RHODES UNIVERSITY
Where leaders learn

A Framework for using Open Source Intelligence as a Digital Forensic Investigative Tool

A thesis submitted in partial fulfillment of the requirements for the degree
of
Master of Science at Rhodes University

By

Samantha Elizabeth Rule

December 2014

Abstract

The proliferation of the Internet has amplified the use of social networking sites by creating a platform that encourages individuals to share information. As a result there is a wealth of information that is publically and easily accessible. This research explores whether open source intelligence (OSINT), which is freely available, could be used as a digital forensic investigative tool.

A survey was created and sent to digital forensic investigators to establish whether they currently use OSINT when performing investigations. The survey results confirm that OSINT is being used by digital forensic investigators when performing investigations but there are currently no guidelines or frameworks available to support the use thereof. Additionally, the survey results showed a belief amongst those surveyed that evidence gleaned from OSINT sources is considered supplementary rather than evidentiary.

The findings of this research led to the development of a framework that identifies and recommends key processes to follow when conducting OSINT investigations. The framework can assist digital forensic investigators to follow a structured and rigorous process, which may lead to the unanimous acceptance of information obtained via OSINT sources as evidentiary rather than supplementary in the near future.

Acknowledgements

First and foremost, I would like to thank Roshan Harneker, for believing in me and encouraging me to pursue a Master of Science degree. Without your support, this thesis would never have been possible.

I would also like to express my sincere gratitude to my supervisor, Dr Karen Bradshaw, for her guidance and support throughout the research process and multiple rewrites. Your endless patience and encouragement are truly appreciated.

Thanks also goes to my colleague and friend, Mr Timothy Haig-Smith, for his support and encouragement.

I would also like to thank the survey respondents, who participated in this study.

Lastly, I would like to thank my loved ones, who have supported me throughout the entire process.

Table of Contents

List of Tables

List of Figures

Chapter One Introduction

"*Publication is self-invasion of privacy.*" (Marshall McLuhan)

The proliferation of the Internet and the rate at which technology is evolving, together with the fact that social networking is currently one of the most popular online activities, ensures that cybercrime is likely to increase at a rapid rate (NW3C, 2013). Across the globe, organised crime groups are making use of technology to communicate and commit crimes, which in turn creates many challenges for law enforcement, forensic investigators, corporate security professionals and members of the legal fraternity (Casey, 2004).

Since digital forensics plays such a pivotal role in the research topic, it is also important to understand what digital forensics is. When asked what digital forensics is, one of the speakers at the 2004 Digital Forensic Research Workshop (DFRWS) Mark Pollitt, described digital forensics as a process consisting of a number of tasks and processes that take place during an investigation (Politt, 2004).

Participants at the DFRWS view digital forensics as a science and define it as "the use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources" (Palmer, 2001). These methods are used to facilitate and assist in reconstructing how potential criminal events took place and possibly in anticipating unauthorized actions shown to be disruptive (Palmer, 2001).

Current digital forensic investigation tools and procedures are focused primarily on traditional forensics, which in turn is focused on evidence preservation and usually involves the imaging of a physical device, such as a hard drive or mobile phone (Garfinkel, Farrell, Roussev, & Dinolt, 2009).

Kohn, Eloff, & Olivier (2006) state that the majority of digital forensic investigation models are generic and have been built from experience with existing models. Thus, some of the investigation models have similar techniques but focus on different areas of an investigation. Zainudin, Merabti, & Llewellyn-Jones (2011) report that there is no standard investigation model specifically developed for social networking, despite crimes committed via social media increasing steadily.

Over the last few years, a new format of online communication has evolved, known as social networking (Mutawa, Baggili, & Marrington, 2012). Social networking sites allow users to interact with each other and encourage their users to share information, especially personal information such as age, location, interests and personal social activities. These sites also allow users to post comments, photographs, and videos, and engage in real time conversations with one another.

Although the intended use of social networking is to enable friends to communicate and socialise online, criminals have spotted vulnerabilities in social networks, which they are exploiting owing to the anonymity and abundance of freely available personal information allowed by these social networks. The wealth of information posted on the social networks, enables criminals to access relevant information easily and use it to their advantage to commit crimes.

By design the Internet is "public" thereby ensuring that a vast amount of information is available to anyone who has access to the Internet via a computer or mobile device (Appel, 2011). Open Source Intelligence (OSINT) is publically available information that has been obtained both legally and ethically. OSINT is the earliest form of intelligence and originated well before the Internet came into existence at a time when information was gathered from newspapers, speeches and radio.

Publically available information is collected and exploited to produce intelligence for a specific targeted audience such as military or law enforcement (Pouchard, Dobson, & Trien, 2006). Hulnick (2010) states that OSINT is the lifeblood of intelligence. In the United States, OSINT accounts for 70% – 80% of all intelligence and the earliest record of OSINT dates back to the Second World War. The vast amount of OSINT information available is, however, a challenge for many intelligence analysts as they have to traverse through the ocean of information to locate that which is relevant to them.

1.1 Problem Statement

The traditional digital forensic investigation frameworks do not offer support for using OSINT sources, specifically social networking because current frameworks focus mainly on acquiring a physical device from which to obtain the information. Technology and the use of the Internet are rapidly evolving, while adults and youths alike are increasingly making use of social networking. Social networking sites are one of the fastest growing online communication tools where individuals are encouraged to publish personal information. To ensure that this wealth of information available to digital forensic investigators is not disregarded, but rather embraced, a framework for using OSINT sources, such as social networking, needs to be developed.

1.2 Research Question and Objectives

The following research question was addressed in this research: Can OSINT be used as a digital forensic investigation tool and will the information obtained be admissible as evidence in a court of law?

To answer this question, the following sub-objectives need to be achieved:

1. Ascertain if the information available via social networking and OSINT could indeed assist a digital forensic investigator during an investigation.
2. Create a framework that can be used by digital forensic investigators when conducting an investigation that requires partial or full use of social networking sites to obtain information.
3. Determine criteria for websites that can be considered social networking sites, and identify the types of information or evidence items that can be obtained from these sites.
4. Address any shortcomings or obstacles that a digital forensic investigator may encounter when making use of social networking sites to conduct a full or partial forensic investigation.

With regard to the research question, the researcher hypothesises that there is legitimate value in the information available from OSINT, and this can be used to assist digital forensic investigators.

1.3 Significance of the Research

The LexisNexis study "Law Enforcement Personnel Use of Social Media in Investigations: Summary of Findings", conducted in 2012, found that four out of five law enforcement officers used social media for investigative purposes (LexisNexis, 2012). Since this survey involved 1221 law enforcement officers, it is apparent that social media is already being used extensively in criminal investigations.

The purpose of creating a framework for using OSINT sources as a digital investigative tool is to provide digital forensic investigators with guidelines to assist them when using OSINT sources as a tool during their investigations. The OSINT information, which is publically available, is of great value to any digital forensic investigator but guidance is required to assist an investigator to ensure that the information acquired from the OSINT source will ultimately be admissible as evidence in a court of law.

1.4 Limitations of the Research

The main limitation is the small number of digital forensics practitioners who participated in the survey. A total of 75 digital forensic practitioners were contacted but only 18 responses were received despite two follow-up emails being sent requesting the recipients' participation in the survey.

