyst assessment.				

Table 5.3: Variation of a Simplified Risk Assessment
Source: Author Rendered

Chapter 6: Conclusion

6.1 Overall Argument and Key Takeaways

The overall intent of this work was to determine the regard to which open source intelligence (OSINT) could be an effective tool for emergency management, especially in relation to public safety. In order to do this, a literature review was presented in Chapter 2 that defined what OSINT was, what its current value and capabilities were, what deterrents to progress it faced, and furthermore, the legal ethical, and policy limitations towards OSINT collection and use. While this chapter brought forward a number of discussions, it also identified a lack of existing or refined legal, ethical, and policy documentation regarding the use of publicly available information. Additionally, it also recognized open source intelligence as a relatively new topic of exploration within academia, with most of the peer-reviewed literature dating back approximately fifteen years. In consideration to the literature review and the results, this work identified several key takeaways:

1. OSINT can be effectively used for situational awareness and horizon scanning;
2. OSINT can be effectively used to find a person, group, or location;
3. OSINT can be effectively used to verify information;
4. OSINT, combined with network theory and social network analysis can be used to improve analytical output by serving as information maps to reveal patterns and information that may not otherwise be apparent;
5. There is a lack of documentation pertaining to ethical conduct in OSINT;
6. There is a lack of standardized training available to analysts conducting OSINT;
7. The legislation in place either lacks a definition or is vague on the following terms:
    a. Publicly Available Information (e.g. breach data)
    b. Personal Identifiable Information (e.g. social media moniker)
    c. CANCIT Information (e.g. a Canadian phone number belonging to a foreign national)
    d. Lack of a standardized and trained policy to enforce collection, use, retention, destruction, and dissemination of information

With these key takeaways in mind, this work found OSINT to be an effective tool in maintaining public safety. As the case studies within this work have shown, the use of open source investigative techniques did fulfil the tasks at hand in an effective manner. However, as identified above, there are also a number of factors that may increase risk towards using OSINT for public safety. That is to say, although the use of OSINT may be effective in maintaining public safety, in

some cases it may not be as conducive to maintaining privacy as it could be. Although it is within the purview of analysts to conduct OSINT ethically, it is not their prerogative to create legislation to protect the information of Canadians. That being said, organizations and analysts should also be responsible to safeguard the information collected to mitigate against data spills and breaches. However, on such a matter as privacy, we must defer to the Government of Canada to create legislation that not only safeguards the privacy of Canadians, but also penalizes those who do not abide by such laws - something which the *Digital Charter*, if enacted to law, will seek to do.

In order to ensure that OSINT continues to be an effective tool for public safety, processes must be in place for analysts to receive standardized mandatory training to ensure they do not put themselves or their respective organization at risk of violating legislation or ethical codes of conduct. On a similar note, standardized training will also ensure that analysts are conducting OSINT at the respected level they are trained and authorised to carry out. Section 5.7.4. discusses several recommendations that can assist with this endeavor. Of course, those recommendations are not exhaustive, and each analyst and organization will have varying considerations to take into account.

Moreover, the deficiencies identified on the previous page are especially critical because of the rate at which technology is evolving as a whole. For example, as artificial intelligence is increasingly integrated into smart devices and applications, there needs to be clear legislation in place that discusses how AI may be used to support OSINT. Further to this, analysts must be trained in how to use AI for their investigations in order to ensure the integrity of the information they hold is kept at all times.

6.2 Limitations of the Research

On the matter of limitations of the research, it is important to note that the reliability of the results in this work was somewhat impacted by the limitations in access to information. Although it is difficult to definitively ascertain the impact of these limitations, it is important to make note of them for informational purposes. In relation to the Access to Information and Privacy requests submitted in November 2022, not all the agencies were able to respond to the initial request. Although some agencies were able to respond, portions of the obtained information were redacted or not disclosed as

a result of the information being classified. However, as seen in section 5.6, part of this limitation was mitigated by referring to legislation.

An additional limitation of the research is that it did not use real-time data. Although this would have been beneficial to the research as a whole - especially in regards to conducting OSINT on social media - its use could have risked negative ethical consequences, including collecting and using personal data and / or disclosing personal identifiable information. In order to mitigate this, much of the case studies focused on conducting OSINT investigations with alternative tool sets (e.g. mapping software, image analysis, et cetera).

In consideration to the two main limitations of this work, their impacts on the overall results were varying. The limited responses from the Access to Information and Privacy requests were likely the most impactful as more information could have allowed for a more thorough understanding and assessment of the training regimen of several government agencies and departments. However, where information was made available, an assessment was made that assisted the overall intent of this work, and so its impact was rather minor.

6.3 Recommendations

As a result of those assessments, this work was able to make recommendations to serve as a way forward for employing OSINT to benefit public safety. Although those recommendations have been provided in section 5.7, they will be briefly summarized here in order to reiterate their significance. In short, the overall recommendation was to address the concerns arising from legislation, policy, training, and infrastructure. Specifically, this could include clarifying legislation, policy, and authority. The recommendations also incorporated details on developing a training plan that would identify how analysts could be educated and qualified in conducting OSINT investigations. Finally, another key recommendation was to establish a technical infrastructure that could efficiently support the conduct of OSINT at different tiers. Together, these recommendations could assist in strengthening an agency's use of OSINT, and more importantly, decrease any risks that may manifest as a result of vague interpretations of legislation and / or policy, lack of training and /or inadequate technical infrastructure.

6.4 Contribution to Knowledge

As a whole, this work has enabled the understanding of how open source intelligence may be employed for public safety. It has done so through the act of conducting several case studies that showcased step-by-step actions to arrive at a solution to a given problem, and also by evaluating applications that can be used for situational awareness. Further to this, the Access to Information and Privacy Requests were quite central to this work. By evaluating and comparing each departmental or agency response, and then comparing it to relevant legislation, an evaluation of how publicly available information is collected, retained, used, and disseminated in Canada, was able to be made.

Bearing this, it must be noted that studies on Canadian OSINT capabilities and how they are employed for public safety is fairly limited. Therefore, the case studies conducted within this work provide information that has previously been overlooked and / or understudied. Thus, this work as a whole will serve as a contribution to a field of study that is currently lacking attention, especially within a Canadian context. It is with hope that additional studies such as this occur more frequently so that deficiencies are addressed prior to their evolution into serious problems.

In terms of practical contributions, this work will benefit public and private sector agencies and individuals who employ OSINT in accordance with their particular mandate. If recommendations are considered, there will likely be an improvement to OSINT policies, to the training regimen of OSINT analysts, and to the overall function of the capability, especially as it relates to decreasing the risk conducting OSINT presents to the individual and agency carrying out an investigation. Further to this, it may also assist in addressing concerns relating to privacy, collection, retention, use, and dissemination of OSINT through the establishment of a clear ethical policy document alongside clear legislation.

6.5 The Future of OSINT

We live in a socio-technological domain. In this regard, technology is the basis for how a vast majority of people communicate, share thoughts, and obtain information. As this becomes more apparent, it is evident that the Internet and World Wide Web as a whole will remain foundational to everyday life. Acknowledging this, open source intelligence is likely to increase in use, especially as it remains cost efficient. Several ways in which OSINT will likely continue to grow include: the development of OSINT tools and techniques, the accessibility of professional OSINT education and training, and the availability of open source intelligence consultants. As these fields continue to evolve, it is also likely that an increase in OSINT partnerships and collaborations will be made between public and private agencies (where legally allowed) to share information, datasets, and lessons learned.

As previously alluded to, open source intelligence is a discipline which is constantly evolving to meet the requirements of the analyst and the task at hand. As we move forward into an information / data-centric age, new evolutions of applications, platforms, artificial intelligence, user-generated content, and a participatory social culture will not only challenge OSINT, but in a number of ways it will also redefine how the intelligence discipline as a whole defines sources, methodology, and analysis. However, being a part of such an information / data-centric age also means living amongst troves of misinformation, and disinformation, in addition to the possibility of being targets of information and psychological warfare. So, not only will information verification be especially crucial to any investigation, so too will checking one's biases and critically assessing the source of information, the content, the audience it was directed at, the media to which it appeared on, and the overall effectiveness the information has generated. OSINT as a whole can assist with this process. In fact, we have already witnessed situations in which OSINT has identified staged videos posted to social media (e.g. geolocation was used to prove a video showing “Ukrainian soldiers” harassing a Russian-speaking woman was staged[159]). Alas, where information warfare is present, OSINT can and will assist in countering such efforts.

[159] See https://www.youtube.com/watch?v=3ePOLY-hNT0

In closing, the epitome of any open source intelligence capability is to align real-time information generation with speedy collection, analysis, and dissemination. Although a number of OSINT collaborators have proven their efficiency in this manner, it will be interesting to see how this will continue to transpire in the future, especially with considerations to artificial intelligence. Figures 6.5.1 and 6.5.2 below provide some insight into how AI may fit into and improve the intelligence cycle in relation to open source intelligence.

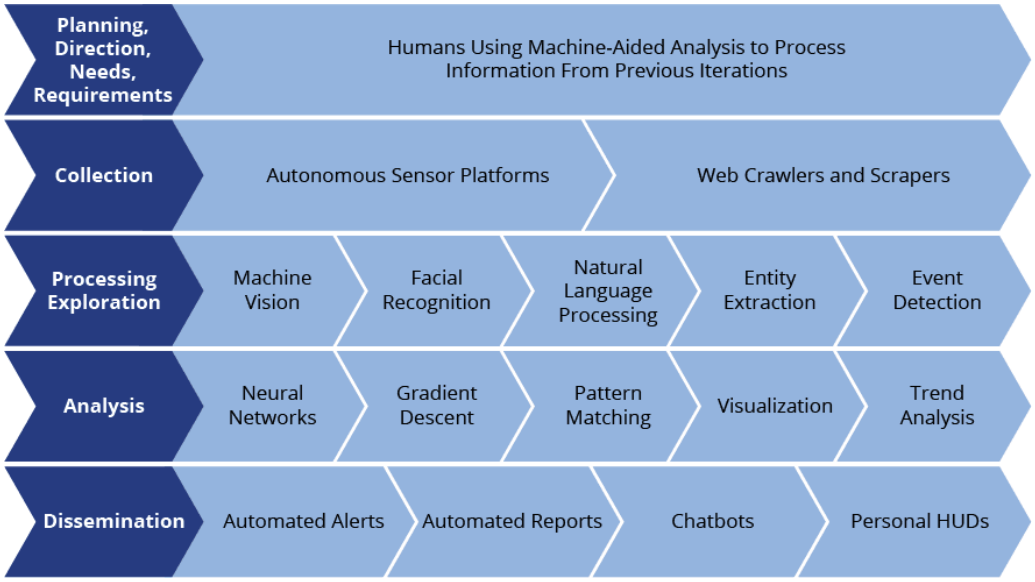

Figure 6.5.1: How AI Fits into the Intelligence Cycle
Source: Recorded Future, 2019

Figure 6.5.2: How AI May Improve the Intelligence Cycle
Source: Recorded Future, 2019

As depicted in Figures 6.5.1 and 6.5.2, there is significant room for artificial intelligence to maintain a role in increasing the efficiency and effectiveness of open source intelligence. Of course, it will still be important to preserve human involvement as the outputs generated from artificial intelligence may be informing real-world operations. Therefore, AI generated outputs need to be reviewed in order to ensure they are valid, otherwise an invalid algorithm could go unchanged and generate data that could be erroneous or altogether meaningless - thereby wasting valuable time in the process. By virtue of the sheer amount of data that is moving through the Internet on a daily basis, using artificial intelligence with open source intelligence is something that will continue to be anticipated. Furthermore, in order for OSINT to remain an employable capability in the future, it will be imperative that both public and private agencies address any potential privacy and ethical concerns. This transparency will assist with gaining public support and / or criticism where necessary. In this way, open source intelligence can and will be a crucial capability to preserving public safety.

References

Access to Information Act, R.S.C., 1985, c. A-1 Retrieved 6 May 2023 from https://laws-lois.justice.gc.ca/eng/acts/a-1/page-1.html#h-227

Adams, S. (2022). *How Government Agencies Utilize OSINT.* Skopenow. Retrieved 18 October 2022 from https://www.skopenow.com/news/osint-for-government#:~:text=OSINT%20enables%20government%20agencies%20to,action%20planning%2C%20and%20tackling%20misinformation.

Akhgar, B., Bayerl P., & Sampson F. (2016). *OSINT as an Integral Part of the National Security Apparatus*. Springer

Akhgar, B., & Wells, D. (2018). Critical Success Factors for OSINT-Driven Situational Awareness. *European Law Enforcement Research Bulletin*, (4 SCE), 67-74. Retrieved from 1 April 2023 https://bulletin.cepol.europa.eu/index.php/bulletin/article/view/332

Ali, F., Khan, F. H., Bashir, S., and Ahmad, U. (2019, March 12). Counter Terrorism on Online Social Networks Using Web Mining Techniques. In Bajwa, I., Kamareddine, F., Costa, A. (eds) *Intelligent Technologies and Applications.* INTAP 2018. Communications in Computer and Information Science, vol 932. Springer Singapore. Pp. 240-250. https://doi.org/10.1007/978-981-13-6052-7_21

Ali, A. (2021, November 26). *From Amazon to Zoom: This is What Happens on the Internet Every Minute*. World Economic Forum. Retrieved 26 November 2022 from https://www.weforum.org/agenda/2021/11/amazon-youtube-zoom-internet-minute-2021/

Amos, D. (2022, June 12). Open source intelligence methods are being used to investigate war crimes in Ukraine. NPR. Retrieved 17 November 2022 from https://www.npr.org/2022/06/12/1104460678/open-source-intelligence-methods-are-being-used-to-investigate-war-crimes-in-ukr

Arena, M. (2016, May 19). *Cyber Threat Intelligence Requirements.* Intel 471. Retrieved 14 November 2022 from https://intel471.com/blog/cyber-threat-intelligence-requirements-what-are-they-what-are-they
-for-and-how-do-they-fit-in-the

Azevedo, R., Medeiros, I., & Bessani, A. (2019, August). PURE: Generating Quality Threat Intelligence by Clustering and Correlating OSINT. 18th IEEE International Conference on Trust, Security, and Privacy in Computing and Communications. *13th IEEE International Conference on Big Data Science and Engineering (TrustCOm/BigDataSE),* 483-490. DOI:10.1109/TrustCom/BigDataSE.2019.00071

Backstrom, L, Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2012, January 5). *Four Degrees of Separation*. Cornell University. Arxiv. Retrieved 3 May 2023 from https://arxiv.org/abs/1111.4570

Baffa, R. (2022). *The Ukraine-Russia War Confirms the Value of OSINT BabelStreet*. Retrieved 3 December 2022 from https://www.babelstreet.com/blog/the-ukraine-russia-war-confirms-the-value-of-osint

Barabasi, A.L. (2003). *Linked: How Everything is Connected to Everything Else and What it means for Business, Science, and Everyday Life*. Plume.

BasuMallick, C. (2022, August 8). *What is Big Data? Definition, Types, Importance, and Best Practices.* Spiceworks. Retrieved 16 January 2023 from https://www.spiceworks.com/tech/big-data/articles/what-is-big-data/

Bazzell, M. (2021). Open Source Intelligence Techniques: Resources for Searching and Analyzing Information. 8th Ed. Amazon: Bolton, On.

Bazzell, M. (2022). Open Source Intelligence Techniques: Resources for Searching and Analyzing Information. 9th Ed. Amazon: Bolton, On.