Despite this limitation, the researcher believes that overall, the relevance of the research documented in this thesis is still valid as the information gleaned from the forensic practitioners was merely used to establish the status quo of OSINT use in digital forensic investigations as a starting point for developing the framework for conducting investigations.

1.5 Thesis Structure

The remainder of the thesis is arranged as follows:

Chapter 2 focuses on the three main areas of the research:

- digital forensics and the current traditional investigation frameworks,

- social networking and the wealth of information available, some examples of social media platforms and the information that is available on each, and
- OSINT and the vast volume of information that can be obtained from various OSINT sources and how it is currently being used.

Chapter 3 describes the research methodology, the participants' demographics and how the data from the electronic survey was analysed.

Chapter 4 presents the results of the research undertaken and the interpretation thereof. A comparison of the results from the electronic survey and the literature survey is also provided and discussed.

Chapter 5 describes the proposed framework to be used when conducting a digital forensic investigation combined with using OSINT sources based on the findings from Chapter 4.

Chapter 6 concludes the research with a summary of the findings from the electronic survey and the proposed framework to be used when performing a digital forensic investigation incorporating OSINT sources. Future research areas are also presented in this final chapter.

1.6 Summary

In this chapter the researcher stated the research problem, discussed the objectives of the research and explained the significance of the research. Additionally, the researcher outlined reasons for the research being conducted. The limitations of the research were acknowledged and finally a brief outline of the structure of the thesis was given.

Chapter Two Literature Review

This chapter reviews in detail the literature upon which the research is based and consists of three distinct parts:

Section 2.1 defines what digital forensics is, examines some of the current investigative processes and reviews digital evidence.

Section 2.2 defines social networking, reviews how social networking is being used for communication as well as how law enforcement is currently using social networking when performing investigations.

Section 2.3 defines OSINT, examines the advantages of OSINT and how OSINT is used in the private sector and by law enforcement.

2.1 Digital Forensics

2.1.1 Definition

In 2001, the first DFRWS took place. The objectives of the workshop were to bring academics and digital forensic practitioners together to form a community that could assist in defining the discipline and help identify the challenges that lay ahead (Palmer, 2001).

At the first DFRWS, digital forensics was formally defined as a science and described as "the use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources" (Palmer, 2001). Digital forensics can also be defined as the identification, preservation and analysis of digital evidence while following a sound methodology that will be legally accepted (Koen, 2009).

Dixon (2005) describes the goal of digital forensics as the identification, preservation, extraction, documentation and finally the interpretation of digital data. Sound forensic practices must always be followed since, without proper procedures, the forensic evidence will not stand up in a court of law.

Forensic science is defined as the “application of science to law” (Casey, 2004). This means that forensic science is ultimately for use in a court of law. Forensic science provides an all-encompassing body of proven investigative techniques and methodologies that digital forensic investigators use when conducting a digital forensic investigation involving electronic evidence (Casey, 2004).

2.1.2 Investigative Process

During the DFRWS in 2001 an investigative process was established including the following categories or steps which experts asserted should be adhered to when conducting a digital forensic investigation (Palmer, 2001):

- Identification
- Preservation
- Collection
- Examination
- Analysis
- Presentation

Ieong (2006) defines the fundamental principles of a digital forensic investigation as reconnaissance, reliability, and relevance. The digital forensic investigator needs to perform a reconnaissance to understand how the evidence can be collected and analysed using various tools, investigative methods and practices. Reliability involves ensuring that the chain of evidence is always preserved during the evidence extraction, thereby always protecting the integrity of the evidence. Finally, the digital forensic investigator should always ensure that the evidence collected and extracted is relevant to the case.

In 2011, Casey revised the most common steps for conducting a comprehensive digital forensic investigation and suggested that the following steps should be followed (Casey, 2011):

- Preparation: The first step is to create a plan to perform a digital forensic investigation and obtain all the required support documentation, forensic tools and hardware before commencing with the actual investigation.

- Identification: This step identifies and surveys all possible sources of digital evidence. This includes obtaining evidence from all possible devices at the crime scene such as routers with built-in storage, mobile devices, and many others. This also pertains to evidence that can be obtained via Internet services accessed on these devices.
- Preservation: This step ensures that all possible changes to the digital evidence are prevented. This step also extends to the collection of the evidence.
- Examination and Analysis: This step involves searching and analysing the evidence and has been described as the "application of a scientific method and critical thinking" (Casey, 2011, p. 39) to be able to answer the fundamental questions of any investigation as stated by Ieong in the section below.
- Presentation: This step involves drafting a report of the investigation findings that will stand up to legal scrutiny.

Ieong (2006) affirms that a digital forensic investigation is the process of determining and relating extracted information and digital evidence to produce accurate factual information for review by a court of law. Ieong (2006) further suggests that the digital forensic investigation procedures developed by traditional forensic scientists focused on the procedures of handling evidence, while procedures developed by technologists focused on the technical details of capturing the evidence. Ioeng (2006) asserts that many digital forensic investigators have chosen to follow the technical procedures and have forgotten about the purpose and core concept of a digital forensic investigation. Legal practitioners sometimes have difficulty in understanding or applying their processes and tasks in digital investigations owing to the complicated procedures and technical details that many digital forensic investigators follow.

Ieong (2006) further provides six key questions that an investigator should ask while performing an investigation:

- *What?* (the data)
- *Why?* (the motivation behind the cybercrime)
- *How?* (the procedures undertaken by the cybercriminal)
- *Who?* (the people affected and involved)
- *Where?* (the location)

- *When?* (the time when the cybercrime took place)

The Oxford dictionary defines a framework as a simple structure supporting or underlying a system or concept. Kohn et al. (2006) describe a computer forensic framework as a structure to support a forensic investigation. Using a forensic framework while performing a digital forensic investigation ensures that the conclusion reached by one digital forensic investigator is more likely to be the same as that by any other digital forensic professional who conducted the same investigation (Von Solms, Lourens, Reekie, & Grobler, 2006).

2.1.3 Digital Evidence

Owing to the increase in the proliferation of digital devices and the explosion of the Internet, digital evidence is likely to be present in almost any crime that is committed (Jordaan, 2012). Peisert, Sishop, & Marzullo (2008) state that more than half of the cases investigated by the Federal Bureau of Investigation (FBI) in the United States involve the use of digital evidence.

Owing to the widespread use of the Internet, contextual information that is available from the Internet can be extremely useful and important for digital forensic investigations. Information that can be obtained from the Internet includes information on individuals relating to who is connecting to which computer devices and what tasks or activities they are performing (Adelstein, 2006).

Digital evidence is defined as any data that can establish that a crime has taken place or provide a link between a crime and the victim or the crime and the perpetrator (Casey, 2004). The Scientific Working Group on Digital Evidence (SWGDE) defines digital evidence as any “information that is useful and of sufficient value that is either stored or transmitted in a digital form”.[1]

While conducting a forensic investigation, and evaluating the evidence for reliability and accuracy, it is imperative that the evidence is accurate and reliable due to the

[1] https://swgde.org/

impact that the findings can have for an individual (Casey, 2002). Jordaan (2012) describes a key factor for any court case being the quality of evidence and that there is a need to assure the quality thereof to improve its value for use in any legal system.