Bazzell, M. (2023) *IntelTechniques Search Tools*. Breaches & Leaks Search Tool. Retrieved 20 April 2023 from https://www.inteltechniques.com/tools/Breaches.html

Bellingcat. (2022, September). *Bellingcat's Online Investigation Toolkit*. Version 6.9. Retrieved 24 April 2023 from https://docs.google.com/spreadsheets/d/18rtqh8EG2q1xBo2cLNyhIDuK9jrPGwYr9DI2UncoqJQ/edit#gid=930747607

Bellingcat. (nd). *Investigations*. Bellingcat. Retrieved 15 June 2023 from https://www.bellingcat.com/category/news/

Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). Social Big Data: Recent Achievements and New Challenges. *An International Journal on Information Fusion, 28*, 45–59. https://doi.org/10.1016/j.inffus.2015.08.005

Bhuiyan, J. (2021, November 8). *LAPD Ended Predictive Policing Programs Amid Public Outcry. A New Effort Shares Many of their Flaws*. The Guardian. Retrieved 29 January 2023 from https://www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform

Bill C-11, *An Act to Enact the Consumer Privacy Protection Act and the Personal Information and Data Protection Tribunal Act and to Make Consequential and Related Amendments to Other Acts*, First Reading, 2nd Session, 43rd Parliament, 2020. Retrieved 7 February 2023 from https://parl.ca/DocumentViewer/en/43-2/bill/C-11/first-reading

Bill C-11, *An Act to Enact the Consumer Privacy Protection Act and the Personal Information and Data Protection Tribunal Act and to Make Consequential and Related Amendments to Other Acts*, 2nd Session, 43rd Parliament, 2021. Retrieved 7 February 2023 from https://www.parl.ca/LegisInfo/en/bill/43-2/c-11

Blackdot Solutions. (2021). *The OSINT Handbook: How to Use Open Source Data to Transform Investigatory Best Practice*. Retrieved 10 February 2023 from https://blackdotsolutions.com/wp-content/uploads/2021/11/The-OSINT-Handbook_.pdf

Block, L. (2021, July 28). *GDPR Essentials for OSINT Research*. Methodology. Retrieved 1 February 2023 from https://www.blockint.nl/methods/gdpr-essentials-for-osint-research/

Bockholt, M., & Zweig, K.A. (2017, June 12). Paths in Complex Networks. Alhajj, R., Rokne, J. (eds) *Encyclopedia of Social Network Analysis and Mining*. 1-11. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-7163-9_110183-1

Borges, E. (2021, March 29). *Top 20 Google Hacking Techniques*. SecurityTrails Blog. Retrieved 24 April 2023 from https://securitytrails.com/blog/google-hacking-techniques

Boyd, D. (2010). Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. *Networked Self: Identity, Community, and Culture on Social Network Sites*, edited by Zizi Papacharissi, 39–58. New York, NY: Routledge. Retrieved 18 October 2022 from https://www.danah.org/papers/2010/SNSasNetworkedPublics.pdf

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication and Society*, *15*(5), 662---679.

Brown, C. (2023). *How to Conduct an Ethical OSINT Investigation*. Blackdot Solutions. Retrieved 10 February 2023 from https://blackdotsolutions.com/blog/ethics-in-data-collection/#:~:text=OSINT%20Regulations%20and%20the%20Ethics%20of%20Data%20Collection&text=Yet%20making%20use%20of%20it,ethics%20of%20mass%20data%20collection

Buchanan, M. (2002). Small Worlds and the Ground-breaking Science of Networks. Norton.

Burgess, M. (2007). *Euro View: Internet Policing*. World Analysis. Retrieved 18 October 2022 from http://worldanalysis.net/postnuke/html/index.php?name=News&file=article&sid=656

Burns, M. (2020, February 14). *Open Source Intelligence: What Social Media Can Tell Us*. Camera Forensics. Retrieved 19 October 2022 from https://www.cameraforensics.com/blog/2020/02/14/open-source-intelligence-what-social-media-can-tell-us/

Byrne, E. (2022, April 20). *Why Social Media is at the Forefront of Protective Intelligence*. Track 24. Retrieved 12 December 2022 from https://www.track24.com/why-social-media-is-at-the-forefront-of-protective-intelligence/

Calo, R. (2014). Digital Market Manipulation. *The George Washington Law Review* 82 (4): 995–1051. Retrieved 18 October 2022 from www.gwlr.org/wp-content/uploads/2014/10/Calo_82_41.pd*f*

Campbell, L. (2006, May 23). *Restoring FOB XXXX*. Defense Visual Information Distribution Service. Retrieved 22 April 2023 from https://dvidshub.net

Canada Border Services Agency Act, S.C. 2005., c. 38. Justice Laws Website: Government of Canada. Retrieved 8 May 2023 from https://laws-lois.justice.gc.ca/PDF/C-1.4.pdf

Canadian Charter of Rights and Freedoms, s 8, Part 1 of the *Constitution Act*, 1982, being Schedule B to the *Canada Act 1982* (UK), 1982, c 11. Retrieved 9 February 2023 from https://laws-lois.justice.gc.ca/PDF/CONST_TRD.pdf

Canadian Security Intelligence Service Act. R.S.C., 1985, c. C-23. Justice Laws Website. Part II: Judicial Control (continued). Government of Canada. Retrieved 9 February 2023 from https://laws-lois.justice.gc.ca/eng/acts/c-23/page-7.html#h-8

Canadian Security Intelligence Services Act, R.S.C., 1985, c. C.-23. Justice Laws Website: Government of Canada. Retrieved 8 May 2023 from https://laws-lois.justice.gc.ca/PDF/C-23.pdf

Cárdenas, A. A., Manadhata, P. K., & Rajan, S. P. (2013, November-December). Big Data Analytics for Security. *IEEE Security & Privacy*, *11*(6), 74-76, DOI: 10.1109/MSP.2013.138. Retrieved 5 January 2023 from https://ieeexplore-ieee-org.ezproxy.library.yorku.ca/document/6682971

Carrillo, A. B. (2018, February 21). What Is Technological Neutrality? Blog Xnoccio Viavansi. *Blog De Viafirma | Especializado En Firma Electrónica e Identidad Digital*. Retrieved 18 October 2022 from https://www.viafirma.com/blog-xnoccio/en/what-is-technological-neutrality/

Carter, J.A., Maher, S., & Neumann, P.R. (2014). *#Greenbirds: Measuring Importance and Influence in Syrian Foreign Fighter Networks*. International Centre for the Study of Radicalisation and Political Violence (ICSR). Retrieved 30 November 2022 from https://icsr.info/wp-content/uploads/2014/04/ICSR-Report-Greenbirds-Measuring-Importance-and-Infleunce-in-Syrian-Foreign-Fighter-Networks.pdf

Castillo, D. (2021, October 29). *Machine Learning Regression Explained*. Seldon Technologies. Retrieved 20 December 2022 from https://www.seldon.io/machine-learning-regression-explained#:~:text=Machine%20Learning%20Regression%20is%20a,used%20to%20predict%20continuous%20outcomes

Central Intelligence Retirees' Association. (Winter 2007). Remarks by Doug Naquin, Director, Open Source Center. *CIRA Newsletter, 32*(4), 2-9. Retrieved October 2022 from https://irp.fas.org/eprint/naquin.pdf

Chen, T.Y.J. (2019). *Advancing Quantitative Risk Analysis for Critical Water Infrastructure. University of Michigan*. University of Michigan. Retrieved 25 January 2023 from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/153423/tyjchen_1.pdf?sequence=1

Chief of Defence Intelligence. (Approval Authority), Director General Intelligence Policy and Partnerships (Office of Primary Responsibility). (2019). Chief of Defence Intelligence Functional directive: Conduct of Open Source Intelligence Activities on the Internet. 1 November 2019. (Effective Date). Department of National Defence. Canada. Retrieved via Access to Information Request.

Chief of Defence Intelligence. (2020). Briefing Note for the Deputy Minister and Chief of the Defence Staff Open Source Intelligence Activities on the Internet. 6 March 2020. Department of National Defence. Canada. Retrieved via Access to Information Request.

Choudhury, N. (2014). World Wide Web and Its Journey from Web 1.0 to Web 4.0. *International Journal of Computer Science and Information Technologies*, *5*(6), 8096-8100. Retrieved 20 September 2022 from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.666.6445&rep=rep1&type=pdf

Cobwebs Technologies. (2021, March 15). *The Advantages of Open Source Intelligence (OSINT)*. Corporate Security. Retrieved 25 January 2023 from https://cobwebs.com/blog/the-advantages-of-open-source-intelligence-osint/

Crisis & Trauma Resource Institute. (2023). *On Demand Workshops. Vicarious Trauma - Strategies for Resilience*. CTRI. Retrieved 10 June 2023 from https://ctrinstitute.com/product/on-demand-workshop-vicarious-trauma/

Dawson, M., Lieble, M., & Adeboje, A. (2018). Open Source Intelligence: Performing Data Mining and Link Analysis to Track Terrorist Activities. *Information Technology - New Generations*. 558, 1-11. Springer International Publishing.

Defence Research and Development Canada. (2020, May). *Standards for Evaluating Source Reliability and Information Credibility in Intelligence Production*. Retrieved 24 Jan 2023 from https://cradpdf.drdc-rddc.gc.ca/PDFS/unc351/p812555_A1b.pdf

de la Torre-Abaitua, G., Lago-Fernández, L.F., & Arroyo, D. (2019, August 1). *A Compression Based Framework for the Detection of Anomalies in Heterogeneous Data Sources*. Cornell University. Retrieved 18 December 2022 from https://arxiv.org/abs/1908.00417

Dekens, N. (2020, June 8). *Vicarious Trauma*. OSINT Curio.us. Retrieved 26 November 2022 from https://osintcurio.us/2020/06/08/vicarious-trauma-and-osint-a-practical-guide/

Deliu, I., Leichter, C., & Franke, K. (2017, December). Extracting Cyber Threat Intelligence From Hacker Forums: Support Vector Machines Versus COnvolutional Neural Networks. *IEEE International Conference Big Data.* pp. 3648-3656. Retrieved 18 October 2022 from https://ieeexplore.ieee.org/document/8258359

Department of Foreign Affairs, Trade and Development Act, S.C. 2013, c. 33 S. 174. Justice Laws Website: Government of Canada. Retrieved 8 May 2023 from https://laws-lois.justice.gc.ca/PDF/F-27.5.pdf

Department of National Defence. (2001). *Land Force Information Operations Field Manual: Intelligence. B-GL-357-001/FP-001*. Government of Canada. Retrieved 18 October 2022 from https://www.scribd.com/doc/48074701/Land-Force-Information-Operations-Intelligence-Field-Manual

Department of Public Safety and Emergency Preparedness Act, S.C. 2005, c. 10. Justice Laws Website: Government of Canada. Retrieved 8 May 2023 from https://laws-lois.justice.gc.ca/eng/acts/P-31.55/page-1.html#h-401518

Dilipraj, E. (2014, July-September). Terror in the Deep and Dark Web. *Air Power Journal*. 9(3), 120-140. MONSOON 2014. Retrieved 18 December 2022 from http://www.academia.edu/9622433/TERROR_IN_THE_DEEP_AND_DARK_WEB

Dixon, S. (2022a). *Most Popular Social Networks Worldwide as of January 2023, Ranked by Number of Monthly Active Users*. Statista. Retrieved 19 October 2022 from https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Dixon, S. (2022b). *Number of Social Media Users Worldwide from 2017 to 2027*. Statista. Retrieved 19 October 2022 from https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Domo. (2022). *Data Never Sleeps 10.0.* Domo. Retrieved 26 November 2022 from https://www.domo.com/data-never-sleeps#

Dragos, Inc. (2021, January 6). *Self-Reflection Time: The OSINT Collection Risk Framework*. Retrieved 27 January 2023 from https://www.dragos.com/blog/industry-news/self-reflection-time-the-osint-collection-risk-framework/

Edunov, S. Diuk, C.G., Filiz, I.O., Bhagat, S., & Burke, M. (2016, February 4). *Three and a Half Degrees of Separation*. Meta. Retrieved 3 may 2023 from https://research.facebook.com/blog/2016/2/three-and-a-half-degrees-of-separation/

Eilers, C. (2022). *40+ Google Search Hacks, Tricks & Tips for 2023*. 13 December 2022. Retrieved 24 April 2023 from https://zety.com/blog/google-search-hacks

Ellis, H. (2018, October 18). *How to Prevent, Identify and Address Vicarious Trauma — While Conducting Open Source Investigations in the Middle East*. Bellingcat. Retrieved 26 November 2022 from https://www.bellingcat.com/resources/how-tos/2018/10/18/prevent-identify-address-vicarious-trauma-conducting-open-source-investigations-middle-east/

Ellis, H. (2019, February 5). *Open Source Investigations: How to Prevent, Address and Identify Vicarious Trauma*. Global Investigative Journalism Network.Retrieved on 27 January 2023 from https://gijn.org/2019/02/05/open-source-investigations-how-to-prevent-address-and-identify-vicarious-trauma/

Estrada, E., & Knight, P. (2015). *A First Course in Network Theory*. Oxford University Press.

European Commission. (2014). *Technical Policy Briefing Notes - 8: Social Network Analysis*. MEDIATION Project. The PROVIA / MEDIATION Adaptation Platform. Retrieved 11 March 2023 from https://www.pik-potsdam.de/~wrobel/mediation-platform/pbs/pb8/strengths_and_weaknesses.html#:~:text=Subjective%20bias%20and%20can%20be,artificial%20boundaries%20(often%20necessarily).

European Commission. (nd). *Who does the data protection law apply to?* Application of the Regulation. Retrieved 26 November 2022 from https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organizations/application-regulation/who-does-data-protection-law-apply_en#:~:text=The%20GDPR%20applies%20to%3A,the%20data%20is%20processed%3B%20or

Evangelista, J.R.G., Sassi, R.J., Romero, M., & Napolitano, D. (2021). Systematic Literature Review to Investigate the Application of Open Source Intelligence (OSINT) with Artificial Intelligence. *Journal of Applied Security Research*, *16*(3), 345-369, DOI: 10.1080/19361610.2020.1761737. Retrieved 15 October 2022 from https://www.tandfonline.com/doi/abs/10.1080/19361610.2020.1761737

Finklea, K. (2017, March 10). Dark Web. *Congressional Research Service*. R44101. Retrieved 15 December 2022 from https://a51.nl/sites/default/files/pdf/R44101%20(1).pdf

Fisher, D., DeLine, R., Czerwinski, M., & Drucker, S. (2012, May-June). Interactions with Big Data Analytics. *Interactions, 19*(3), 50-59. https://doi.org/10.1145/2168931.2168943 Retrieved 18 October 2022 from https://dl.acm.org/doi/abs/10.1145/2168931.2168943

Fitsanakis, J. (2011, November 8). Analysis: CIA Open Source Center monitors Facebook, Twitter, blogs. Intel News. Retrieved 11 October 2022 from https://intelnews.org/2011/11/08/01-861/#more-7508

Fitzgibbons, L. (2022, November). *What is a Zettabyte?* Tech Target. Retrieved 18 October 2022 from https://www.techtarget.com/searchstorage/definition/zettabyte

Forensically. (2003). *Forensically (Beta)*. Retrieved 1 May 2023 from https://29a.ch/photo-forensics/#forensic-magnifier

Flashpoint. (2021, March 9). *5 Ways Governments Can Apply Open Source Intelligence.* Flashpoint. Retrieved 11 October 2022 from https://flashpoint.io/blog/how-governments-apply-open-source-intelligence/

Freeman, L. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*. Empirical Press.

Gephi. (2022). *The Open Graph Viz Platform*. Gephi Makes Graphs Handy. Retrieved 10 June 2023 from https://gephi.org/

Getting, B. (2007, April 18). *Basic Definitions: Web 1.0, Web. 2.0, Web 3.0* Practical Ecommerce. Retrieved 24 November 2022 from https://www.practicalecommerce.com/Basic-Definitions-Web-1-0-Web-2-0-Web-3-0

Google. (2023). *Google Data Analytics Professional Certificate*. Corsea Website. Retrieved 16 January 2023 from https://gb.coursera.org/professional-certificates/google-data-analytics

Google Alerts. (2023). Alerts. Retrieved 27 April 2023 from https://www.google.com/alerts

Google Earth Pro. (6 March 2015a). IHEC Government Office and Presidential Palace Complex. Mosul, Iraq. 36°23’19.14”N 43°07’31.36”E, Eye Alt 1.44km

Google Earth Pro. (6 March 2015b). IHEC Government Office. Mosul, Iraq. 36°23’25.70” N43°07’29.95”E. Eye Alt 388m.

Google Earth Pro. (22 June 2017). IHEC Government Office. Mosul, Iraq. 36°23’25.70” N43°07’29.95”E. Eye Alt 389m.

Google Earth Pro. (28 August 2015). IHEC Government Office. Mosul, Iraq. 36°23’29.44” N43°07’18.54”E. Eye Alt 331m.

Google News. (2023). Your Briefing. Google News. Retrieved 27 April 2023 from https://news.google.com/

Granovetter, M. (1973, May). The Strength of Weak Ties. *American Journal of Sociology*, 78, 1360-1380. Retrieved 1 April 2023 from https://edisciplinas.usp.br/pluginfile.php/4865899/mod_resource/content/0/The%20Strength%20of%20Weak%20Ties.pdf

Granovetter, M.S. (1982). The Strength of Weak Ties: A Network Theory Revisited. In *Social Structure and Network Analysis*, edited by P. Marsden and N. Linn. Beverly Hills, Calif.: Sage.

Granovetter, M.S. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Social Structure and Network Analysis*, edited by P. Marsden and N. Linn. Beverly Hills, Calif.: Sage.

Greenberg, A. (2015, March 30). *Feds Demand Reddit Identify Users of a Dark-Web Drug Forum*. Wired. Retrieved 18 December 2022 from https://www.wired.com/2015/03/dhs-reddit-dark-web-drug-forum/

Gritzalis, D., & Stavrou, V. (2016, February). Exploiting Open Source Intelligence Capabilities for the Benefit of the Hellenic Air Force. *Information Security & Critical Infrastructure Protection Laboratory.* Department of Informatics, Athens University of Economics & Business. Retrieved 19 October 2022 from https://infosec.aueb.gr/Publications/HAF-2016%20OSINT%20NEREUS.pdf

Hall, J. (2015, June 24). S*ickening New ISIS Video Shows Caged Prisoners Lowered into a Swimming Pool and Drowned, Shot with an RPG and Blown Up with Explosive-filled' Necklaces.'* Daily Mail. Retrieved 23 March 2023 from https://www.dailymail.co.uk/news/article-3135913/Sickening-new-ISIS-video-shows-caged-prisoners-lowered-swimming-pool-drowned-shot-RPG-blown-explosive-filled-necklaces.html

Hao, D. & Li, C. (2011. December). The Dichotomy in Degree Correlation of Biological Networks. *PloS One*. *6*(12), 1-13, e28322. 10.1371/journal.pone.0028322. Retrieved 8 March 2023 from https://www.researchgate.net/publication/51873746_The_Dichotomy_in_Degree_Correlation_of_Biological_Networks

Harris, S. (2020, November 1). *Ten Minute Tip: Image Geolocation – Part 1*. OSINTCurio.us. Retrieved 18 October 2022 from https://osintcurio.us/2020/11/01/ten-minute-tip-image-geolocation-part-1/

Hassan, N. (2018, August 12). *An Introduction to Open Source Intelligence (OSINT) Gathering*. Secjuice. Retrieved 18 October 2022 from https://www.secjuice.com/introduction-to-open-source-intelligence-osint/

Health Canada. (2008, August). *Privacy: A Fundamental Right in Canada*. Retrieved 7 February 2023 from https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/occup-travail/privacy-protection/privacy-protection-eng.pdf

Hewson, C., & Buchanan, T. (2013). *Ethics Guidelines for Internet---Mediated Research*. The British Psychological Society, Leicester. Retrieved 15 February 2023 from http://oro.open.ac.uk/42132/

Hewson, C., Buchanan, T., Kaye, L., Coulson, N., Branley-Bell, D., Fullwood, C., & Devlin, L. (2021, June 7). *Ethics Guidelines for Internet---Mediated Research*. British Psychological Association Retrieved 15 February 2023 from https://www.bps.org.uk/guideline/ethics-guidelines-internet-mediated-research

Higgins, E. (2023, March 29). *How Online Investigators Proved Video of Ukrainian Soldiers Harassing Woman was Staged*. Bellingcat. Retrieved 29 April 2023 from https://www.bellingcat.com/news/2023/03/29/how-online-investigators-proved-video-of-ukrainian-soldiers-harassing-woman-was-staged/

Hunchly. (2023). *OSINT Software for Cybersecurity, Law Enforcement, Journalists, Private Investigators and More*. https://www.hunch.ly/

IBM. (2017). IBM i2 Analyst's Notebook: Discover and Deliver Actionable Intelligence. *IBM Security Solution Brief*. Retrieved 10 June 2023 from https://www.ibm.com/downloads/cas/QNGO6RNA

IBM. (2023, March 24). *ASCII Characters*. 24 March 2023. Retrieved 1 May 2023 from https://www.ibm.com/docs/en/aix/7.1?topic=support-ascii-characters

Image Engineering. (2023). *Noise: Image Quality Factors*. Image Quality Factors. Retrieved 1 May 2023 from https://www.image-engineering.de/library/image-quality/factors/1080-noise

Innovation, Science and Economic Development Canada. (2022a, June 16). *New Laws to Strengthen Canadians' Privacy Protection and Trust in the Digital Economy*. Government of Canada. Retrieved 1 February 2023 from https://www.canada.ca/en/innovation-science-economic-development/news/2022/06/new-laws-to-strengthen-canadians-privacy-protection-and-trust-in-the-digital-economy.html

Innovation, Science and Economic Development Canada. (2022b, March 13). *Canada's Digital Charter: Trust in a Digital World*. Government of Canada. Retrieved 1 February 2023 from https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter-trust-digital-world

Innovation, Science and Economic Development Canada. (2022c, August 18). *Bill C-27 Summary: Digital Charter Implementation Act, 2022*. Government of Canada. Retrieved 1 February 2023 from https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter/bill-summary-digital-charter-implementation-act-2020

Irwin, D., & Mandel, D. (2020, December). Standards for Evaluating Source Reliability and Information Credibility in Intelligence Production. *Defence Research and Development Canada - Toronto Research Centre*. STO-TR-SAS-114, 7-1 - 7-16. Retrieved 18 October 2022 from https://cradpdf.drdc-rddc.gc.ca/PDFS/unc351/p812555_A1b.pdf

Jang-Jaccard, J., & Nepal, S. (2014). A Survey of Emerging Threats in Cybersecurity. *Journal of Computer and System Sciences*. *80*(5), 973-993. DOI:10.1016/j.jcss.2014.02.005

Janjeva, A., Harris, A., & Byrne, J. (2022, June 7). The Future of Open Source Intelligence for UK National Security. *Royal United Services Institute for Defence and Security Studies*. Retrieved 18 October 2022 from https://rusi.org/explore-our-research/publications/occasional-papers/future-open-source-intelligence-uk-national-security

Joint Chiefs of Staff. (2014, November 20). *Joint Publication 3-13. Information Operations*. Incorporating Change 1. U.S. Army Joint Staff. Retrieved 27 January 2023 from https://irp.fas.org/doddir/dod/jp3_13.pdf

Jolliffe, I. T., & Cadima, J. (2016, April 13). Principal Component Analysis: A Review and Recent Developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Kadushin, C. (2004, February 17). Introduction to Social Network Theory. Retrieved 1 March 2023 from http://melander335.wdfiles.com/local--files/reading-history/kadushin.pdf

Kadushin, C. (2012). *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford University Press.

Kamel Boulos, M.N., & Wheeler, S. (2007, March 24). The Emerging Web 2.0 Social Software: An Enabling Suite of Sociable Technologies in Health and Health Care Education. *Health Libraries Group.Health Information and Libraries Journal, 24,* 2-23. DOI: 10.1111/j.1471-1842.2007.00701.x. Retrieved 24 November 2022 from https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1471-1842.2007.00701.x

Kenausis, L. (2021, October). *Talking with Journalists and Analysts about the Ethics of Open Source*. Stanley Center for Peace and Security. Ethical Journalism Network. Retrieved 10 February 2023 from https://stanleycenter.org/publications/open-source-ethics-journalists-analysts/

Khanzode, K.C., & Sarode, R.D. (2016, April-June). Evolution of the World Wide Web: From Web 1.0 to 6.0. *International Journal of Digital Library Services*. 6(2), 1-11. https://www.edtech1.com/Evolution%20of%20the%20World%20Wide%20Web,%20From%201.0%20to%206.0.pdf

Kilduff, M., & Tsai, W. (2003). *Social Networks and Organizations*. SAGE Publications.

Kim, N., Lee, S., Cho, H., Kim, B.I., & Jun, M. (2018, January). Design of a Cyber Threat Information Collection System for Cyber Attack Correlation. *2018 International Conference Platform Technology Service (PlatCon)*. Jeju, Korea (South). pp. 1-6. DOI: 10.1109/PlatCon.2018.8472775. Retrieved 18 October 2022 from https://ieeexplore.ieee.org/document/8472775

Klausen, J. (2014). Tweeting the Jihad: Social Media Networks of Western Foreign Fighters in Syria and Iraq. *Studies in Conflict & Terrorism*, *38*(1), 1-22. https://doi.org/10.1080/1057610X.2014.974948 https://www.tandfonline.com/doi/full/10.1080/1057610X.2014.974948

Kleckova, A. (2021, February). Open Source Intelligence and Terrorism. Prague Security Studies Institute. Alumni Brief -1. Retrieved 26 November 2022 from https://www.pssi.cz/download/docs/8539_pssi-alumni-brief-01-osint-4.pdf

Kost, E. (2023, January 5). *5 Step Guide: How to Perform a Cyber Risk Analysis in 2023*. UpGuard. Retrieved 27 January 2023 from https://www.upguard.com/blog/how-to-perform-a-cyber-risk-analysis#:~:text=Cyber%20risk%20is%20calculated%20by,x%20Vulnerability%20x%20Information%20Value

Krebs, B. (2014, November 6). *Feds Arrest Alleged 'Silk Road 2 Admin,' Seize Servers.* Krebs on Security. Retrieved 18 December 2022 from https://krebsonsecurity.com/2014/11/feds-arrest-alleged-silk-road-2-admin-seize-servers/

La Pierre, H. (2020). Intelligence Preparation of the Operating Environment: Building Towards a Better Understanding of Cyber Operations. Canadian Forces College. Retrieved 1 June 2023 from https://www.cfc.forces.gc.ca/259/290/22/305/LaPierre.pdf

Larsen, H.L., Blanco, J.M., Pastor, R.P., & Yager, R.R. (Eds). (2017). *Using Open Data to Detect Organised Crime Threats: Factors Driving Future Crime*. Springer International Publishing.

LaSorsa, J.M. (2022). *Modern Tools for a Traditional Task*. Circuit Magazine. Retrieved 19 October 2022 from https://circuit-magazine.com/protective-intelligence/

Lee, S. (2022, January 20). *8 Best Dark/Deep Web Browsers for Anonymous Web Searching in 2022*. Wondershare. Retrieved 15 December 2022 from https://drfone.wondershare.com/dark-web/dark-web-browser.html

Leiner, B., Cerf, V.G., Clark, D.D., Khan, R.E., Kleinrock, L., Lynch, D., Postel, J., Roberts, L.G., & Wolff, S. (1997). *Brief History of the Internet*. Internet Society. Retrieved 24 November 2022 from https://www.internetsociety.org/internet/history-internet/brief-history-internet/

Leinwand Leger, D. (2014, May 15). *How FBI Brought Down Cyber-Underworld Site Silk Road*. USA Today. Retrieved 18 December 2022 from https://www.usatoday.com/story/news/nation/2013/10/21/fbi-cracks-silk-road/2984921/

LexisNexis. (2014, February). *Social Media Use in Law Enforcement*. Retrieved 19 October 2022 from https://risk.lexisnexis.com/-/media/files/government/infographics/2014-social-media-use-in-law-enforcement-infographic%20jpg.jpg?h=614&iar=0&w=900&hash=B23283DEAD007D1DA80B557E541DE18B

Library and Archives of Canada Act, S.C. 2004, c. 11. Justice Laws Website: Government of Canada. Retrieved 8 May 2023 from https://laws-lois.justice.gc.ca/PDF/L-7.7.pdf

Loehrke, B., al-Kaisy, A., Kenausis, L., Terrill, D., & Smits, K. (2021). *Feeling the Burden: Ethical Challenges and Practices in Open Source Analysis and Journalism*. Ethical Journalism Network. Stanley Center for Peace and Security. Planet Labs, Inc. Retrieved 18 October 2022 from https://stanleycenter.org/wp-content/uploads/2022/01/NWRPT-FeelingtheBurden_122-v2.pdf

Mahadevan, P. (2019, June 27). *A Decade on from the 2008 Mumbai Attack: Reviewing the Question of State*-Sponsorship. International Centre for Counter-Terrorism. Retrieved 30 November 2022 from https://www.icct.nl/publication/decade-2008-mumbai-attack-reviewing-question-state-sponsorship

Makkula Center for Applied Ethics. (2021, November 8). *A Framework for Ethical Decision Making*. Santa Clara University. Retrieved 16 February 2023 from https://www.scu.edu/ethics/ethics-resources/a-framework-for-ethical-decision-making/

Markham, A., & Buchanan, E. (2012). *Ethical Decision---Making and Internet Research*. Association of Internet Research. Retrieved 15 February 2023 from http://aoir.org/reports/ethics2.pdf

Marx, G. (2013, September/October). The Public as Partner? Technology Can Make Us Auxiliaries as Well as Vigilantes. *Copublished by the IEEE Computer and Reliability Societies*. 1540-7993/13/$31.00. Retrieved 26 November 2022 from https://web.mit.edu/gtmarx/www/marx-publicas.html

Maxar Technologies. (2022, June 16). *Maxar's WeatherDesk Predicts a Significant Decline in Ukrainian Crop Harvests*. Earth Intelligence. Retrieved 29 April 2023 from https://blog.maxar.com/earth-intelligence/2022/maxars-weatherdesk-predicts-a-significant-decline-in-ukrainian-crop-harvests

MilOsintCTF. (2023). *Challenges*. Retrieved 29 March from https://milosintctf.com/challenges

Milgram, S. (1967, May). The Small World Problem. *Psychology Today*. 60-67.

Misbahuddin, S. (2019). What's the difference between the internet and the world wide web? BBC. Retrieved 24 November 2022 from https://www.bbc.co.uk/newsround/47523993#:~:text=The%20world%20wide%20web%2C%20or,connect%20towns%20and%20cities%20together.

Ministry of Defence. (2014, September). *NATO Standard AJP-3.10.1 Allied Joint Doctrine for Psychological Operations*. Edition B Version 1 + UK National Elements. Retrieved 27 January 2023 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/450521/20150223-AJP_3_10_1_PSYOPS_with_UK_Green_pages.pdf

Mitchell, A., Blazina, C., Stocking, G., Walker, M., & Fedeli, S. (2019, June 5). *Many Americans Say Made-Up News Is a Critical Problem that Needs to be Fixed*. Pew Research Center. Retrieved on 6 February 2023 from https://www.pewresearch.org/journalism/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/

Montgomery, D., Horwitz, S., & Fisher, M. (2013), *Police, Citizens and Technology Factor into Boston Bombing Probe*. The Washington Post. Retrieved 26 November 2022 from https://www.washingtonpost.com/world/national-security/inside-the-investigation-of-the-boston-marathon-bombing/2013/04/20/19d8c322-a8ff-11e2-b029-8fb7e977ef71_story.html

Moses, L.B., Oboler, A., & Parnaby, L. (2021, June 2). *Tracking Violent Extremism Online and the Challenge of Open-Source Intelligence*. Global Network on Extremism & Technology. Retrieved 1 February 2023 from https://gnet-research.org/2021/06/02/tracking-violent-extremism-online-and-the-challenge-of-open-source-intelligence/

Myre, G. (2022, April 26). How does Ukraine Keep Intercepting Russian Military Communications? NPR. Retrieved 17 November 2022 from https://www.npr.org/2022/04/26/1094656395/how-does-ukraine-keep-intercepting-russian-military-communications

Nachsin, J. S. (2010). Terrorism and the Use of Social Media in Mumbai Attacks. *Journal of Counterterrorism & Homeland Security International*, *16*(4).

National Defence. (2000). *Land Force Information Operations Intelligence Field Manual*. B-GL-357-001/FP-001. Government of Canada. Retrieved 25 June 2023 from www.castpics.net/pdfs/m/CA/b-gl-357-001%20-%20Intelligence%20Field%20Manual.pdf

National Defence Act, R.S.C. 1985, c. N-5. Justice Laws Website: Government of Canada. Retrieved 8 May 2023 from https://laws-lois.justice.gc.ca/PDF/N-5.pdf

Nespoli, P., Papamartzivanos, D., Marmol, F.G., & Kambourakis, G. (2018, 2nd Quarter). Optimal Countermeasures Selection Against Cyber Attacks: A Comprehensive Survey on Reaction Frameworks. *IEEE Communications Surveys & Tutorials*. *20*(2), 1361-1396. DOI:10.1109/COMST.2017.2781126

NexusXplore. (2021). *What is NexusXplore?* NexusXplore: Find People, Places, & Information. Developed by OSINT Combine. Retrieved 10 June 2023 from https://www.nexusxplore.com/

Nhan, J., Huey, L., & Broll, R. (2015, December 5). Digilantism: An Analysis of Crowdsourcing and the Boston Marathon Bombings. *The British Journal of Criminology*, *57*(2), 341–361. https://doi.org/10.1093/bjc/azv118. Retrieved 26 November 2022 from https://academic.oup.com/bjc/article/57/2/341/2623876?login=false

Nordine, J. (2023). The OSINT Framework. Retrieved 20 April 2023 from https://osintframework.com/

Norton Rose Fulbright. (2018, April). *What Should Canadian Businesses Know About GDPR*. Retrieved 7 February 2023 from https://www.nortonrosefulbright.com/en/knowledge/publications/db9b8e2b/what-should-canadian-businesses-know-about-gdpr#:~:text=The%20GDPR%20will%20also%20apply,of%20individuals%20in%20the%20EU.