Arthur (2010) reports that levels of certainty can be associated with evidence sources and digital evidence. An evidence certainty scale was developed to address the uncertainties related to digital evidence by allowing certainty assessments to be associated with digital evidence (Arthur, 2010).

Owing to the inconsistencies in the use of terminology by digital forensic investigators when describing the level of certainty for a particular evidence finding, Casey (2011) proposed a scale for categorizing levels of certainty when handling digital evidence. The evidence certainty scale consists of seven levels, C0 to C6. This scale is depicted in Table 1.

Evidence that is labelled C0 is contradictory to the known facts, while evidence at level C6 is tamperproof and contains a high level of assurance. Casey's certainty scale therefore allows digital forensic investigators to identify the level of confidence in the evidence finding. It is clear, therefore, that if the evidence finding has a low level of confidence the digital forensic investigator may not be able to conclude the finding without additional corroborating information.

There are both advantages and disadvantages to the certainty scale. One of the advantages is that it is flexible and can be used to measure the evidential weight of both the process that generated the evidence and the evidence finding itself such as a Microsoft Word document. The one major disadvantage of the certainty scale is that it is subjective since digital forensic investigators will need to use their own judgment when assessing the certainty levels of evidence (Casey, 2011).

Kohn et al. (2006) report that the overriding goal of any digital forensic investigation is to ensure that concrete evidence is produced, which can eventually be presented in a court of law.

Table 1: Casey's Evidence Certainty Scale (taken from Casey, 2011)

Certainty Level	Description	Qualification
C0	Evidence is contradictory to known facts.	Incorrect
C1	Evidence is extremely questionable.	Highly uncertain
C2	One source of evidence is tamperproof.	Somewhat uncertain
C3	The evidence contains some irregularities even though the evidence sources are difficult to tamper with.	Possible
C4	Multiple independent sources concur that either the evidence is protected against tampering or the evidence is not protected against tampering.	Probable
C5	Multiple independent sources that are protected from tampering concur, although a minute uncertainty still exists.	Almost certain
C6	Evidence is tamperproof and contains a high level of assurance.	Certain

2.1.4 Testing of Digital Evidence

Carrier (2002) devised four categories for use when assessing a digital forensic procedure known as the Daubert test, which comprises generally accepted guidelines for evaluating scientific evidence including the potential rate of error (Casey, 2002). This test is an expansion of the previous United States of America court's approach to the admission of scientific evidence.

The four categories in the Daubert test are:

- Testing: This addresses whether the procedure can be and has been tested.

- Error Rate: This questions whether there is a known error rate for the procedure.
- Publication: This asks if the procedure has been published and subjected to peer review.
- Acceptance: This relates to whether this procedure is generally accepted in the relevant scientific community.

When collecting digital evidence, digital forensic investigators must ensure that only specific forensically-sound tools and techniques are used in order to maintain the integrity of the evidence and ensure that the evidence will be admissible in a court of law (Koen, 2009). Adelstein (2006) concurs that the integrity of digital evidence must be guaranteed during the entire forensic investigation and further asserts that hash sums should be calculated on the evidence source system and the extracted evidence. These should then be compared to ensure the evidence authenticity and integrity.

Altheide & Carvey (2011) explain that a key activity performed during a forensic investigation is the creation of a cryptographic hash. ”A cryptographic hash function takes an arbitrary amount of data as an input and returns a fixed size string as output and the resulting value is a hash” (Altheide & Carvey, 2011, p. 56). Hashing usually takes place during the verification phase of the disk imaging process, as any modification, even to a single bit of data, will produce a completely different hash value. This means that a hash generated of the source drive can be compared with the hash of the forensic image and if the hashes match, this confirms that the two items are exactly the same (Altheide & Carvey, 2011).

2.1.5 Live Acquisition

Adelstein (2006) states that the nature of digital forensic investigations is changing, as the traditional approach to performing a digital forensic investigation has several disadvantages. Traditional digital forensic investigations strive to preserve all hard disk evidence in a forensically sound state, which means that no data is changed and is thus admissible to a court of law. Moore’s law is a metric that reflects the rapid advancement of computer technology and asserts that computer technology will double every two years (Glassman, 2012). As a result of Moore’s law, it is not always possible to perform a traditional forensic investigation owing to the increase in hard drive capacity, large RAID arrays and network accessible storage units (Choo, Smith, & McCusker, 2007). It is therefore sometimes necessary to perform a live system acquisition making use of live digital forensic techniques. Over and above the challenges listed above, cloud computing also does not allow for a traditional investigative approach due to the remote evidence and lack of physical access (Dykstra & Sherman, 2012).

A live forensic acquisition captures a snapshot of the computer's state, which cannot be reproduced at a later stage (Adelstein, 2006). Live system acquisitions are becoming a critical part of the digital forensic investigative process as these enable investigators to capture the volatile system state. A number of forensic toolkits are available to assist with automatic live acquisition (Cohen, Bilby, & Caronni, 2011).

When performing a real time analysis or a live acquisition and analysis, the evidence findings are useful for gathering time-sensitive information and can lead to further acquisitions of computer systems or data. Adelstein (2006) believes that live digital forensic analysis will become an accepted standard in time.

Beebe (2009) states that digital forensics lacks processes and standards; moreover, there is the problem of scalability and non-standard computing devices. Garfinkel (2010) reiterates that there are two major problems with current forensic tools. Firstly, today's tools were not designed to assist with the investigation process but rather to find specific evidence. Secondly, the tools were created to solve crimes against people where the evidence resides on a computer. Garfinkel (2010) further reports that any forensic tools developed must be able to support the Internet and social network information associated with a user, for example a collection of social networking accounts or a user's digital footprint.

Huber et al. (2011), report that traditional digital forensics is based on the analysis of evidence from computers and network traffic. There is thus, a need for the development of new methods of evidence extraction specifically from social networks. Research has shown that there is currently very little academic research being conducted to address the collection of evidence from social networks and that currently, data extraction of information from social networks is being conducted via web crawlers (Huber et al., 2011).

A traditional forensic investigation is dependent on the seizure of hardware to perform an analysis, but with the increase in social networking sites and online communication, traditional forensic methods have become useless (Huber et al., 2011).

Mulazzani, Huber, & Weippl (2012) concur that the extraction of forensic data from social networks has become a research problem since all user communications from social networks are stored online at the social network provider without direct access by investigators. The explosion in the number of social networking sites will undoubtedly change the manner in which digital forensic investigations are conducted in the future (Mulazzani et al., 2012).

2.2 Social Networking

2.2.1 Definition

Social media can be defined as a medium that provides a platform for users to create individual profiles within a site containing personal photos, a screen name, an email address and contact details amongst others. The profile provides the user with a unique identification by which other members of the social platform can identify them (Brunty & Helenek, 2013).

Boyd & Ellison (2007) define social networking sites as Internet-based services, which create a platform for individuals to create public or semi-public profiles. A social network site allows individuals to choose who they wish to share a connection with and to view and navigate through their lists of connections, friends and friends of friends within the same social network site.

2.2.2 Social Networking Websites

Starin, Baden, Bender, Spring, & Bhattacharjee (2009) report that social networking sites are popular with millions of Internet users as these sites allow users to share information with their social network friends.

Social networking sites have been described as "fingerprints" of the 21st century by Marsico (2010). Van Manen (2010) likens social networking sites to the Greek god Momus, where an individual's innermost thoughts are published for all to see, as by choice people are giving others access to information that was once viewed as personal, private, and hidden.

Duncan (2008) states that social networking websites have become very popular over the last few years and that social networking websites allow users to create profiles, view other user's profiles, post blogs and participate in bulletins. Additionally, the FBI estimates that there are in excess of 200 different social networking sites. Social networking sites create a platform for people to meet, but the same social networking sites also allow their users a certain sense of anonymity. Duncan (2008) asserts that another risk of social networking websites is the disclosure of personal information regardless of whether such disclosure is intentional or ends up occurring unwittingly.

The explosion of social media has ensured several new communication methods for individuals and organisations but has also led to new methods by which organisations can search or investigate individuals. Law enforcement agencies have already started using social networking sites for investigative purposes (Edgar-Nevill & Qi, 2011).

Today most conversations between the youth occur on social networking sites. Whereas in the 20th century most young people would meet in shopping malls, today interactions with friends take place in the privacy of a bedroom or in a packed movie house because of the various social networking platforms available. Social networking sites have created a convenient communication medium for millions across the globe; however, as a result of the information made available on these social networking sites, law enforcement is also able to benefit from this (Marsico, 2010).

Xu, Ryan, Prybutok, & Wen (2012) report, that individuals make use of various social networking platforms for immediate access to friends, to organise social activities and to spread news in an efficient manner. Social networks also play an important role as a vehicle for political communications and influence (Xu et al., 2012). An example of the use of social media during political turmoil is the Arab Spring revolutions that took place in Egypt and Tunisia in January 2011. During this time, social media were used to gather crowds and organise protests (Lang & De Sterck, 2014).

Social networking sites are increasingly being used to seek information for employment, reconnaissance and investigations. Social networking sites are also being used in legal and criminal investigations to find and provide evidence of individuals not abiding by the law or incriminating themselves. The evidence obtained from social networking sites includes confirming location information, proving or disproving alibis, and establishing motives and personal relationships (Edgar-Nevill & Qi, 2011).

Edgar-Nevill & Qi (2011) affirm that due to the explosion of social networking sites, a wealth of valuable information is generated; this information subsequently needs to be investigated as to how it can be used as evidence.

2.2.3 Information Disclosed on Social Networking Sites

When a profile is created on a social networking site, a series of questions are required to be completed including information such as location, age, interests and an “about me” section. Users are also encouraged to include a profile photo. The user controls the visibility and privacy of their social network profile; however, in some social networking sites, such as Facebook, by being part of the same ‘network’ users are able to view each other’s profiles unless the user has increased the default security and privacy of the profile (Boyd & Ellison, 2007).

Users share an enormous amount of information on the various social networking sites; this information includes personal information, friends, colleagues and acquaintances. The user profiles can include personally identifiable information such as photos, contact information and addresses. Users within the various social networking sites also post information such as date of birth and school alumni (Ellickson & Lynch, 2010).

Social networking sites can be extremely revealing as users provide personal information and the social networking platform allows comments to be made by other participants and friends. This information provides the viewer with insight about the user’s values, activities, locations and interests (Edgar-Nevill & Qi, 2011).

Choo, Smith, & McCusker (2007) concur that social networking sites allow users to post their personal information, upload photographs and interact with other users in real time. The information available on the social networking sites can easily be used to identify or profile users, while criminals can use the information available from the social networking sites to commit identity fraud. Choo et al. (2007) also report that terrorists are able to make use of social networking sites as a means to reach an international audience, lobby for funding, recruit new members and distribute Internet radicalisation.

2.2.4 Types of Social Media Sites

Table 2 provides examples of some of the main social networking sites, the primary use of the particular social networking site, the privacy settings that are made available in the social networking site and the type of information that can be accessed via the social networking site. This information was obtained by the researcher

Table 2: Social Networking Sites (summarised from information available on the respective sites and corresponding webpages)

Social Networking Site	**Primary Purpose**	**Privacy**	**Information Available**
Facebook	To provide a platform that enables people to share and connect with others	Granular privacy module allows the user to determine what they are willing to share with their friend list, groups, and other users of the social platform Different information can be shared with different groups or users	True names of the users are encouraged when individuals create a profile but cannot be enforced Messaging includes; mail, real time chat and a timeline or "wall" Over 1 billion photos uploaded on a monthly basis Used to conduct private background checks Data is organised by user ID or group ID With a subpoena or court order the following information is available: photographs, contact information, credit card details,

			IP logs, inbox messages and chat logs Cooperative with emergency requests
Twitter	Micro-blogging	Most content is public Direct messages are private and kept until deleted by the user Simplified privacy model, updates are either private or public Short URLs can be used to serve malicious links or code	No contact number available Only last IP retained Twitter does not preserve data without legal process Most multimedia is handled by 3rd partly links
MySpace	Myspace promotes a creative community of likeminded people to connect collaborate and inspire one another	True names are not required Privacy is less granular	Messaging through chat, friend updates Profiles contain publically viewable information Data is organised by friend ID Subpoena is required for private messages, data older than 181 days, and friend lists User info and stored files are indefinitely retained IP logs and information for deleted accounts retained for 1 year
LinkedIn	Business focused providing a platform for professionals to connect	Granular privacy module, allows the user to determine what they are willing to share with their connections, groups, and public profile which is available via any web search	Profiles focused on education and work experience Used to perform background checks Profile information is not checked to ensure that it is reliable

2.2.5 Law Enforcement and Social Networking

Social media is changing the manner in which law enforcement agencies conduct criminal investigations as well as how they identify crimes. According to the 2012 online research project conducted by LexisNexis (LexisNexis, 2012), four out of five law enforcement officers are using social media for investigative purposes. The LexisNexis 2012 report sought to understand how law enforcement was using social media for the purpose of investigations, the acceptability of social media and the investigative techniques currently used by law enforcement, and the processes required to leverage social media in investigations. A total of 1221 law enforcement personnel participated in this online research study.

The LexisNexis 2012 report also provided further evidence of how law enforcement officers used social media when conducting criminal investigations. One law enforcement officer stated that “social media is a valuable tool because you are able to see the activities of a target in his comfortable stage” (LexisNexis, 2012). Sometimes suspects post incriminating evidence in the form of bragging, statements and pictures.

Zainudin et al. (2011) report that social networking websites are ideal for exploitation by criminals owing to the opportunities available to commit crimes arising from the following key properties:

- large user base that is widely distributed
- users are grouped together with common and shared interests
- social networking has created a platform that can be used to deploy fraudulent resources and applications, enticing users to install them

A law enforcement officer who participated in the 2012 LexisNexis report provided an example of how social media were useful in managing a potentially volatile situation. Law enforcement officers used social media to monitor anticipated potential civil unrest “in relation to a contract negotiation impasse/strike and Occupy group activities”. Monitoring the various social media allowed law enforcement to identify the organisers quickly, make personal contact with the organisers, and provide assistance in organising the event. This resulted in avoiding problems that were experienced in many other jurisdictions (LexisNexis, 2012).