Nykamp D.Q. (2022). *Node definition*. Math Insight. Retrieved 1 April 2023 from https://mathinsight.org/search/?q=node+definition#search

Office of the High Commissioner for Human Rights (OHCHR), and the Human Rights Center at the University of California, Berkeley (UC Berkeley), School of Law. (2022, January 3). *The Berkeley Protocol on Digital Open Source Investigations*. Human Rights Center. New York: United Nations. Retrieved 16 February 2023 from https://humanrights.berkeley.edu/berkeley-protocol-digital-open-source-investigations

Office of the Privacy Commissioner of Canada. (2019, May). *PIPEDA in Brief*. Retrieved 7 February 2023 from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/

Office of the Privacy Commissioner of Canada. (2013, May). *The Case for Reforming the Personal Information Protection and Electronic Documents Act*. Retrieved 7 February 2023 from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_r/pipeda_r_201305/

Office of the Privacy Commissioner of Canada. (2021, October 14). *Strengthening National Security and Privacy in the Digital Era.* Speech delivered by Daniel Therrien, Privacy Commissioner of Canada on 14 October 2021. Last modified 13 January 2022. Retrieved 9 February 2023. from https://www.priv.gc.ca/en/opc-news/speeches/2021/sp-d_20211014/

Office of the Privacy Commissioner of Canada. (2022a, September 21). *Privacy Act Reform*. The Privacy Act. Retrieved 9 February 2023 from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-privacy-act/pa_r/

Office of the Privacy Commissioner of Canada. (2022b, May 26). *The State of Privacy As I End My Term. Remarks at the International Association of Privacy Professionals (IAPP) Canada Privacy Symposium 2022*. Address by Daniel Terrien, Privacy Commissioner of Canada. Retrieved 9 February 2023 from https://www.priv.gc.ca/en/opc-news/speeches/2022/sp-d_20220526/

OSINT Combine. (2023). *Available Free Tools*. OSINT Combine. Retrieved 20 April 2023 from https://www.osintcombine.com/tools

Palantir. (2023) *Palantir Foundry*. Palantir. Retrieved 10 June 2023 from https://www.palantir.com/docs/foundry/platform-overview/overview/

Panimalar, A., Shree, V., & Kathrine, V.A. (2017, September). The 17 V's of Big Data. *International Research Journal of Engineering and Technology*. *4*(9), 329-333. Retrieved 16 January 2023 from https://www.irjet.net/archives/V4/i9/IRJET-V4I957.pdf

Parliament of Canada - House of Commons. (2016, December). Calkins, B (Chair). *Protecting The Privacy of Canadians: Review of the Privacy Act: Report of the Standing Committee on Access to Information, Privacy and Ethics*. 42nd Parliament, 1st Session. Retrieved 9 February 2023 from https://www.ourcommons.ca/Content/Committee/421/ETHI/Reports/RP8587799/ethirp04/ethirp04-e.pdf

Pastor-Galindo, J. Nespoli, P., Marmol, F.G., & Perez, G.M. (2020, January 5). *The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends*. IEEE Access. DOI 10.1109/ACCESS.2020.2965257. Retrieved 16 January 2023 from https://ieeexplore.ieee.org/ielx7/6287639/8948470/08954668.pdf

PBS. (1998). *Cyrillic Alphabet*. PBS WETA. Retrieved 29 March 2023 from https://www.pbs.org/weta/faceofrussia/reference/cyrillic.html

Pellet, H., Shiaeles, S., & Stavrou, S. (2019, March). Localising Social Network Users and Profiling Their Movement. *Computer Security*. *81*, 49-57. Retrieved 18 October 2022 from https://www.sciencedirect.com/science/article/abs/pii/S0167404818301524

Penny, J.W. (2019, October 28). *Protecting Information Consumers*. Centre for International Governance Innovation. Retrieved 1 February 2023 from https://www.cigionline.org/articles/protecting-information-consumers/

Personal Information Protection and Electronic Documentation Act, S.C. 2000, c. 5. Government of Canada. Retrieved 26 November 2022 from https://laws-lois.justice.gc.ca/PDF/P-8.6.pdf

Poulsen, K. (2013, September 13). *FBI Admits It Controlled Tor Servers Behind Mass Malware Attack.* Wired. Retrieved 18 December 2022 from http://www.wired.com/threatlevel/2013/09/freedom-hosting-fbi/

Pournouri, S., Zargari, S., & Akhgar, B. (2019, January). An Investigation of Using Classification Techniques in Prediction of Type of Targets in Cyber Attacks. IEEE 12th International Conference Global Security, Safe, Sustainable (ICGS3). pp. 202-212. DOI:10.1109/ICGS3.2019.8688266.

Powers, S. M., & Jablonski, M. (2015). *The Real Cyber War: The Political Economy of Internet Freedom*. University of Illinois Press.

Privacy Act, R.S.C., 1985, c. P-21. Government of Canada. Retrieved 7 February 2023 from https://laws-lois.justice.gc.ca/eng/ACTS/P-21/page-1.html#h-397177

Privy Council Office. (2023). Intelligence Analyst Learning Program. Transport Canada. 17 February 2023. Retrieved via Access to Information Request.

Proceeds of Crime (Money Laundering) and Terrorist Financing Act, S.C. 2000, c. 17. Justice Laws Website: Government of Canada. Retrieved 7 May 2023 from https://laws-lois.justice.gc.ca/eng/acts/P-24.501/page-8.html#h-399036

Public Order Emergency Commission. (2023). *Public Order Emergency Commission Institutional Report – Royal Canadian Mounted Police*. Retrieved 18 May 2023 from https://publicorderemergencycommission.ca/files/exhibits/DOJ.IR.00000011.pdf

Public Safety Canada. (2023). Access to Information and Privacy - Response. 8 May 2023.

Puente, M. (2019, March 12). *LAPD Data Programs Need Better Oversight to Protect Public, Inspector General Conclude*s. Los Angeles Times. Retrieved on 29 January 2023 from https://www.latimes.com/local/lanow/la-me-ln-lapd-data-20190312-story.html

Puiu, T. (2022, March 15). *How Open-Source Intelligence (OSINT) is Exposing the Ukraine War in Real-Time*. ZME Science Retrieved 13 October 2022 from https://www.zmescience.com/science/news-science/open-source-intelligence-ukraine/

Puleri, J. (2021). Law Enforcement and Open Source Intelligence: Evolution, Technologies, and Privacy Issues. *Utica College ProQuest Dissertations Publishing*. Retrieved 15 October 2022 from https://www.proquest.com/docview/2532617701?parentSessionId=8PCf%2FVtXPH75%2FuigdTyHtBDzSZMOF34wadd7l45QpCw%3D&pq-origsite=primo&accountid=15182

Purdue University. (2022). *Glossary: Technological Neutrality*. Purdue University. Retrieved 26 May 2023 from https://purdue.edu/critical-data-studies/collaborative-glossary/technological-neutrality.php

Quick, D., & Choo, K.K.R. (2018, January). Digital Forensic Intelligence: Data Subsets and Open Source Intelligence (DFINT+OSINT): A Timely and Cohesive Mix. *Future Generation Computer Systems*. 78, 558-567, Jan. 2018.

Qusef, A., & Alkilani, H. (2022) The effect of ISO/IEC 27001 Standard Over Open Source Intelligence. *PeerJ Computer Science, 8*(810). DOI: https://doi.org/10.7717/peerj-cs.810. Retrieved 26 Jan 2023 from https://peerj.com/articles/cs-810/#p-10

Rancich, A. M., & Gelpi, R. J. (1998). Análisis de los Principios Eticos en Juramentos Médicos Utilizados en las Facultades de Medicina de la Argentina en Relación al Hipocratico [Analysis of the Ethical Principles in Medical Oaths Used by Medical Schools of Argentina in Relation to the Hippocratic Oath]. *Medicina*, *58*(2), 147–152. Retrieved 15 February 2023 from https://pubmed.ncbi.nlm.nih.gov/9706247/#:~:text=The%20consensus%20was%20on%20the,and%20the%20rule%20of%20confidentiality

Recorded Future. (2019, January 9). *How Artificial Intelligence is Shaping the Future of Open Source Intelligence.* Recorded Future. Retrieved 15 May 2023 from https://www.recordedfuture.com/open-source-intelligence-future

Renshaw, M., & Chaplin, L. (2020, December 22). *Berkeley Protocol on Digital Open Source Investigations*. Leigh Day. Retrieved 16 February 2023 from https://www.leighday.co.uk/latest-updates/blog/2020-blogs/berkeley-protocol-on-digital-open-source-investigations/

Ressler, S. (2006, July). Social Network Analysis as an Approach to Combat Terrorism: Past, Present, and Future Research. *Homeland Security Affairs*; *2*(2), 1-11. Retrieved 8 March 2023 from https://www.proquest.com/openview/dae6d26ba2112d9cda358bbde499ce7a/1?pq-origsite=gscholar&cbl=1336360

Riebe, T., Kaufhold, M.A., Biselli, T., & Reuter, C. (2023). Privacy Concerns and Acceptance Factors of OSINT for Cybersecurity: A Representative Survey. *Proceedings on Privacy Enhancing Technologies 2023(1), 477-493*. Retrieved 28 January 2023 from https://petsymposium.org/popets/2023/popets-2023-0028.pdf

Ringin, S. (2022, March 17). *Open Source Intelligence Investigations - Benefits and Risks*. Duxton Hill. Retrieved 25 January 2023 from https://duxtonhill.com.au/open-source-intelligence-investigations-benefits-and-risks/

Roache, M, Tewa, S., Cadier, A., Labbe, C., Padovese, V., Schmid, R., O'Rielly, E., Richter, M., Konig, K., Sadeghi, M., Vercellone, C., Fishman, Z., Adams, N., Pavilonis, V., Walid, S., Griffin, K., Palmer, C., Slomka, A., Vallee, L., Kapoor, A., Maitland, E., Wang, M., & Palmer, K. (2023). *Russia-Ukraine Disinformation Tracking Center: 372 Websites Spreading War Disinformation and The Top Myths They Publish*. NewsGuard. Retrieved 29 April 2023 from https://www.newsguardtech.com/special-reports/russian-disinformation-tracking-center/

Royal Canadian Mounted Police Act, R.S.C. 1985. c. R-10. Justice Laws Website: Government of Canada. Retrieved 8 May 2023 from https://laws-lois.justice.gc.ca/PDF/R-10.pdf

Royal Canadian Mounted Police. (2021, January). *Audit of Open Source Information*. Royal Canadian Mounted PoliceRetrieved 18 October 2022 from https://www.rcmp-grc.gc.ca/en/audit-open-source-information?wbdisable=true

Royal Canadian Mounted Police Regulations (SOR/2014-281). Justice Laws Website: Government of Canada. Retrieved 18 May 2023 from https://laws.justice.gc.ca/PDF/SOR-2014-281.pdf

Sandor, A. (2020). An Intelligence Perspective on Privacy and Data Protection Risks in Social Media. Sciendo. *International Conference Knowledge-Based Organization*, XXVI(1), 151-156. NATO HUMINT Centre of Excellence, Oradea, Romania. DOI: 10.2478/kbo-2020-0023. Retrieved 28 January 2023 from https://sciendo.com/article/10.2478/kbo-2020-0023

SANS. (2023). *MGT433: Managing Human Risk*. SANS. Retrieved 10 June 2023 from https://www.sans.org/cyber-security-courses/managing-human-risk/

Scassa, T. (2020). *Replacing Canada's 20-Year-Old Data Protection Law*. Centre for International Governance Innovation. Retrieved 7 February 2023 from https://www.cigionline.org/articles/replacing-canadas-20-year-old-data-protection-law/#:~:text=First%2C%20PIPEDA%20has%20fallen%20sadly,demands%20of%20the%20data%20economy.

Schaurer, F., & Storger, J. (2013, Winter/Spring). The Evolution of Open Source Intelligence (OSINT). *The Intelligencer, 19*(3), 53-56. Retrieved 20 September 2022 from https://www.afio.com/publications/Schauer_Storger_Evo_of_OSINT_WINTERSPRING2013.pdf

Shneiderman, B. (2008, June). Extreme Visualization: Squeezing a Billion Records into a Million Pixels. (3)12. DOI: 10.1145/1376616.1376618. Retrieved 18 October 2022 from https://www.cs.umd.edu/~ben/papers/Shneiderman2008Extreme.pdf

Siddiqui, S. (2019, June 7). *Half of Americans See Fake News as Bigger Threat than Terrorism, Study Finds*. The Guardian. Retrieved 6 February 2023 from https://www.theguardian.com/us-news/2019/jun/06/fake-news-how-misinformation-became-the-new-front-in-us-political-warfare

Simonson, W., & Alsbrooks, W. (1975). A DBMS for the U.S. Bureau of the Census. *ACM DL Digital Library*. Association for Computing Machinery. Retrieved 18 December 2022 from https://dl.acm.org/doi/10.1145/1282480.1282521

Smith-Boyle, V. (2022, June 22). *How OSINT Has Shaped the War in Ukraine.* American Security Project. https://www.americansecurityproject.org/osint-in-ukraine/

Solmaz, T. (2022, February 25). 'Hybrid Warfare': One Term, Many Meanings. *Small Wars Journal*. Retrieved 20 May 2023 from https://smallwarsjournal.com/jrnl/art/hybrid-warfare-one-term-many-meanings

Southern Adventist University. (2022, October 31). *Google & Google Scholar: Boolean Operators*. Boolean Operators. Retrieved 24 April 2023 from https://southern.libguides.com/google/boolean

Srnicek, N. (2016). *Platform Capitalism*. Wiley.

Standing Committee on Access to Information, Privacy and Ethics (ETHI). (2016, November 1) *Evidence*. No 032, 1st Session, 42nd Parliament. Retrieved 15 January 2023 from https://publications.gc.ca/collections/collection_2016/parl/xc73-1/XC73-1-2-421-32-eng.pdf

Talkwalker. (2023a). *Talkwalker Alerts*. Retrieved 27 April 2023 from https://www.talkwalker.com/alerts

Talkwalker. (2023b). *Talkwalker Social Media Analytics Search*. Retrieved 27 April 2023 from https://www.talkwalker.com/social-media-analytics-search

Tamburello, T., & Gravelle, D. (2006a, January). *Soldiers Rung in the New Year with a Splash*. Defense Visual Information Distribution Service. January 2006. Retrieved 22 April 2023 from https://www.dvidshub.net/image/14822/polar-bear-swim

Tamburello, T., & Gravelle, D. (2006b, January). *Soldiers Rung in the New Year with a Splash*. Defense Visual Information Distribution Service. January 2006. Retrieved 22 April 2023 from https://www.dvidshub.net/image/14893/polar-bear-swim

Tamburello, T., & Gravelle, D. (2006c, January). *Soldiers Rung in the New Year with a Splash*. Defense Visual Information Distribution Service. January 2006. Retrieved 22 April 2023 from https://www.dvidshub.net/news/5236/soldiers-ring-new-year-with-splash

Taylor, P. (2022, September 8). *Volume of Data/Information Created, Captured, Copied, and Consumed Worldwide from 2010 to 2020, with Forecasts from 2021 to 2025*. Statista. Retrieved 26 November 2022 from https://www.statista.com/statistics/871513/worldwide-data-created/

Tow, J., & Yeo, W. (2005, August, 3). *The Role of Open Source Intelligence in the Global War on Terror*. Rajaratnam School of International Studies, Nanyang Technical University. Retrieved 14 October 2022 from https://www.rsis.edu.sg/rsis-publication/idss/712-the-role-of-open-source-intell/?doing_wp_cron=1665755229.9112360477447509765625#.Y0loYnbMKUk

Townsend, L., & Wallace, C. (2016, July 12). *Social Media Research: A Guide to Ethics*. The University of Aberdeen. Retrieved 13 February 2023 from https://ahrecs.com/resources/social-media-research-a-guide-to-ethics-by-townsend-and-wallace-guidance-dr-leanne-townsend-prof-claire-wallace-2016/

Troitter, D. (2015). Open Source Intelligence, Social Media and Law Enforcement: Visions, Constraints and Critiques. *European Journal of Cultural Studies*. *18*(4-5), 530-547. DOI: 10.1177/1367549415577396. DOI: 10.2478/kbo-2020-0023 https://www.researchgate.net/publication/280291560_Open_source_intelligence_social_media_and_law_enforcement_Visions_constraints_and_critiques

Tsvetovat, M., & Kouznetsov, A. (2011). *Social Network Analysis for Startups: Chapter 4: Cliques, Clusters and Components*. O'Reilly Media, Inc: Cambridge. https://www.oreilly.com/library/view/social-network-analysis/9781449311377/ch04.html

Tunney, C. (2019, November 5). *RCMP Launches Review of its Social Media Monitoring Operation*. CBC News: Politics. Retrieved 18 October 2022 from https://www.cbc.ca/news/politics/rcmp-social-media-review-1.5346741

Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011, November 18). *The Anatomy of the Facebook Social Graph. Cornell University*. Arxiv. Retrieved 3 May 2023 from https://arxiv.org/abs/1111.4503