Another law enforcement officer stated in the LexisNexis 2012 report, “I use it passively most of the time so that probationers are not aware that we use it” (LexisNexis, 2012).

Social media is very useful to see who probationers are talking to and what they are talking about. A probationer's friends are good contacts as sometimes the probationer is cautious and careful not to post about his/her activities, but the friends do not always take the same precautions (LexisNexis, 2012).

Marsico (2010) concurs that law enforcement agencies are turning to social networking sites to investigate criminals and the various crimes committed. Law enforcement has also been using social networking sites to gather intelligence relating to criminal gangs. Gang members reportedly tweeted a warning message on Twitter stating that they had a snitch in their company shortly after a gang member was arrested but released from police custody without any charges being laid. Law enforcement was actively monitoring the social networking site Twitter, and soon after the gang's initial tweet, others joined the conversation thereby provided incriminating information (Marsico, 2010).

In the 2011 London riots, various social media platforms played a key role as British Police were able to positively identify rioters using social networking sites. One rioter was arrested after encouraging looting on Facebook, while another rioter bragged on Facebook by uploading photographs of himself with his loot (Edgar-Nevill & Qi, 2011). In another case, US law enforcement was able to positively identify students who denied knowing each other; by using Facebook they were able to prove that the two students were in fact Facebook "friends" (Ellickson & Lynch, 2010).

One of the challenges that law enforcement faces when tracking individuals on social networking sites is trying to find the alias or screen name that they are using. In addition, a great deal of the information on social networking sites has nothing to do with crimes and therefore extensive time is wasted by combing through the available information. Trying to find evidence on social networking sites can be compared to finding a needle in the haystack (Marsico, 2010).

Quayle & Taylor (2011) emphasise that 90% of terrorist activity on the Internet is via a social networking site as the social networking sites offer anonymity, by safeguarding the identities of those individuals who participate. Social networking is also a method that can be used to attract new members and followers.

Evidence from social networking sites can reveal information such as personal communications, and can be used to establish motives, a crime or criminal enterprises and the instrumentalities or fruits of the crime (Ellickson & Lynch, 2010).

A US prosecutor admitted to using the Web search engine Google to perform online searches about the victims, suspects and witnesses connected to his cases. These searches provided valuable information such as photos, status updates, and blogs. The US prosecutor confirmed using Facebook, Twitter, YouTube, and MySpace when performing intelligence gathering (Edgar-Nevill & Qi, 2011).

Gangs are reportedly using social networking sites to display photos and videos of gang members holding illegal firearms and making hand gestures, and law enforcement has had successful prosecutions from the evidence available on the social networking sites YouTube and MySpace. Gang members have become wise to the fact that there are undercover law enforcement officers on the social networking sites and have started to ask for assistance in identifying these individuals (Marsico, 2010).

2.3 OSINT

2.3.1 Definition

Steele (2006) defines OSINT as unclassified information that is intentionally discovered, then categorised, separated and communicated to a select audience in order to address a specific question. Gibson (2004) defines OSINT as the legal systematic exploitation of publically available information.

Neri, Geraci, & Pettoni (2011) describe OSINT as an intelligence gathering discipline that consists of collecting information from public or open sources and analysing this information to produce valuable intelligence.

2.3.2 Sources for OSINT

OSINT came into existence well before the digital information age and was associated with intelligence gathering from open sources of information such as newspapers and public speeches largely by the military but also by other institutions with their own agendas for gathering intelligence. As the Internet developed and made available more and more unrestricted information sources, the urgency for OSINT grew exponentially (Glassman & Kang, 2012).

OSINT can be in the form of either printed or electronic formats, such as journals, television, newspapers and more recently, also applies to information available on the Internet. Intelligence analysts have made use of OSINT to supplement classified information for many years (Best & Cumming, 2007).

The Internet has become a vital resource for any intelligence analyst owing to the abilities of Internet browsers, searching, indexing and search engines (Appel, 2011). The majority of people have embraced the Internet and social networking sites, thereby ensuring that certain records of individuals' lives are embedded in the public, semi-public and deep web of the Internet (Appel, 2011).

Appel (2011) also reports that OSINT relies on the collection of various sources of information including Internet data. This intelligence is of value for national security, market research and market competitors.

Koops, Hoepman, & Leenes (2013) report that Facebook and Twitter are mined for law enforcement purposes, while online news channels are also monitored to detect and prevent terrorist activity.

2.3.3 Benefits of OSINT

There are numerous benefits of OSINT, including the fact that gathering information from open sources is usually less expensive and risky compared with information that is collected from other intelligence sources. OSINT can provide insights into new developments such as new political movements, new technologies, political gatherings and the mass movement of people (Best & Cumming, 2007). It can also diminish the burden placed on classified intelligence collection by restricting requests for information to only that which is not available and accessible via OSINT resources (Steele, 2006).

With the volume of information available publically, the main problem facing intelligence analysts is the overload of information. Most of the time there is no value to the information available and it becomes time and resource intensive to wade through the mountains of data available (Neri et al., 2011).

Steele (2006) suggests that intelligence analysts should use their proven classified intelligence methods to exploit OSINT, as this will provide them with an all-inclusive intelligent suite of products.

2.3.4 Value of OSINT

Best and Cummings (2007) assert that the intelligence community are re-examining the value of OSINT because it is freely accessible through the Internet. Some intelligence communities state that one of the drawbacks of using OSINT is the slow uptake of the development of analytical tools enabling analysts to analyse, collect and distribute the large volume of open source information (US Dept of Homeland Security & US Secret Service, 2007). Steele states that excluding the information available via the Internet is equivalent to excluding the greatest freely available data source as the Internet enables commerce, encourages and supports human interaction and provides entertainment (Steele, 2006).

OSINT is able to combine all available resources and expertise without the need for security clearance and to produce intelligence that can be shared with everyone. This is extremely valuable for early warnings and law enforcement investigations. One of the downsides of the freely available information from the Internet is that this information needs to be assessed for its source and reliability (Steele, 2006). Gibson (2004) concurs that OSINT is criticised as it cannot easily be verified or evaluated mainly because the majority of the information is obtained via the Internet. OSINT has no boundaries and is a continuous process of collecting, processing, sharing and analysing information.

The world is changing with the advancement of technology and ecommerce. Today information that is freely available on the Internet often proves to have more value in helping intelligence analysts understand the world than the results obtained from traditional cloak and dagger intelligence (Qureshi & Memon, 2012). Intelligence professionals are in agreement that OSINT is useful and should therefore be collected and analysed in the same way that classified intelligence is gathered and analysed (Best & Cumming, 2007).

Cuijpers (2013) argues that intelligence communities require advanced tools to navigate their way through the oceans of available information. Even though the Internet has provided intelligence analysts with search engines and translation applications, more refined tools are required. Cuijpers (2013) confirms that existing tools lack the functionality of propriety tools, which have been designed with the purpose of processing data for evidence.

VIRTUOSO is a project sponsored by the European Commission to develop OSINT tools. Cuijpers (2013) recognises that these tools will be of great value to law enforcement and intelligence agencies, but that they will threaten the freedom of citizens whose personal information is at the core of the open sources.