United Nations Conference on Trade and Development. (2021). *Data Protection and Privacy Legislation Worldwid*e. UNCTAD. Retrieved 1 February 2023 from https://unctad.org/page/data-protection-and-privacy-legislation-worldwide

United Nations Office on Drugs and Crime. (2012). *Countering the Use of the Internet for Terrorist Purposes*. United Nations Office on Drugs and Crime. Retrieved 18 October 2022 from https://www.unodc.org/unodc/en/terrorism/news-and-events/use-of-the-internet.html

U.S. Department of the Army. (2012, July). *Open-Source Intelligence, ATP 2-22.9*. U.S. Department of the Army. Retrieved 18 October 2022 from https://irp.fas.org/doddir/army/atp2-22-9.pdf

Van der Weide, Y. (2020, December 3). *Using the Sun and the Shadows for Geolocation*. *Bellingcat.* Retrieved 3 December 2022 from https://www.bellingcat.com/resources/2020/12/03/using-the-sun-and-the-shadows-for-geolocation/

Varda, D.M., Forgette, R., & Banks, D. (2009). Social Network Methodology in the Study of Disasters: Issues and Insights Prompted by Post-Katrine Research. *Population Research and Policy Review*, *28(*1), 11-29. DOI: 10.1007/s11113-008-9110-9. Retrieved 1 March 2023 from https://www.researchgate.net/publication/225163941_Social_Network_Methodology_in_the_Study_of_Disasters_Issues_and_Insights_Prompted_by_Post-Katrina_Research

Vargas, J. A. (2014, August 21) *Comprehensive Assessment of the Impact of Social Media in the United States Intelligence Community*. Contingent Security Services, Ltd. the International Association of Airport and Seaport Police. Retrieved 26 November 2022 from https://www.linkedin.com/posts/joel-vargas-jd-864aa144_assessment-of-the-the-impact-of-social-media-activity-7001042213602258944-Ytse?utm_source=share&utm_medium=member_desktop

Vejvodová, P. (2019). Information and Psychological Operations as a Challenge to Security and Defence. *Vojenské Rozhledy*. *28*, 83-96. DOI: 10.3849/2336-2995.28.2019.03.083-096. Retreived 27 January 2023 from https://www.researchgate.net/publication/335947264_Information_and_Psychological_Operations_as_a_Challenge_to_Security_and_Defence

Verma, P. (2022, March 23). *The Rise of the Twitter Spies*. The Washington Post. Retrieved 13 October 2022 from https://www.washingtonpost.com/technology/2022/03/23/twitter-open-source-intelligence-ukraine/

Visano, L. (2013). "Servitude of the Certitude in the 9/11 Hauntology: A Case Study of (in)securities in Cyber Security". In C. Colaguori (ed) *Security, Life and Death: Governmentality and Biopower in the post-9/11 Era* Toronto: deSitter Press.

Volkert, R. (2019, May 21). *Dark Web: Basics of Breach Data*. OSINT Curious. Retrieved 27 Jan 2023 from https://osintcurio.us/2019/05/21/basics-of-breach-data/#:~:text=%E2%80%9CBreach%20Data%E2%80%9D%20is%20data%20that,very%20useful%20in%20OSINT%20investigations.

Wagner, J. (2015, August 2015). *Photo Forensics Tutorial [Video]*. YouTube. Retrieved 1 May 2023 from https://www.youtube.com/watch?v=XRCq8CJrI_s&t=601s

Wang, M.H., Tsai, M.H., Yang, W.C., & Lei, C.L. (2018a, April). Infection Categorization Using Deep Autoencoder. *IEEE Conference on Computing and Communications.* Workshops (INFOCOM Workshops). pp. 1-2. DOI:10.1109/INFCOMW.2018.8406878. Retrieved 18 October 2022 from https://www.researchgate.net/publication/326702940_Infection_categorization_using_deep_autoencoder

Wang, R., Ji, W., Liu, M., Wang, X., Weng, J., Deng, S., Gao, S., & Yuan, C.A. (2018b). Review on Mining Data from Multiple Data Sources. *Pattern Recognition. Letters*. July 2018. 109, 120-128.

Watts, D. (1999). Networks, Dynamics and the Small World Phenomenon. *American Journal of Sociology*, *105*(2), 493-527.

Watts, D. (2003). *Six Degrees: The Science of a Connected Age*. W.W. Norton & Company.

Weimann, G. (2015). Going Dark: Terrorism on the Dark Web. *Studies in Conflict & Terrorism*, *39*:3, 195-206, DOI: 10.1080/1057610X.2015.1119546 Retrieved 18 December 2022 from https://www.tandfonline.com/doi/abs/10.1080/1057610X.2015.1119546?journalCode=uter20

Wheaton, K.J., & Richey, M.K. (2014, January 9). *The Potential of Social Network Analysis in Intelligence*. E-International Relations. Retrieved 8 March 2023 from https://www.e-ir.info/2014/01/09/the-potential-of-social-network-analysis-in-intelligence/

Widdis, J. (2023). Email Correspondence. ATIP Response - 11 January 2023. (See Appendix 13).

Williams, H.J., & Blum, I. (2018). *Defining Second Generation Open Source Intelligence (OSINT) for the Defense Enterprise*. RAND Corporation: Santa Monica, CA. Retrieved 20 September 2022 from https://www.rand.org/pubs/research_reports/RR1964.html

Windward. (2022, July 15). *Alleged Russian Grain Laundering*. Retrieved 29 April 2023 from https://windward.ai/wp-content/uploads/2022/07/Windward-Insights-Maritime-Meeting-Reveals-a-New-Phenomenon_-Alleged-Russian-Grain-Laundering-July-2022-2.pdf

Yeboah-Ofori, A., & Brimicombe, A. (2017). Cyber Intelligence and OSINT: Developing Mitigation Techniques Against Cybercrime Threats on Social Media. *International Journal of Cyber-Security and Digital Forensics, 7*(1), 87-98. https://doi.org/10.17781/P002378. Retrieved 18 October 2022 from https://repository.uel.ac.uk/item/88300

Younas, M.A. (2014, March). Digital Jihad" and its Significance to Counterterrorism. *Counter Terrorist Trends and Analyses, 6*(2), 10-17. Retrieved 26 November 2022 from https://www.jstor.org/stable/26351231

Young, S. (2018, February 22). Technology and Function Creep. Office of the Saskatchewan Information and Privacy Commissioner. https://oipc.sk.ca/technology-and-function-creep-2/#:~:text=%E2%80%9CFunction%20creep%E2%80%9D%20occurs%20when%20information,sign%2Dout%20of%20the%20workplace

Zuboff, S. (2015). Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology, 30*(1), 75–89. https://doi.org/10.1057%2Fjit.2015.5

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.

Appendices

Appendix 1: Data Prefix and Multiplier Conversion Table

Unit	Abbreviation	Approximate Size
bit	b	Binary digit, a single 0 or 1
Byte	B	8 bits
kilobyte	KB	1,024 bytes or 10^3 bytes
megabyte	MB	1,024 KB or 10^6 bytes
gigabyte	GB	1,024 MB or 10^9 bytes
terabyte	TB	1,024 GB or 10^{12} bytes
petabyte	PB	1,024 TB or 10^{15} bytes
exabyte	EB	1,024 PB or 10^{18} bytes
zettabyte	ZB	1,024 EB or 10^{21} bytes
yottabyte	YB	1,024 or 10^{24} bytes

Table A1: Data Prefix and Multiplier Conversion Table
Source: Fitzgibbons, 2022

Appendix 2: 17 V's of Big Data

No	Big Data Characteristics	Elucidation	Description
1	Volume	Size of Data	Quantity of collected and stored data. Data size is in TB
2	Velocity	Speed of Data	The transfer rate of data between source and destination
3	Value	Importance of Data	It represents the business value to be derived from big data
4	Variety	Type of Data	Different type of data like pictures, videos and audio arrives at the receiving end
5	Veracity	Data Quality	Accurate analysis of captured data is virtually worthless if it's not accurate
6	Validity	Data Authenticity	Correctness of accuracy of data used to extract result in the form of information
7	Volatility	Duration of Usefulness	Big data volatility means the stored data and how long is useful to the user
8	Visualization	Data Act / Data Process	It is a process of representing abstract
9	Virality	Spreading Speed	It is defined as the rate at which the data is broadcast / spread by a user and received by different users for their use
10	Viscosity	Lag of Event	It is a time difference the event occurred and the event being described
11	Variability	Data Differentiation	Data arrives constantly from different sources and how efficiently it differentiates between noisy data or important data
12	Venue	Different Platform	Various types of data arrived from different sources via different platforms like personnel system and private & public cloud
13	Vocabulary	Data Terminology	Data terminology likes data model and data structures
14	Vagueness	Indistinctness of existence in a Data	Vagueness concern the reality in information that suggested little or no thought about what each might convey
15	Complexity	Correlation of Data	Data comes from different sources and it is necessary to figure out the changes whether small or large in data with respect to the previously arrived data so that information can get out quickly
16	Verbosity	Massive data that comes from different sources	The redundancy of the information available at different sources
17	Voluntariness	Data that can be used by different organizations without interference	The will full availability of big data to be used according to the context
18	Versatility	The evolution of Big Data	The ability of big data to be flexible enough to be used differently for different context

Table A2: 17 V's of Big Data

Source: Panimalar, Shree, & Kathrine, 2017, 330-332

Appendix 3: Source Credibility Scales

Evaluation will occur during the processing stage of the intelligence cycle. The two-character evaluation (e.g. F3) represents an assessment on the reliability of the source, followed by the assessed level of confidence of the information. For example, an F3 source would depict a source whose reliability cannot be judged, but is delivering information that has been assessed as possibly true through corroboration of other assessments or information. It is important to note that the reliability of a source and the credibility of the information are considered in isolation of each other as a means of ensuring accuracy.

Reliability of the Collection Capability		Credibility of the Information	
A	Completely reliable	1	Completely credible
B	Usually reliable	2	Probably true
C	Fairly reliable	3	Possibly true
D	Not usually reliable	4	Doubtful
E	Unreliable	5	Improbable
F	Reliability cannot be judged	6	Truth cannot be judged

Table A3: NATO AJP 2.1 2016 Source Reliability and Information Credibility Scales
Source: Irwin & Mandel, 2020, 7-2

Reliability of Source		
A	Completely reliable	Refers to a tried and trusted source which can be depended upon with confidence
B	Usually reliable	Refers to a source which has been successful in the past but for which there is still some element of doubt in a particular case
C	Fairly reliable	Refers to a source which has occasionally been used in the past and upon which some degree of confidence can be based
D	Not usually reliable	Refers to a source which has been used in the past but has proved more often than not unreliable
E	Unreliable	Refers to a source which has been used in the past and has proved unworthy of any confidence
F	Reliability cannot be judged	Refers to a source which has not been used in the past

Table A4: NATO STANAG[160] 2511 Source Reliability Scale
Source: Irwin & Mandel, 7-2

[160] STANAG: Standardization Agreement

Credibility of Information		
1	Confirmed by other source	If it can be stated with certainty that the reported information originates from another source than the already existing information on the same subject, it is classified as "confirmed by other sources" and is rated "1".
2	Probably true	If the independence of the source of any item or information cannot be guaranteed, but if, from the quantity and quality of previous reports its likelihood is nevertheless regarded as sufficiently established, then the information should be classified as "probably true" and given rating of "2".
3	Possibly true	If, despite there being insufficient confirmation to establish any higher degree of likelihood, a freshly reported item of information does not conflict with the previously reported behaviour pattern of the target, the item may be classified as "possibly true" and given rating of "3".
4	Doubtful	An item of information which tends to conflict with the previously reported or established behaviour pattern of an intelligence target should be classified as "doubtful" and given a rating of "4".
5	Improbable	An item of information which positively contradicts previously reported information or conflicts with the established behaviour pattern of an intelligence target in a marked degree should be classified as "improbable" and given a rating of "5".
6	Truth cannot be judged	Any freshly reported item of information which provides no basis for comparison with any known behaviour pattern of a target must be classified as "truth cannot be judged" and given a rating of "6". Such a rating should be given only when the accurate use of a higher rating is impossible.

Table A5: NATO STANAG 2511 Information Reliability Scale

Source: Irwin & Mandel, 2020, 7-3

Appendix 4: Principal OSINT workflows and derived intelligence

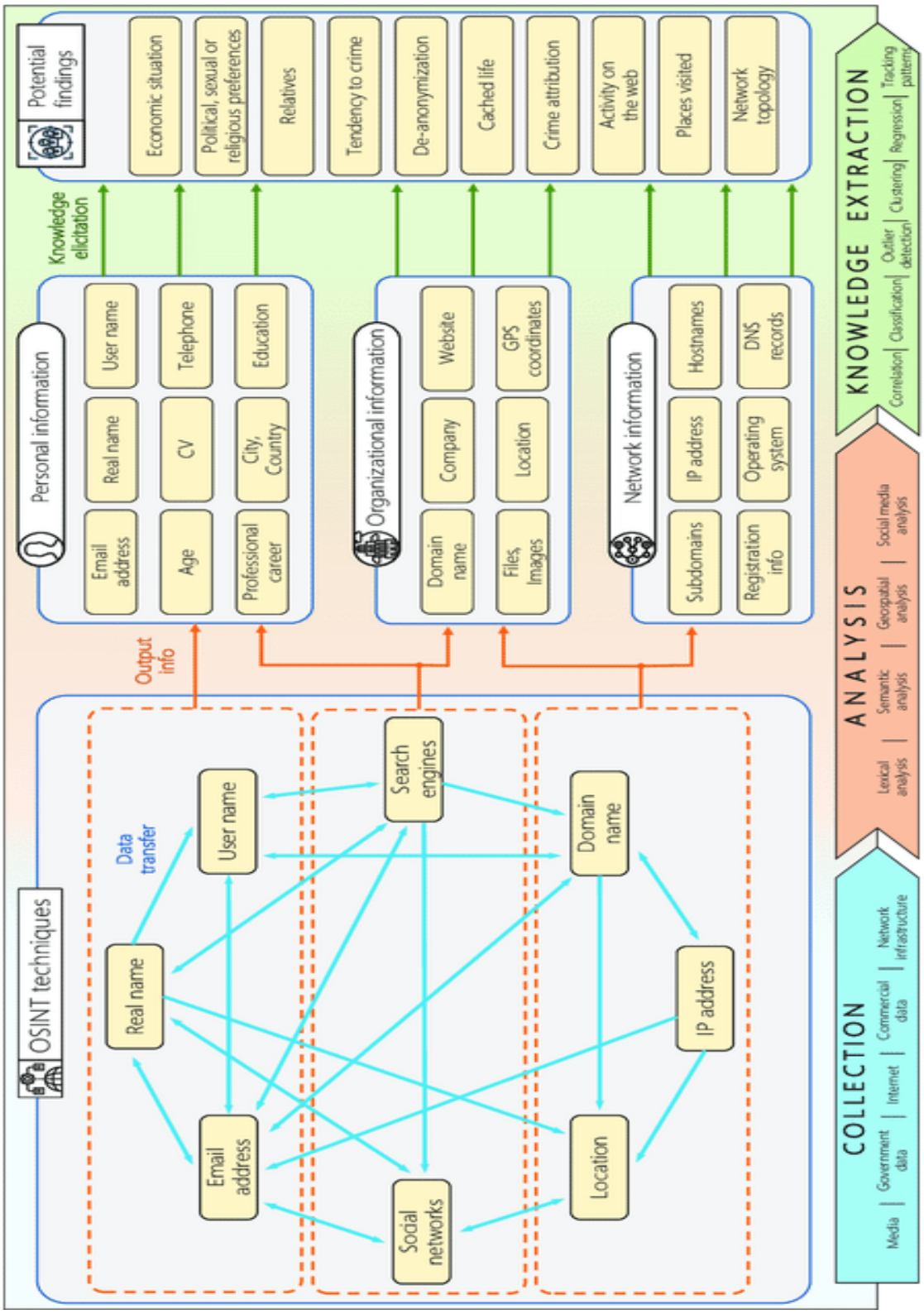

Figure A1: Principal OSINT workflows and derived intelligence
Source: Pastor- Galindo, et al., 2020, 10287

Appendix 5: Preventing, Addressing, and Identifying Vicarious Trauma

Source: Ellis, 2019

Prevention:

- Know Yourself: What images affect you most? Understanding your (and your teams) sensitivities allows for mental preparation and prevents trauma.
- Warn colleagues about traumatic media that you intend to show them, or that they may come across.
- Understand your background and environment: Having a personal connection to your work may intensify trauma.

Identification:

- Sleep: Are you not able to sleep, or are you having nightmares?
- Nutrition: Has your appetite decreased?
- Relationships: Are you spending more time alone? Are you pushing friends and family away?
- Substance Abuse: Has your relationship to drugs and/or alcohol changed?

Address:

- Take a step back: Pause and reflect on the effects your trauma might be having on you
- Connect with friends and family: Connecting with loved ones may provide a welcome break.
- Connect with colleagues: Your colleagues may be going through the same effects as you are. They know what you have been exposed to on a daily basis. Taking the time to reflect on this and support each other assists in addressing and overcoming trauma.
- Relaxation: Exercise and meditation have been shown to alleviate the symptoms of vicarious trauma as they help to reduce anxiety and depression through providing an aspect of your life you can control.
- Professional Counseling: Counselling may provide additional techniques to alleviate trauma.

Appendix 6: The 10 Principles of the Digital Charter

Source: Directly adapted from (Innovation, Science and Economic Development Canada, 2022b)

1. Universal Access:
All Canadians will have equal opportunity to participate in the digital world and the necessary tools to do so, including access, connectivity, literacy and skills.

2. Safety and Security:
Canadians will be able to rely on the integrity, authenticity and security of the services they use and should feel safe online.

3. Control and Consent:
Canadians will have control over what data they are sharing, who is using their personal data and for what purposes, and know that their privacy is protected.

4. Transparency, Portability and Interoperability:
Canadians will have clear and manageable access to their personal data and should be free to share or transfer it without undue burden.

5. Open and Modern Digital Government:
Canadians will be able to access modern digital services from the Government of Canada, which are secure and simple to use.

6. A Level Playing Field:
The Government of Canada will ensure fair competition in the online marketplace to facilitate the growth of Canadian businesses and affirm Canada's leadership on digital and data innovation, while protecting Canadian consumers from market abuses.

7. Data and Digital for Good:
The Government of Canada will ensure the ethical use of data to create value, promote openness and improve the lives of people—at home and around the world.

8. Strong Democracy:
The Government of Canada will defend freedom of expression and protect against online threats and disinformation designed to undermine the integrity of elections and democratic institutions.

9. Free from Hate and Violent Extremism:
Canadians can expect that digital platforms will not foster or disseminate hate, violent extremism or criminal content.

10. Strong Enforcement and Real Accountability:
There will be clear, meaningful penalties for violations of the laws and regulations that support these principles.

Appendix 7: Reviewing the Privacy Act: List of Recommendations

Source: Directly adapted from Parliament - House of Commons, 2016, 65-69

RECOMMENDATION 1
a) That the purpose clause in section 2 of the Privacy Act be expanded to reinforce the quasi-constitutional nature of privacy rights by including generally accepted and technologically neutral privacy principles similar to those in contained in the Personal Information Protection and Electronic Documents Act, including accountability; identifying purposes; limiting collection; limiting use, disclosure, and retention; accuracy; safeguards; openness; individual access; and challenging compliance.
b) That the Privacy Act be modified to clarify that the privacy principles in the amended purpose clause shall guide the interpretation of the Act.

RECOMMENDATION 2
That the definition of "personal information" in section 3 of the Privacy Act be amended to ensure that it be technologically neutral[161] and that it includes unrecorded information.

RECOMMENDATION 3
That the Government of Canada define metadata in the Privacy Act, in a technologically neutral way and with an emphasis on the information it can reveal about an individual.

RECOMMENDATION 4
That the Privacy Act be amended to require that all information sharing under paragraphs 8(2)(a) and (f) of the Privacy Act be governed by written agreements and that these agreements include specified elements.

RECOMMENDATION 5
That the Privacy Act be amended to create an explicit requirement that new or amended information-sharing agreements be submitted to the Office of the Privacy Commissioner of Canada for review, and that existing agreements should be reviewable by the Privacy Commissioner upon request.

RECOMMENDATION 6
a) That the Privacy Act be amended to create an explicit requirement that departments be transparent about the existence of any information-sharing agreements.
b) That the Privacy Act be amended to require, except in appropriate circumstances, the publication of the content of information-sharing agreements between departments or with other governments.

RECOMMENDATION 7
That the Privacy Act be amended to create an explicit requirement for institutions to safeguard personal information with appropriate physical, organizational and technological measures commensurate with the level of sensitivity of the data.

[161] Technological neutrality can be defined as individuals and organizations having the freedom to choose technology that is most appropriate to their requirements, especially as they relate to acquisition, development, or commercialization (Carillo, 2018; Purdue University, 2022).

RECOMMANDATION 8
That the Privacy Act be amended to set out clear consequences for failing to safeguard personal information.

RECOMMENDATION 9
That the Privacy Act be amended to create an explicit requirement for government institutions to report material breaches of personal information to the Office of the Privacy Commissioner of Canada in a timely manner.

RECOMMENDATION 10
That the Privacy Act be amended to create an explicit requirement for government institutions to notify affected individuals of material breaches of personal information, except in appropriate cases, provided that the notification does not compound the damage to the individuals.

RECOMMENDATION 11
That section 4 of the Privacy Act be amended to explicitly require compliance with the criteria of necessity and proportionality in the context of any collection of personal information, consistent with other privacy laws in effect in Canada and abroad.

RECOMMENDATION 12
That the Privacy Act be amended to clarify that a recipient federal institution that receives personal information through information sharing with another federal institution is collecting personal information within the meaning of section 4 of the Privacy Act, and must meet the criteria of necessity and proportionality that apply to the collection of personal information.

RECOMMENDATION 13
That section 6 of the Privacy Act be amended so as to explicitly require compliance with the criteria of necessity and proportionality in the context of any retention of personal information.

RECOMMENDATION 14
That the Privacy Act be amended to set clear rules governing the collection and protection of personal information that is collected on the internet and through social media.

RECOMMENDATION 15
a) That the Government of Canada strengthen the oversight of privacy rights by adopting an order-making model with clear and rigorously defined parameters.
b) That, in order to ensure the most effective use of resources, the Government of Canada explore ways of finding efficiencies, by, among other things, combining the adjudicative functions of the Office of the Privacy Commissioner of Canada and the Office of the Information Commissioner of Canada.

RECOMMENDATION 16
That the Government of Canada further examine the possibility of expanding judicial recourse and remedies under the Privacy Act.

RECOMMENDATION 17
That the Privacy Act be amended to include a requirement for government institutions to conduct privacy impact assessments for new or significantly amended programs and submit them to the Office of the Privacy Commissioner of Canada in a timely manner.

RECOMMENDATION 18
That the Privacy Act be amended to require federal government institutions to consult with Office of the Privacy Commissioner of Canada on draft legislation and regulations with privacy implications before they are implemented.

RECOMMENDATION 19
That the Privacy Act be amended to explicitly confer the Privacy Commissioner with:
a) the authority to conduct, on his own initiative, research and studies on issues of public importance, and
b) a mandate to undertake public education and awareness activities.

RECOMMENDATION 20
That the Privacy Act be amended to require an ongoing five-year parliamentary review.

RECOMMENDATION 21
That section 64 of the Privacy Act be amended to create an exemption from confidentiality requirements to provide the Privacy Commissioner with the discretionary authority to report proactively on government privacy issues where he considers it in the public interest to do so.

RECOMMENDATION 22
That the Privacy Act be amended to expand the ability of the Office of the Privacy Commissioner of Canada to collaborate with other data protection authorities and review bodies on audits and investigations of shared concern in connection with Privacy Act issues.

RECOMMENDATION 23
That section 32 of the Privacy Act be amended to grant the Privacy Commissioner discretion to discontinue or decline complaints on specified grounds, including when the complaint is frivolous, vexatious or made in bad faith, and that the Commissioner's decision to discontinue or decline a complaint be subject to a right of appeal by the complainant.

RECOMMENDATION 24
That reporting requirements on broader privacy issues dealt with by federal institutions be reinforced by requiring the addition of a descriptive element so as to make the information in the reports accessible and relevant.

RECOMMENDATION 25
That there be specific transparency requirements for lawful access requests from agencies involved in law enforcement.

RECOMMENDATION 26
That the Government of Canada explore extending the scope of the Privacy Act to all federal government institutions, including ministers' offices and the Prime Minister's Office.

RECOMMENDATION 27
That the Government of Canada consider extending the right of access to personal information to foreign nationals.

RECOMMENDATION 28
That the Government of Canada examine the possibility of limiting exemptions to access to personal information requests under the Privacy Act.

Review Of The Privacy Act – Revised Recommendations Of The Privacy Commissioner Of Canada
Source: Directly adapted from Parliament - House of Commons, 2016, 71

THEME ONE: TECHNOLOGICAL CHANGES

1. Clarify requirements for information-sharing agreements: Require that all information sharing under paragraphs 8(2)(a) and (f) of the Privacy Act be governed by written agreements and that these agreements include specified elements. Further, all new or amended agreements should be submitted to the Office of the Privacy Commissioner of Canada (OPC) for review, and existing agreements should be reviewable upon request. Finally, departments should be required to be transparent about the existence of these agreements.

2. Create a legal obligation for government institutions to safeguard personal information: Create an explicit requirement for institutions to safeguard personal information with appropriate physical, organizational and technological measures commensurate with the level of sensitivity of the data;

3. Make breach reporting mandatory: Create an explicit requirement for government institutions to report material breaches of personal information to the OPC in a timely manner and to notify affected individuals in appropriate cases;

THEME TWO: LEGISLATIVE MODERNIZATION

4. Create an explicit necessity requirement for collection: Amend section 4 of the Privacy Act to create a more explicit necessity requirement for the collection of personal information, consistent with other privacy laws in Canada and abroad;

5. Replace the ombudsman model for the investigation of complaints with OPC powers to issue binding orders;

6. Consider creating a statutory mechanism to independently review privacy complaints against the OPC;

7. Require government institutions to conduct privacy impact assessments (PIAs) for new or significantly amended programs and submit them to OPC prior to implementation;

8. Require government institutions to consult with OPC on draft legislation and regulations with privacy implications before they are tabled;

Appendix 8: Ethical Challenges - Interview with Analysts and Journalists

Source: Directly adapted from Loehrke, et al, 2021, 5-8

Background: The Ethical Journalism Network and the Stanley Center for Peace and Security interviewed 8 journalists and 20 analysts with a view to obtain examples of circumstances in which they or their teams had felt uneasy about publishing findings from their open source investigations.

Excerpt 1: Diplomacy, Crises, and Conflict (Loehrke, et al., 2021, 5)

In these interviews analysts had discussed situations when they were concerned about how something they published might inadvertently create diplomatic tension, or result in an escalation of pressure during ongoing crises.

Excerpt A: Iran's Nuclear program (Loehrke, et al., 2021, 5)

"I'd been tracking Iran's nuclear program for a while. Satellite imagery showed what seemed like a significant, undeclared expansion of its nuclear activities.

If made public, it was the kind of story that could complicate sensitive talks with Iran. On the other hand, the information was already out there in the open source. It's my job to help explain the situation to the public and to policymakers.

I was confident in the analysis. But I worried that publishing it could add to international tensions. So we checked in with official contacts. They clearly wanted us to hold off. I assumed that meant there was some kind of negotiation underway, likely to get access to the site for international inspectors. Public pressure could disrupt that.

We went ahead and published. In the end, inspectors got access to the site."

Excerpt B: North Korean Missile Base (Loehrke, et al., 2021, 6)

"There was this North Korean missile base. Hardly secret. The base had been there a while. I wrote an update on the site, based on some recent imagery. The report ended up getting some media coverage, reframed, and put in context with ongoing diplomatic efforts. This was during some sensitive months between the US, South Korea, and North Korea.

The media coverage of the report ballooned. The story got big enough that the US and South Korean governments commented on the missile developments.

I thought it was useful analysis. Nothing flashy. But it took on a life of its own after we published."

Excerpt C: Iranian Missile Attack on US Bases in Iraq (Loehrke, et al., 2021, 6)

"Remember after the US assassinated Qassam Soleimani, and Iran responded with missile attacks on American bases in Iraq? The administration said, you know, troops weren't really under threat. The attacks missed. Some people got a headache or something.

I looked at imagery after the attacks. Seemed like the Iranians had hit what they said they would. I paused before posting about it. Because could pointing this out worsen the crisis, if it gets caught up in a public narrative and encourages the US president to escalate the conflict?

Open source work is important for evaluating real-world crises and not allowing governments to get away with misleading claims. But in this case, it kinda felt like being accurate, being quick, being transparent, and being truthful still could lead to bad consequences."

Excerpt 2: Unintended Consequences (Loehrke, et al., 2021, 6)

> In some situations, analysts may inadvertently discover information that, if published, could likely assist nefarious actors, or result in collateral damage to systems in place that are designed to sustain a common good. For a number of reasons, these situations can get difficult to recognize.
>
> Excerpt A: Tracking Terrorists in Africa (Loehrke, et al., 2021, 6)
>
> "I was using satellite imagery to track a terrorist organization in Africa. I identified some activity but also inadvertently discovered a US and allied military operation. I shared it with a contact, who asked me to not publish because it'd compromise the operation.
>
> More than happy to do that. Not a problem. That's why we do reviews like that."
>
> Excerpt B: Troubleshooting North Korean Missiles (Loehrke, et al., 2021, 6)
>
> "The last thing I want to do is end up helping the North Koreans troubleshoot their missiles. And I definitely worry about that often.
>
> Like, if I'm looking at photos of a missile. And I were to notice something clearly wrong with the design or performance parameters. And then post something online saying, 'Well, that's your problem right there.
>
> Fix that, and you got a better missile.'"
>
> Excerpt C: Russian and Syrian Bombing Sorties (Loehrke, et al., 2021, 6-7)
>
> "We obtained cockpit recordings of Russian and Syrian pilots on bombing sorties. Using some open source and geospatial mapping, documenting evidence and cross-corroborating data, we could validate the tapes. And show Russian involvement in strikes on hospitals in Syria.
>
> We reported it. They bombed one of the hospitals again. We reported it, again. Our story was held up to the Russian ambassador in a UN Security Council meeting. Soon after, Russia withdrew from an additional protocol to the Geneva Conventions and, later, the UN deconfliction system.
>
> Obviously, they're aware now that there's skilled groups like us who can call them out for war crimes using evidence that they can't deflect like they used to."
>
> Excerpt D: North Korean Missiles (Loehrke, et al., 2021, 7)
>
> "I was looking at videos of missiles coming out of North Korean state news. They often censor their footage before uploading, but sometimes you can see something come into the shot that shouldn't be there.
>
> I pointed out something that, for a single frame, was briefly uncensored. Pretty sure that got the censor in deep trouble."

Excerpt 3: Privacy (Loehrke, et al., 2021, 7)

> "As open source information makes it easier to identify individuals and sources, it is getting much more difficult for teams to protect the identity of their sources. This issue is compounded by the growing requirement to consider the safety and security of sources, subjects and bystanders."
>
> Excerpt A: Protecting Sources (Loehrke, et al., 2021, 7)
>
> "We have to be extremely careful when using bystander video in our reporting.

Immediately after an event—say an explosion or some act of violence—it's fairly common to find firsthand video of the event posted on social media. It can be immensely valuable for verifying a story and adding visuals. But by showing video evidence and how we verify a story, it can endanger the person who uploaded the video.

We were doing a story about a missile strike in a country in the Middle East. We found video of the strike and got in touch with the person who took the video. We wanted to include the video in our reporting. But based on the video, it wouldn't be hard to figure out which building, apartment, or window our contact was standing in when filming. That could get the person arrested or bring harm to a family.

In this case, we didn't publish the video with our reporting.

But you have to anticipate this stuff. You can ask your source, "Is this okay? Can we do this?" You can ask the same questions internally. But it doesn't change the reality that a wrong choice could put someone in danger."

Excerpt B: Due Diligence (Loehrke, et al., 2021, 7)

"We were researching a sanctions-evasion case and looked at several ships that we suspected were engaged in illicit trade. Using corporate records, we were able to identify a particular company involved and its sole shareholder.

We did our due diligence before publishing a paper on the case. We were confident in the analysis, enough to name individuals in the report. But in hindsight, you still ask if it was appropriate. There is so much risk for collateral damage on people who could otherwise be innocent. Did we take all necessary precautions before crossing that line?

It really forced us to think more comprehensively about risk, how we approach risk, and how we disseminate our analyses."

Excerpt 4: Interaction Between Analysts and Journalists (Loehrke, et al., 2021, 8)

Some open source analysts may work with journalists as expert sources as a means to ensure the perspective the journalist is providing is valid, precise, and within the correct context. However, things can still be misinterpreted. In the interviews, analysts and journalists both acknowledged their lack of coordination on their shared responsibilities if ethical challenges emerge.

Excerpt A: Misinformation (Loehrke, et al., 2021, 8)

"You have to be really careful with what you're saying and not.

There was a recent report about an explosion at a nuclear facility in Iran. Some people in the non-proliferation community used geospatial data to validate the location. Those analysts did nothing wrong. It was fine.