2.3.5 OSINT as a Business Tool

OSINT is said to be an accepted practice within the private sector, and is becoming more advanced especially with the development of tools and techniques. OSINT is no longer being questioned regarding its usefulness or validity but rather how quickly it can be developed into a discipline for government and private sector intelligence analysts (Gibson, 2004).

2.3.6 OSINT Source Verification

Gibson (2004) states that OSINT has always been available but in the last few years it has received recognition and been widely used. Moreover, OSINT does not always have to be obtained openly; some information can be discreetly obtained. However, Gibson (2004) reiterates that intelligence must always be accurate, reliable, timely, and verifiable.

OSINT is in the public domain and should not be confused with “available to the public”, as there are some barriers to accessing OSINT. The barriers are usually resources and the amount of effort required to gather the OSINT. Gibson (2004) suggests that a checklist (see Table 3) be drafted to confirm the usability of all OSINT sources.

2.3.7 Covert Intelligence

Covert operations are also conducted to access information that is not publically available and to communicate with suspects and chart social relationships (Ellickson & Lynch, 2010).

An example of this is given in the LexisNexis report of 2012, which investigated the use of social media by law enforcement personnel (LexisNexis, 2012). One of the law enforcement officers stated that, while trying to locate a suspect in connection with various drug related charges, the suspect's Facebook profile was viewed and a Facebook friend request was sent to the suspect from a fictitious Facebook profile. The suspect accepted the Facebook friend request and using Facebook's location services, the suspect kept checking in, thereby allowing the law enforcement officer to follow the movements of the suspect and eventually track him down.

Table 3: OSINT Source Checklist (taken from Gibson, 2004)

Checklist	Description
Authority	Does the OSINT source command high opinion or respect from peers or customers?
Accuracy	How accurate is the OSINT source for example, can it be validated or benchmarked?
Objectivity	Is the OSINT source biased in any way?
Timely	Is the OSINT source date/time/location tagged?
Relevancy	How relevant is the OSINT source?

2.4 Summary

Chapter 2 discussed in detail what digital forensics, social networking and OSINT are. Examples of how social networking and OSINT are currently being used by law enforcement when they conduct investigations were discussed as well as the value of the information available on the Internet. It has been determined from the literature review that there is currently no investigative framework available to assist with using OSINT sources as a digital forensic investigative tool. Chapter 3 discusses the methodology used in conducting the research on which this thesis is based.

Chapter 3 Methodology

This chapter details the approach and methodology used in this study. The topics covered include reasons for the use of the interpretive philosophy and how the representative sample was obtained. This chapter also includes a discussion of the techniques and tools used to gather and analyse the data.

3.1 Research Strategy

The purpose of the research is exploratory as it seeks to establish a framework for using OSINT and social networking as a digital forensic investigative tool.

The research philosophy followed is interpretive as the research was conducted using people as subjects rather than inanimate objects. The philosophy is also interpretive because the survey questionnaires allowed for more than one answer to be given.

The research conducted is qualitative as most of the data collected is subjective, making it difficult to quantify. For the data to be useful the survey responses were analysed and interpreted using qualitative data analysis processes thus allowing a theory to be developed from the data (Saunders, Philip, & Thornhill, 2009).

The time frame for conducting the research was cross-sectional; such a time horizon is useful for a study of a particular fact to explain how factors are related, but in different entities such as organisations or disciplines such as digital forensics (Saunders et al., 2009, p. 155).

3.2 Data Collection and Analysis

The primary data was collected by means of an electronic survey questionnaire sent to digital forensic practitioners. The electronic surveys were sent by email and were completed via a web link.

The survey questionnaire was compiled using mostly rating questions which are often used to collect opinion data. Due to the collected data being qualitative, the researcher chose to use a Likert-style rating scale. Each respondent was asked how strongly he or she agreed or disagreed with a statement or series of statements when completing the questionnaire (Saunders et al., 2009). Several open-ended questions were also included to obtain more in-depth information about the response to a previous question.

The survey was divided into three sections, namely demographic data, social media and lastly OSINT gathering. The first section of the survey related to demographic data. The information provided by the responders would provide the researcher with data relating to the respondent's education, how many years of digital forensic investigation experience they hold, the industry they are currently working in and, lastly, what their role in digital forensics is.

The second section contained questions relating to social media. Respondents were asked to indicate if and how they use social media when they conduct digital forensic investigations. This section is important to the researcher as the answers provided by the respondents will allow the researcher to determine how social media is being used and the respondents' opinions of the value of social media when conducting digital forensic investigations.

The third section of the survey contained questions relating to OSINT. Respondents were asked if they made use of OSINT when conducting digital forensic investigations and whether the OSINT was for reactive or proactive digital forensic investigations. The answers to the questions posed in this section were directly related to the information needed to define the framework for using OSINT as a digital forensic tool.

The complete questionnaire is attached as Appendix A.

3.3 Survey Participants

The sample group chosen for the research comprised digital forensic practitioners only. An electronic survey questionnaire was sent to 75 respondents via email. This particular sample group was chosen as digital forensics is a highly specialised field and all of the respondents were individually contacted about participation in the research. However, despite all respondents being known to the researcher, completion of the questionnaire was done anonymously. Eighteen completed electronic questionnaires were received from the total of 75 initially sent out.

Participation in the survey was voluntary. When first contacted via email, each participant was informed of the following:

- Participation in the survey questionnaire would be confidential and the participant's identity would not be revealed in the study;
- Participants could stop the survey questionnaire and cease to participate if the process became too intrusive;
- The findings would be shared with all participants upon request.

The survey was approved by the ethics committee of Rhodes University (case number: CS 13-06).

3.4 Participants' Demographic Information

Fig. 1 displays the respondents' professional certifications. Nine of the 18 respondents currently hold the vendor specific certification, AccessData Certified Examiner (ACE); four hold the Association of Certified Fraud Examiners (CFE) certificate, while four hold certifications that were not listed on the questionnaire and are shown on the chart as other. Two of the respondents do not hold any digital forensic certifications and this is displayed on the chart as none. Two of the respondents hold the Certified Information Systems Security Professional (CISSP) and the Certified Information Systems Auditor (CISA) credentials. Two of the respondents hold a digital forensics specific qualification called the GIAC Certified Forensic Analyst (GCFA), one respondent holds the GIAC Certified Forensic Examiner (GCFE), and finally, one respondent holds the vendor-specific certification, Encase Certified Examiner (EnCE).

Table 4 presents the demographic information about the sample group. This information, supplied by the respondents, includes their location, the highest tertiary qualification achieved, the industry they work in, the number of years they have been performing digital forensic investigations, and their current digital forensics job role.