But a reporter picked up the story of the explosion and then just sprinkled in, like, "Oh, by the way," an international inspector was detained last year with an assertion that they detected nitrate residue on the person's hands.

That reporting was horribly irresponsible. All the facts were correct but presented in a way that creates the false impression that an international inspector had planted a bomb in a nuclear facility. That's how misinformation starts, and it could erode trust in the essential work that inspectors do."

Appendix 9: The Markkula Framework for Ethical Decision Making

Source: Directly adapted from Markkula Center for Applied Ethics, 2021, np

Identify the Ethical Issues

1. Could this decision or situation be damaging to someone or to some group, or unevenly beneficial to people? Does this decision involve a choice between a good and bad alternative, or perhaps between two "goods" or between two "bads"?
2. Is this issue about more than solely what is legal or what is most efficient? If so, how?

Get the Facts

3. What are the relevant facts of the case? What facts are not known? Can I learn more about the situation? Do I know enough to make a decision?
4. What individuals and groups have an important stake in the outcome? Are the concerns of some of those individuals or groups more important? Why?
5. What are the options for acting? Have all the relevant persons and groups been consulted? Have I identified creative options?

Evaluate Alternative Actions

6. Evaluate the options by asking the following questions:
    - Which option best respects the rights of all who have a stake? (The Rights Lens)
    - Which option treats people fairly, giving them each what they are due? (The Justice Lens)
    - Which option will produce the most good and do the least harm for as many stakeholders as possible? (The Utilitarian Lens)
    - Which option best serves the community as a whole, not just some members? (The Common Good Lens)
    - Which option leads me to act as the sort of person I want to be? (The Virtue Lens)
    - Which option appropriately takes into account the relationships, concerns, and feelings of all stakeholders? (The Care Ethics Lens)

Choose an Option for Action and Test It

7. After an evaluation using all of these lenses, which option best addresses the situation?
8. If I told someone I respect (or a public audience) which option I have chosen, what would they say?
9. How can my decision be implemented with the greatest care and attention to the concerns of all stakeholders?

Implement Your Decision and Reflect on the Outcome

10. How did my decision turn out, and what have I learned from this specific situation? What (if any) follow-up actions should I take?

Six Ethical Lenses

If our ethical decision-making is not solely based on feelings, religion, law, accepted social practice, or science, then on what basis can we decide between right and wrong, good and bad? Many philosophers, ethicists, and theologians have helped us answer this critical question. They have suggested a variety of different lenses that help us perceive ethical dimensions. Here are six of them:

The Rights Lens

Some suggest that the ethical action is the one that best protects and respects the moral rights of those affected. This approach starts from the belief that humans have a dignity based on their human nature per se or on their ability to choose freely what they do with their lives. On the basis of such dignity, they have a right to be treated as ends in themselves and not merely as means to other ends. The list of moral rights—including the rights to make one's own choices about what kind of life to lead, to be told the truth, not to be injured, to a degree of privacy, and so on—is widely debated; some argue that non-humans have rights, too. Rights are also often understood as implying duties—in particular, the duty to respect others' rights and dignity.

The Justice Lens

Justice is the idea that each person should be given their due, and what people are due is often interpreted as fair or equal treatment. Equal treatment implies that people should be treated as equals according to some defensible standard such as merit or need, but not necessarily that everyone should be treated in the exact same way in every respect. There are different types of justice that address what people are due in various contexts. These include social justice (structuring the basic institutions of society), distributive justice (distributing benefits and burdens), corrective justice (repairing past injustices), retributive justice (determining how to appropriately punish wrongdoers), and restorative or transformational justice (restoring relationships or transforming social structures as an alternative to criminal punishment).

The Utilitarian Lens

Some ethicists begin by asking, “How will this action impact everyone affected?”—emphasizing the consequences of our actions. Utilitarianism, a results-based approach, says that the ethical action is the one that produces the greatest balance of good over harm for as many stakeholders as possible. It requires an accurate determination of the likelihood of a particular result and its impact. For example, the ethical corporate action, then, is the one that produces the greatest good and does the least harm for all who are affected—customers, employees, shareholders, the community, and the environment. Cost/benefit analysis is another consequentialist approach.

The Common Good Lens

According to the common good approach, life in community is a good in itself and our actions should contribute to that life. This approach suggests that the interlocking relationships of society are the basis of ethical reasoning and that respect and compassion for all others—especially the vulnerable—are requirements of such reasoning. This approach also calls attention to the common conditions that are important to the welfare of everyone—such as clean air and water, a system of laws, effective police and fire departments, health care, a public educational system, or even public recreational areas. Unlike the utilitarian lens, which sums up and aggregates goods for every individual, the common good lens highlights mutual concern for the shared interests of all members of a community.

The Virtue Lens

A very ancient approach to ethics argues that ethical actions ought to be consistent with certain ideal virtues that provide for the full development of our humanity. These virtues are dispositions and habits that enable us to act according to the highest potential of our character and on behalf of values like truth and beauty. Honesty, courage, compassion, generosity, tolerance, love, fidelity, integrity, fairness, self-control, and prudence are all examples of virtues. Virtue ethics asks of any action, "What kind of person will I become if I do this?" or "Is this action consistent with my acting at my best?"

The Care Ethics Lens

Care ethics is rooted in relationships and in the need to listen and respond to individuals in their specific circumstances, rather than merely following rules or calculating utility. It privileges the flourishing of embodied individuals in their relationships and values interdependence, not just independence. It relies on empathy to gain a deep appreciation of the interest, feelings, and viewpoints of each stakeholder, employing care, kindness, compassion, generosity, and a concern for others to resolve ethical conflicts. Care ethics holds that options for resolution must account for the relationships, concerns, and feelings of all stakeholders. Focusing on connecting intimate interpersonal duties to societal duties, an ethics of care might counsel, for example, a more holistic approach to public health policy that considers food security, transportation access, fair wages, housing support, and environmental protection alongside physical health.

Using the Lenses

Each of the lenses introduced above helps us determine what standards of behavior and character traits can be considered right and good. There are still problems to be solved, however.

The first problem is that we may not agree on the content of some of these specific lenses. For example, we may not all agree on the same set of human and civil rights. We may not agree on what constitutes the common good. We may not even agree on what is a good and what is a harm.

The second problem is that the different lenses may lead to different answers to the question "What is ethical?" Nonetheless, each one gives us important insights in the process of deciding what is ethical in a particular circumstance.

Making Decisions

Making good ethical decisions requires a trained sensitivity to ethical issues and a practiced method for exploring the ethical aspects of a decision and weighing the considerations that should impact our choice of a course of action. Having a method for ethical decision-making is essential. When practiced regularly, the method becomes so familiar that we work through it automatically without consulting the specific steps.

The more novel and difficult the ethical choice we face, the more we need to rely on discussion and dialogue with others about the dilemma. Only by careful exploration of the problem, aided by the insights and different perspectives of others, can we make good ethical choices in such situations.

The following framework for ethical decision-making is intended to serve as a practical tool for exploring ethical dilemmas and identifying ethical courses of action.

Appendix 10: The Berkeley Protocol on Digital Open Source Investigations - Professional, Methodological, and Ethical Principles

Source: Directly adapted from OHCHR 7 UC Berkeley, 2022, 11-15; Renshaw & Chaplin, 2020

Professional Principles

1. Accountability: Records of investigations and materials must be kept, and documentation must be clear.

   Open source investigators must be accountable for their actions, which can often be ensured through clear documentation, record-keeping and oversight. Transparency in investigative methods and procedures is an essential element in ensuring accountability. Thus, to the extent possible and reasonable, open source investigators should maintain records of their activities. The steps of an open source investigation – from identification of relevant material through collection, analysis and reporting – should be consistently and clearly documented. Any individuals engaged in the collection or handling of online information should be aware of the potential for their methodology to be questioned, including the possibility of being called to testify at trial. Documentation of open source investigations may be done manually or by using automated processes provided by various software. As long as documentation is consistent and sufficiently thorough, either manual or automatic methods can be used. Automated processes and software must be understood by users and be explainable in court either by users or developers. In addition, open source investigators should record any tools or software used in the course of their work.

2. Competency: Investigators must 'have proper training and technical skills'.

   Open source investigators must have proper training and technical skills to execute the activities in which they engage. They must conduct online activities in a professional and ethical manner, avoiding the appropriation of others' work; crediting all those who participate in an investigation (when safe to do so and when desired by participants); and accurately reporting data, including acknowledging any gaps that may exist in online content. Open source investigators and investigation processes must also remain flexible, stay up to date with new developments and adopt new technologies and techniques as appropriate. In addition, organizations and investigation teams should have mechanisms in place to ensure that procedures are consistently implemented and adhered to.

3. Objectivity: Researchers must be aware of their inherent bias, and take steps to mitigate this.

   Objectivity is a foundational principle that applies to all investigations, whether online or offline. Open source investigators should understand the potential for personal, cultural and structural biases to affect their work and the need to take countermeasures to ensure objectivity. Open source investigators must ensure that they approach their investigations objectively, developing and deploying multiple working hypotheses and not favouring any particular theory to explain their cases. For open source investigations conducted online, objectivity is particularly important because of the way in which information on the Internet is structured and presented to users. The browser, search engine, search terms and syntax used may lead to very different results, even when the underlying query is the same. Inherent biases in the Internet's architecture and algorithms employed by search engines and websites can threaten the objectivity of search results. Search results may

also be influenced by a number of technical factors, including the device used and its location, and the user's prior search history and Internet activity. Open source investigators should counterbalance such biases by applying methodologies to ensure that search results are as diverse as possible, for example, by running multiple search queries and using a variety of search engines and browsers. Investigators should be aware that search results may also be influenced by other factors, including as a result of the discrepancy in the digital environment whereby online information may be unevenly available from certain groups or segments of society. Finally, investigators should always strive to be aware of and correct for their own biases, which may be either conscious or subconscious.

4. Legality: Applicable laws, including data protection and privacy rights, must be complied with.

   Open source investigations should comply with applicable laws, which means that investigators need to have a baseline understanding of the laws that apply to their work. In particular, investigators should be aware of data protection laws and the right to privacy, which is protected under international human rights law.25 Even though information may be publicly available, it does not mean that there are no privacy implications in its collection and use. Open source investigators must consider the privacy implications of their actions, including a person's reasonable expectation of privacy in different digital spaces. Investigators should also be aware of the mosaic effect, whereby public data, even when anonymized, may become vulnerable to reidentification if enough data sets containing similar or complementary information are released or combined.26 In addition, investigators should be aware that, in some jurisdictions, the ongoing and persistent monitoring of individuals online, or the systematic collection and long-term retention of personal data, may require additional permissions and safeguards due to the heightened privacy concerns raised by such activities.

5. Security Awareness: Those 'conducting investigations online should have basic operational security awareness to ensure that they minimise their digital trail and are aware of the potential risks.'

   While security by design addresses the architecture and infrastructure of an investigation and any collateral activities, the principle of security awareness focuses on considerations that individuals must take into account in the course of their work – in particular, awareness of their online behaviour. All individuals conducting investigations online should have basic operational security awareness to ensure that they minimize their digital trail and are aware of the potential risks. Organizations conducting open source investigations should ensure that their investigators are provided with information security training to understand the risks that they may face and have an understanding of the three core pillars of information security: (a) confidentiality (e.g. only allowing permitted users to access data); (b) integrity (ensuring data is not tampered with or otherwise altered by unauthorized users); and (c) availability (ensuring systems and data are available to authorized users when they need it). Training should also focus on the Internet's governance structure. Threat and risk assessments should be conducted before commencing online investigative activities and should be periodically reviewed and amended as necessary. Security is everyone's responsibility, not only the responsibility of information technology units or security risk managers.

Methodological Principles

1. Accuracy: There is a methodological and ethical imperative to ensure the accuracy – and thus the quality – of investigations by only relying on credible materials. Open source investigators should seek to be as truthful and precise as possible in the course of their investigations and in the presentation of any results, especially when it comes to acknowledging weaknesses in the underlying data or the overall case. Accuracy is often improved through the use and testing of multiple working hypotheses and/or peer review, both of which can help minimize the chances of biased selection, interpretation and presentation of data. Analytical conclusions should not be exaggerated or overstated. The use of clear, objective, fact-based language and the avoidance of emotive language will protect the actual and perceived objectivity of an investigation and its results.

2. Data Minimization: The principle of data minimization prescribes that digital information should only be collected and processed if it is: (a) justified for an articulable purpose; (b) necessary for achieving that purpose; and (c) proportional to the ability to fulfil that purpose. In the context of open source investigations, online content should only be collected if it is relevant to a particular investigation. This principle favours itemized, manual collection over bulk, automated collection, while noting that the latter may be appropriate in some cases. Applying this principle to the collection of online content will help avoid over-collection, which is important for several reasons. Over-collection – a particular concern when using automated collection processes – may create or exacerbate security vulnerabilities, in particular if it leads to investigators being unaware of the types of information within their possession. Over-collection may also raise privacy and data protection concerns if an automated process does not discriminate according to the type of content. Finally, avoiding over-collection serves the practical purposes of minimizing storage costs and preventing downstream bottlenecks at various stages of the investigation cycle, such as review, analysis and, in the event that an investigation leads to legal proceedings, disclosure.

3. Preservation: It is just as important to prevent undercollection as it is to avoid over-collection of relevant information. This may be of particular concern in the context of online information, the permanence and availability of which is often precarious. The principle of preservation is designed to avoid undercollection so that relevant and potentially probative evidence is not lost. Social media platforms, for example, may remove content that violates their terms of service even if that content has potential value for investigators. Unless a timely preservation request is made to the platform or content is otherwise preserved by investigators, such information may be lost forever. In addition, users may choose to delete or edit their own content, making once public information unavailable. Furthermore, information on the Internet can be easily decontextualized, lost, erased or corrupted. If digital material is to remain accessible and usable for future accountability mechanisms, it needs to be actively and carefully preserved in both the short and long term.

4. Security by Design: The principle of security by design requires that, to the extent possible, digital information and online operations be secure by default. Organizations conducting online open source investigations should invest in and implement appropriate technical and structural measures to ensure that, by default, infrastructure – including hardware and software – is properly anonymized and non-attributable when investigators go online. All

equipment should have up-to-date software to protect against malware, and appropriate privacy and security settings. Security measures should be in place before online investigative activities commence; they should be continuously monitored, updated and adjusted as needed. Investigators, investigation teams or organizations may want to arrange for ongoing testing, including penetration testing, to ensure that their security systems work as designed.

Ethical Principles

1. Dignity: Investigations should be conducted with an awareness of and sensitivity to any underlying dignity-related issues, especially those interests that are protected by international human rights law. For example, investigators should adhere to the principles of non-discrimination, which may affect what gets investigated and who does the investigating or is credited with the investigation, and integrate safeguards concerning the digital, physical and psychosocial security of witnesses, survivors, other investigators, those accused and others who may be negatively affected. Adherence to the principle of dignity may also affect what is shared publicly about an investigation, including in writing and in any visual materials – for example, not showing the full extent of suffering or violence if it is not necessary to do so. This principle ensures that human rights norms are a guiding set of standards for conducting ethical open source investigations.

2. Humility: Open source investigators should be humble, recognizing their own limitations and having an awareness of what they do not know. Proper understanding and interpretation of open source information may require specialized training or consultations with experts. Humility also means taking responsibility for errors. If investigators find that they have made an error, that error should be corrected or reported to those who can minimize the resulting harm. Ideally, there should be a mechanism to report errors and for corrections to be issued, especially for investigations that are public and widely distributed.

3. Inclusivity: Open source investigators must ensure that a range of perspectives and experiences are incorporated into investigations. Factors to consider that may influence the overall inclusivity of an online investigation include its geographic scope, the violations and/or international crimes being investigated and an awareness of the uneven nature of online information with respect to different segments of society. Investigation teams should also be diverse, which includes having a gender balance. In addition, the principle of inclusivity, together with the principle of dignity, may affect the materials an investigator chooses to collect and use in an investigation and how they are presented to different audiences.

4. Independence: Open source investigators should protect themselves and their investigations from inappropriate influence. They should identify and avoid any real or perceived conflicts of interest and put in place safeguards to mitigate those conflicts that cannot be avoided. Transparency of process, methods and funding can help with assessments of independence and protect the actual and perceived independence of an investigation.

5. Transparency: While the principle of accountability requires transparency in an investigator's methods and results, the ethical principle of transparency refers to how open source investigators conduct themselves online and to the outside world. This means avoiding misrepresentation. While anonymity and non-attribution – including the use of virtual

identities – can be important for security reasons, investigators should be aware of the potential negative ramifications of misrepresentation, such as damaging the reputation and credibility of an investigation, team or organization, or contaminating the information collected. Procuring information through misrepresentation may violate a targeted individual's right to privacy and/ or taint an investigation, especially if the misrepresentation is illegal in the relevant jurisdiction(s).