Table 4: Demographic Information of the Respondents (n=18)

Participant	Location	Highest tertiary qualification achieved	Industry	Years of digital forensic experience	Current digital forensics role	Are you comfortable with social media investigations
A	South Africa	Honours Degree	Consulting	3 - 5 years	Corporate Investigator	Comfortable
B	South Africa	National Diploma	Consulting	3 - 5 years	Corporate Investigator	Not comfortable
C	South Africa, Zambia	Honours Degree	Education	15 years or more	Corporate Investigator	Neutral
D	South Africa	Bachelor's Degree	Insurance	5 – 10 years	Corporate Investigator	Neutral
E	South Africa	Diploma	Consulting	5 – 10 years	Corporate Investigator	Neutral
F	South Africa	Master's Degree	Government	15 years or more	Law Enforcement	Extremely comfortable
G	South Africa, Kenya, Zimbabwe	Master's Degree	Financial	5 – 10 years	Corporate Investigator	Neutral
H	South Africa	Post Graduate Diploma	Consulting	10 – 15 years	Corporate Investigator	Neutral
I	South Africa	Honours Degree	Government	3 – 5 years	Law Enforcement	Neutral
J	South Africa	Honours Degree	Consulting	5 – 10 years	Corporate Investigator	Neutral
K	South Africa	Bachelor's Degree	Consulting	5 – 10 years	Corporate Investigator	Neutral
L	Netherlands	Professional Certifications	Consulting	10 – 15 years	Incident Responder	Comfortable
M	South Africa	Honours Degree	Information Technology	5 – 10 years	Criminal/Civil Defence	Neutral
N	South Africa	Honours Degree	Insurance	3 - 5 years	Corporate Investigator	Neutral
O	South Africa	Honours Degree	Education	3 - 5 years	Corporate Investigator	Extremely comfortable
P	Australia	Master's Degree	Consulting	10 – 15 years	Corporate Investigator	Neutral
Q	South Africa	Professional Certifications	Consulting	10 – 15 years	Corporate Investigator	Neutral
R	South Africa	Professional Certifications	Financial	2 – 3 years	Incident Responder	Extremely comfortable

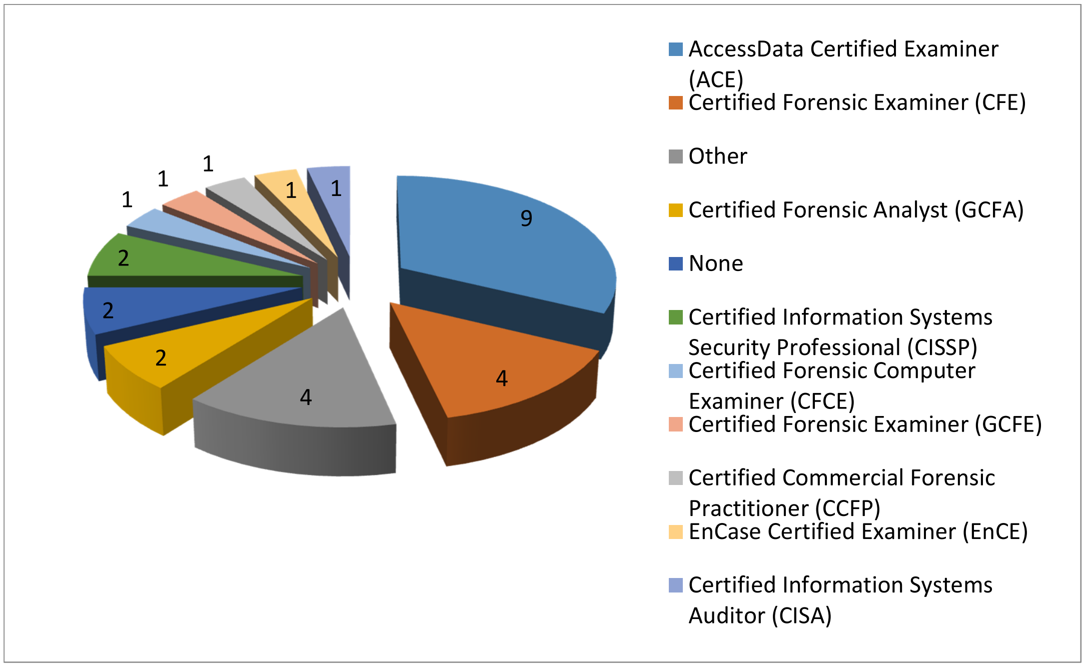


Fig. 1. Professional certifications held by the respondents.

3.5 Summary

Chapter 3 discussed the research methodology used in this study, and provided demographical information relating to the survey respondents.

Chapter 4 discusses the analysis of the data which was collected through the survey.

Chapter 4 Analysis of Data

This chapter presents the results of analysing the data collected through the survey and relates these to the research question and objectives.

Since analysis of the data using mean and standard deviation formulas was not meaningful owing to the small sample size of 18 respondents, the data from the electronic questionnaire was analysed using descriptive analysis.

4.1 Analysis of Survey Questions

4.1.1 Use of Social Media to Obtain Specific Information

The first question asked the respondents to rate the strength of their belief that social networking sites could assist digital forensic investigators in ascertaining certain kinds of information using the scale 5 (strongly agree) to 1 (strongly disagree). Fig. 2 illustrates the respondents' answers. In this and the subsequent figures in this chapter, agree and strongly agree, and disagree and strongly disagree responses are combined as agree and disagree, respectively.

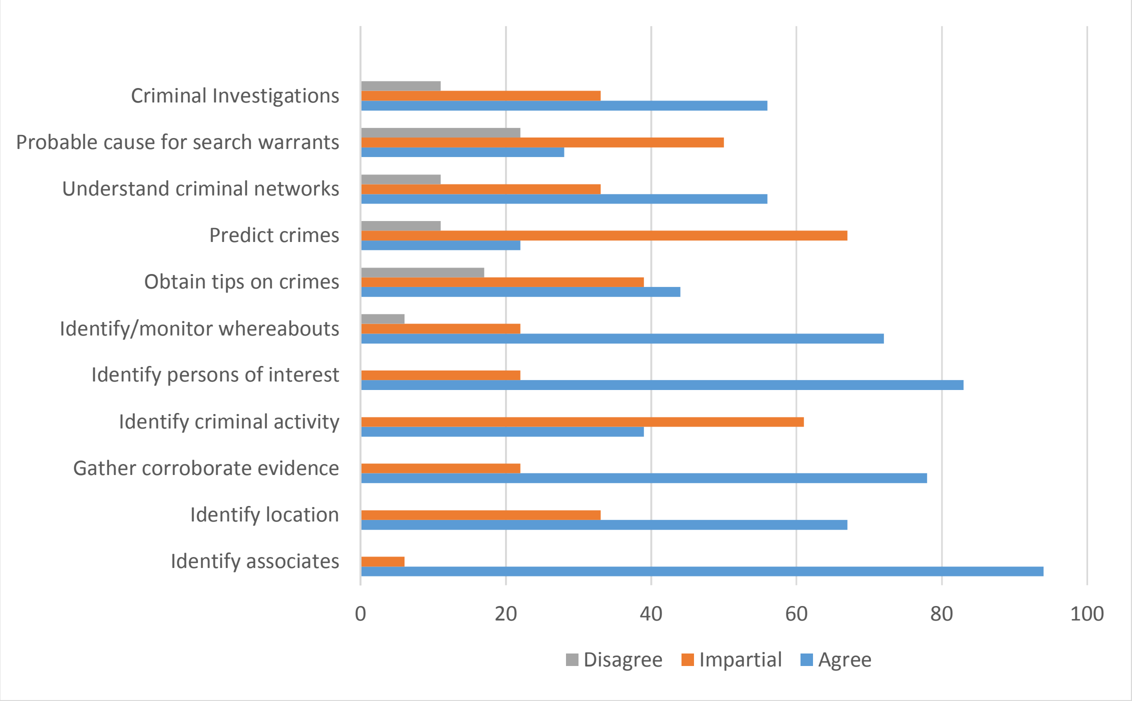


Fig. 2. Responses (n=18) to whether social media sites can assist digital forensic investigators to ascertain certain kinds of information.