Appendix 11: Department of National Defence – Response[162]

s.15(1)
s.21(1)(a)
s.21(1)(b)
s.23

SECRET//REL CAN,

CLASSIFICATION: SECRET

BRIEFING NOTE FOR THE DEPUTY MINISTER AND CHIEF OF THE DEFENCE STAFF

OPEN SOURCE INTELLIGENCE ACTIVITIES ON THE INTERNET –

Executive Summary

-
-
-
-
-
-

Issue

1. To advise the DM and CDS

Background

2. CDI is in the process of revising a functional directive on the conduct of Open Source Intelligence Activities on the Internet (OAI). The previous directive, issued in February 2017, is outdated.

The draft directive is enclosed at **Tab A**.

SECRET//REL CAN,

[162] Please note that while the overall classification of this document is SECRET, all SECRET information has been removed from this document by the Department of National Defence and released at an UNCLASSIFIED level under the Access to Information Act, as indicated in the top right corner of the document.

s.15(1)
s.21(1)(a)
s.21(1)(b)
s.23

CLASSIFICATION: SECRET

15.

16.

17.

Recommendation

18.

Prepared by:
Reviewed by:
Maj. S. O'Blenes, D Law I&IO, JAG, 992-0011
Ms. C. Sully, PLL, CF LA, 943-4018
Responsible DG: Mr. P. Hébert, DGIPP, 945-7658
Responsible Commander: RAdm S.E.G. Bishop, Comd CFINTCOM, 945-5120

Date: 06 Mar 20

Enclosures: 3

Chief of Defence Intelligence Functional Directive: Conduct of Open Source Intelligence Activities on the Internet

1.0 IDENTIFICATION

File Number:

Effective Date: 1 November 2019

Supersedes: *Chief of Defence Intelligence Functional Directive: Framework for the Conduct of Intelligence Activities Using the Internet*, February 2017

Office of Primary Responsibility: Director General Intelligence Policy and Partnerships (DGIPP)

Approval Authority: Chief of Defence Intelligence (CDI)

References:

A. *Charter of Rights and Freedoms*
B. *National Defence Act*
C. *Privacy Act*
D. *Communications Security Establishment Act*
E. *Treasury Board Directive on Privacy Practices*
F. *Ministerial Directive on Defence Intelligence*, September 2013
G. *Ministerial Directive on Defence Intelligence Priorities*, June 2019
H. *Ministerial Directive on Avoiding Complicity in Mistreatment by Foreign Entities*, November 2017
I. DAOD 1000-10, *Policy Framework for Corporate Administration Management*
J. DAOD 1002-0, *Administration of the Privacy Act*
K. DAOD 1002-1, *Privacy Act Requests and Correction of Personal Information*
L. DAOD 1002-3, *Personal Information Management*
M. *Interim CDI Functional Directive on Intelligence and Intelligence-derived Information Sharing with External Entities*, April 2018
N. *CDI Functional Directive: Guidance on the Collection of Canadian Citizen Information*, August 2018

PURPOSE

APPLICATION

1.2 (U) This CDI directive is an order for all officers and non-commissioned members of the Canadian Armed Forces (CAF), and a directive for all employees of the Department of National Defence (DND).

APPROVAL AUTHORITY

1.3 (U) This Functional Directive is issued on the authority of the CDI, under the delegated authority of the Deputy Minister (DM) and the Chief of the Defence Staff (CDS) to issue functional direction in respect of defence intelligence matters in accordance with refs F (*Ministerial Directive on Defence Intelligence*) and I (DAOD 1000-10, *Policy Framework for Corporate Administration Management*).

OFFICE OF PRIMARY RESPONSIBILITY

1.4 (U) DGIPP is the Office of Primary Responsibility for this directive, and all policy issues pertaining to Open Source Intelligence (OSINT) when conducting defence intelligence missions and/or activities.

TABLE OF CONTENTS

1.5 This document contains the following subjects:

s.15(1)
s.21(1)(a)

2.0 DEFINITIONS

9.0 Approval

9.1 (U) This functional direction takes immediate effect and shall remain in effect until otherwise directed.

S.E.G. Bishop
Rear-Admiral
Chief of Defence Intelligence

[Month] 2020

Appendix 12: Transport Canada - Response

*Original document available upon request.

Analytic and Critical Thinking for Intelligence Analysts (ACT) - *Online*	**4 days, $550** **8 one-hour sessions**
Thinking and reasoning are essential to effective and incisive intelligence analysis. This course helps develop analysts' reasoning, evaluation and interpretation skills to help them become more effective and confident intelligence analysts. This is an academically-oriented course, taught in an interactive lecture format. Concepts and techniques in analytic and critical thinking are presented in a manner that relates them to concepts, practices and methodologies used in intelligence. Active student participation is central to getting the most out of this course. Students will cover 21 different learning modules spread over 8 classes designed to develop their reasoning, evaluating and interpreting skills. You will also learn and practice several structured analytic techniques. This course is taught over 4 days. Each day has two one hour classes separated by a 30 minute health break: Day 1 Class 1: Situating Analytic and Critical Thinking Class 2: Fundamental Reasoning Day 2 Class 3: Fallacies: Reduce the influence of faulty reasoning Class 4: Assumptions & Cognitive Biases: Reduce the influence of faulty reasoning Day 3 Class 5: Evaluate claims that others make Class 6: Interpreting Meaning Day 4 Class 7: Consider the Perspectives of Others Class 8: Inferring Mean from Patterns and Evidence The course material is progressive with each class building on what students learned in previous modules. Students who will miss more than one class may be required to make up the missed classes in another course serial to complete this course. There are no prerequisites required to take Analytic and Critical Thinking for Intelligence Analysts. Likewise, this course is not a prerequisite for any other IALP course. However, completing this course can provide intelligence analysts with a greater depth of understanding of topics addressed in some other courses. **Participant must ensure they have access to Webex on their device.**	*May 17-20* *June 7-10*
Introduction to Canadian Intelligence (ICI) - *Online*	**2 days, $550** **4 one-hour sessions**
This course introduces the history and structure of the Canadian intelligence community as well as basic vocabulary for working with intelligence analysis and classified materials. It is meant for those working with intelligence analysts perhaps for the first time, such as managers, policy analysts or support staff. *This course is not intended for intelligence analysts. It introduces information covered in more detail during the Entry-Level Course (ELC).* **Participant must ensure they have access to Webex on their device.**	*May 26-27* *June 23-24*
OSINT Level 1 - *Online*	**5 days, $550** **10 one-hour sessions**
Learn strategies and tactics for collecting and exploiting open source information. Go beyond standard search tools, to find new sources of information or access it faster; evaluate online information sources, use online graphics and social media, and maintain online operational security. **Participant must ensure they have access to Webex on their device.**	*May 9-13* *June 6-10* *June 20-24 (French)*

OSINT Level 2 - *Online*	**5 days, $550** **10 one-hour sessions**
Course Objective: Participants will learn to plan and use efficient search strategies to address OSINT researches as it relates to intelligence problems. **Target Audience**: Intelligence analysts who have taken the introductory (level 1) course. **Key Course Topics & Generic Activities**: • Collection planning for the Internet • Advanced security • Ethics • Case management (chain of evidence, best evidences, computer preparation, etc.) • Browsers and extensions • Research strategies • Target-surface mapping • Sock puppet accounts • Internet of Things (IoT) and CCTVs • Identifying people and associates • ID codes embedded in Websites • Social media newcomers • Corroborating photos and videos • Identifying locations • Identifying times • Extracting information from e-mails headers Many short online exercises will challenge analysts to apply techniques presented in a controlled environment. **Required Resources**: • Online access (Including WebEx) • Presentation (PPT) • Handouts (electronic handouts) **Participant must ensure they have access to Webex on their device.**	*May 16-20* *June 13-17*
Effective Writing Course (EWC) - *Online*	**5 days, $550** **10 one-hour sessions**
Intelligence analysts seek to communicate effectively with their clients through providing clear, concise and client-focused reporting to support timely decision-making. Consumers of intelligence, however, have limited time and patience, and need to be able to absorb critical information and ideas quickly. Using 'hands-on', practical activities, this course aims to provide new and practicing intelligence analysts with techniques and strategies to convey their written message effectively and efficiently. The course will combine virtual instructor-led training sessions in our virtual classroom with practical 'take home' writing exercises and activities completed on your own time and emailed to the instructor for review and feedback. At the end of the course, participants will have a better understanding of the intelligence writing style, how to structure written intelligence products, and how to focus their writing for better understanding and support to their client's decision making. **Participant must ensure they have access to Webex on their device.**	*To be announced*

Appendix 13: Financial Transactions and Reports Analysis Centre - Response Email

Good afternoon Ms. Cioffi,

This is further to your request, for which we received your clarification on December 12, 2022, and submitted under the *Access to Information Act* (the Act) for:

"1. Please provide a list of all training available to/necessary for FINTRAC analysts to conduct open source intelligence analysis to conduct research into trends and developments in the area of money laundering and the financing of terrorist activities and improved ways of detecting, preventing and deterring money laundering and the financing of terrorist activities.

2. What is the overall training budget per fiscal year for intelligence analysts.

3. Please provide any policy document that governs the policy and authorities surrounding OSINT / PAI collection?

The time frame for these documents is between January 1, 2017 and the date of clarification so December 2022."

Pursuant to subsection 4(3) of the Act, FINTRAC may produce a record from its data holdings (using regularly available information systems / tools and expertise) in circumstances where the record does not already exist. However, this obligation is subject to the limitations set out in section 3 of the Act's Regulations which states that FINTRAC does not have to do so if the effort would unreasonably interfere with its operations. As such, for such situations FINTRAC must assess the level of effort required to produce the information.

No responsive records existed for your request and, based on an assessment as described above, it was determined reasonable to produce the information to respond to your request. The produced records are enclosed and, following a review in accordance with the *Act*, we are pleased to advise you that they are being disclosed in their entirety.

Please be advised that you are entitled to complain to the Information Commissioner concerning the processing of your request within sixty days of the receipt of this notice. In the event you decide to avail yourself of this right, your notice of complain should be addressed to: The Office of the Information Commissioner of Canada, 30 Victoria Street, Gatineau, Quebec, K1A 1H3 / 1-800-267-0441 (Toll Free).

Should you have any questions, please do not hesitate to contact Alexandre Dery by email at atip-aiprp@fintract-camafe.gc.ca or at the address noted above. When communicating with us, we would appreciate it if you would quote the file number we have assigned to your request.

Yours sincerely,

John Widdis (he / him / il / lui)
Access to Information and Privacy Coordinator
Financial Transactions and Reports Analysis Centre of Canada

234 Laurier Avenue West, Ottawa, On, K1P 1H7
Government of Canada
ATIP-AIPRP@fintract-canafe.gc.ca |

Please do not hesitate to respond in the official language of your choice. | I am located on the traditional, unceded territory of the Anishinaabe Algonquin Peoples.

Financial Transactions and Reports Analysis Centre - Email Attachment

*Original document available upon request

A-2022-00065

Training for FINTRAC analysts to conduct open source intelligence analysis.

1. Please provide a list of all [training] **available to/necessary for FINTRACT analysts to conduct open source intelligence analysis to conduct research into trends and developments in the area of money laundering and the financing of terrorist activities and improved ways of detecting, preventing, and deterring money laundering and the financing of terrorist activities.**

Strategic intelligence Analysts undergo internal introductory training, developed and conducted by more experienced strategies intelligence analysts, related to the development of and concepts associated with research into trends and developments in the area of money laundering and the financing of terrorist activities and improved ways of detecting, preventing and deterring money laundering and the financing of terrorist activities. Strategic intelligence analysts leverage external training sources, both within the public sector and offered by the private sector subject-matter experts and formal certification processes, which includes specific courses on Open Source Intelligence.

Moreover, the development of research and strategic intelligence – including those leveraging publically available information – involves multiple steps and oversight. It is conducted by a strategic intelligence analyst under the guidance and mentorship of a team lead, goes through peer review, and must receive managerial approval before proceeding through appropriate internal governance for endorsement.

2. What is the overall training budget per fiscal year for intelligence analysts.

Within the Strategic Intelligence, Research and Analytics unit, $43,750.67 has been forecasted as the total expenditures for the cost of training courses for members of the strategic intelligence and research teams for FY2022-23. Of this amount, some $2,750 was allocated to introductory and intermediate courses on the subject of Open Source Intelligence specifically.

The Intelligence Sector has provided the following budget figures for training:
FY 2017-2018 to FY 2022-23, the budget allocated to training per fiscal year is as follows,

FY 2017-18: $ 107, 550.14
FY 2018-19: $ 62,289.49
FY 2019-20: $ 62,481.82
FY 2020-21: $ 81,708.62
FY 2021-22: $ 92,190.29
FY 2022-23: $ 82,211.38

3. Please provide any policy documents that governs the policy and authorities surrounding OSINT / PAI collection?

The authorities related to the collection of publically available information are found in section 54(1)(b)(i) of the *Proceeds of Crime (Money Laundering) and Terrorist Financing Act*. https://laws-lois.justice.gc.ca/eng/acts/P-24.501/page-8.html#h-399036

54(1) The Centre
(b) may collect information that the Centre considers relevant to money laundering activities or the financing of terrorist activities and that
(i) is publicly available, including in a commercially available database

Appendix 14: Public Safety Canada - Response

*Original document available upon request.

Hybrid CoE: Mission

How Centre achieves its mission:
The Centre serves as a **networked hub and integrator of expertise** with the key task being to build participating states' capabilities to prevent and counter hybrid threats.

- This networking and integration is designed to share best practices, testing new ideas and approaches, and providing training courses and exercises.

- This occurs in workshops, seminars, webinars comprised of PS identified practitioners (government experts, planners, policy and decision makers,), academics/scholars and the private sector.

- The results are reported and disseminated to PS through the networks of the PS Points of Contact.

- The Centre also Identifies and applies external capabilities as provided by PS (Examples: Elections Training from US Global Environment Centre; Open Source Intelligence (OSINT) Training from UK; Wargaming from US Naval Institute.

- The Centre's also aims to lead the conversation on hybrid threats by independent research and analysis, and publishing a wide variety of publications and engaging with various partners in the field through conferences and workshops.

Intelligence Analyst Learning Program

OSINT Level 1	In-Person 2 full days Online 5 days 10 one-hour sessions
Learn strategies and tactics for collecting and exploiting open source information. Go beyond standard search tools to find new sources of information or access it faster; evaluate online information sources, use online graphics and social media, and maintain online operational security. **Online participants must ensure they have access to Webex on their device.**	In-Person *October 3-4* *October 17-18* *December 5-6* Online *November 14-18*
OSINT Level 2	Online 5 days 10 one-hour sessions
OSINT Level 1 or equivalent experience is a prerequisite for this course Participants will learn to plan and use efficient search strategies to address OSINT researches as it relates to intelligence problems. Target Audience: Intelligence analysts who have taken the introductory (level 1) course or an equivalent course. The course focuses on four main analytical functions: 1) Identifying, 2) Locating, 3) Timing; and, 4) Verifying Many short online exercises will challenge analysts to apply techniques presented in a controlled environment. A final exercise (@1h30) summaries the learning. **Online participants must ensure they have access to Webex on their device.**	Online *December 12-16*

Appendix 15: Canadian Security Intelligence Service – Response

*Original document available upon request.

Our file: 117-2022-802 June 16, 2023
Dear Giovanna Cioffi,

This refers to your Access to Information Act request of November 9, 2022, for "**Documents such as memos, briefing notes, policies, training guides, and reference materials which instruct analysts on the use of open source intelligence (OSINT), Documents dated 2017 – present."**

Please be advised that the Service has now processed your request, however, all the information requested has been exempted from disclosure by virtue of one or more of sections 15(1) (as it relates to the efforts of Canada towards detecting, preventing or suppressing subversion or hostile activities), 16(1)(c), 19(1) or 24(1) of the *Act*.

With regards to the information exempted pursuant to subsection 19(1), I wish to inform you that the disclosure provisions contained in subsection 19(2) were considered however, none of them applied.

You may use the contact information located in the footer to contact us should you wish to obtain clarification concerning your request. Please provide the file number at the top of this letter in any subsequent correspondence.

Please be advised that you are entitled to file a complaint to the Information Commissioner concerning the processing of your request within sixty days of the receipt of this notice. In the event you decide to avail yourself of this right, your notice of complaint should be addressed to **Information Commissioner of Canada, 30 Victoria Street, Gatineau, Quebec, K1A 1H3**. You may also submit your notice at **https://www.oic-ci.gc.ca/en/submitting-complaint**.

Yours truly,

Coordinator
Access to Information and Privacy