Overall, the majority of respondents agreed that social networking sites could assist digital forensic investigators to ascertain all but two of the kinds of information presented.

- 67% of the respondents were ambivalent as to whether social networking sites could assist in the prediction of crimes.
- 61% of the respondents were ambivalent as to whether social networking sites could assist in the identification of criminal activities.

The results from the LexisNexis 2012 Report provide evidence that the top three uses of social media for crime investigations are: "Identifying people and locations, discovering criminal activity and lastly gathering evidence." (LexisNexis, 2012). A law enforcement officer who participated in the 2012 LexisNexis report stated that social networking sites had assisted them in crime prevention. The law enforcement officer believed that they had thwarted a "Columbine type shooting" by performing an investigation utilizing Facebook. There was sufficient evidence revealing that the threats were credible and additional investigations were conducted that provided evidence that a student was in the process of acquiring weapons and had the intent to harm others (LexisNexis, 2012).

4.1.2 Value of Information Obtained from Social Media

The next question attempted to ascertain whether information obtained from social media was valuable to digital forensics practitioners. A total of 83% of the respondents indicated that they agreed that there is value in obtaining information from social networking sites, while the remainder (17%) were undecided.

In support of these results, another participant in the LexisNexis 2012 report provided the following statement; "It is amazing that people still "brag" about their actions on social media sites. Yeah, even their criminal actions" (LexisNexis, 2012). Despite denying any involvement in an assault involving a victim being struck with brass knuckles, the suspect bragged about hurting a kid on his Facebook profile and stated that the item used to strike the kid was dumped in a trashcan in a park. The law enforcement officers searched a number of parks and located the brass knuckles. Together with the evidence item and the Facebook post, the suspect confessed to the assault during a follow up interview (LexisNexis, 2012).

4.1.3 Tools Used to Mine Social Networking Sites

The next question attempted to ascertain which tools digital forensics practitioners made use of when exploiting social media for digital forensic investigations.

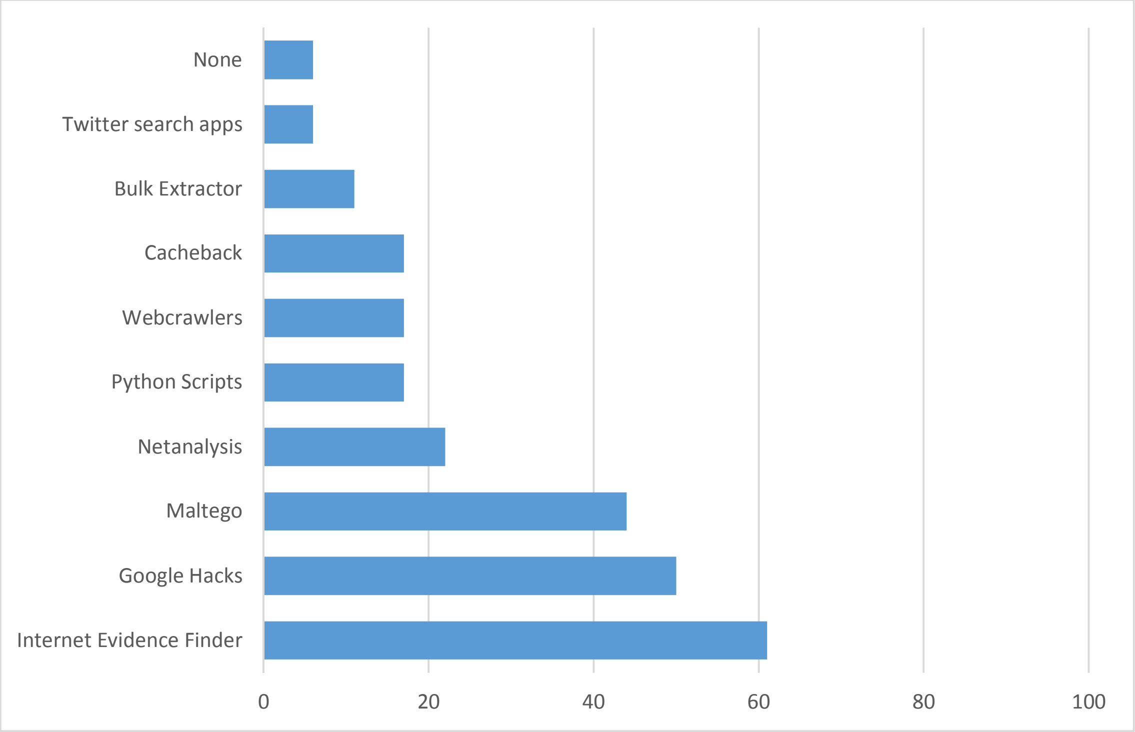


Fig. 3. Responses (n=18) to which tools are used when using social media for digital forensic investigations.

Internet Evidence Finder (IEF) is propriety software and is capable of recovering hundreds of Internet-related artefacts according to the information made available on the website[2]. It is used by 61% of the survey respondents. IEF is a useful tool as it is able to recover evidence relating to social media, webmail, cloud services, web activity and many others.

[2] http://www.magnetforensics.com/mfsoftware/internet-evidence-finder/ief-artifact-modules/

Google hacking involves using the Google search engine[3] together with specific search operators that have a specific meaning to the Google search engine (Long, 2008). When a search operator such as "filetype:pdf" is used it will return Google search results containing sites with the document extension pdf. This search is very useful when searching for a specific pdf document on the Internet. Bradbury (2011) concurs by stating that search engines are a useful resource and that Google hacking is a useful tool. Another example of a useful Google hack is "cache:website" as this returns a cached version of a website, which is useful when looking for an older version of a website. Half of the respondents make use of Google Hacks.

The OSINT and forensic tool Maltego[4], is used by 44% of the respondents. Maltego is used for analysing and visualising connections in data. For example, taking information from the social networking tool LinkedIn and running this information through Maltego returns results such as memberships of other social networking sites and contact information if the LinkedIn profile owner has added this information (Bradbury, 2011). Another useful feature of Maltego is the ability to search the web for valuable information such as email addresses (O'Connor, 2010). When using the Maltego email search feature and entering a name and surname, Maltego searches the web and returns results for all email addresses that contain the name and surname specified.

The tool Netanalysis[5], which is able to extract and analyse data from various Internet browsers is used by 22% of the respondents. The software is useful as it is able to import Internet history and cache data from the various Internet browsers.

Python[6] is referred to as a hacker's programming language as it is not as complex as some of the other programming languages. It also has limitless third–party libaries and is an excellent development platform for creating customized offensive tools (TJ O'Connor, 2013). When using social networking for digital forensic investigations, 17% of the respondents use Python scripts. An example of using Python when performing a

[3] https://www.google.co.za

[4] https://www.paterva.com/web6/products/maltego

[5] http://www.digital-detective.net/digital-forensic-software/netanalysis

[6] https://www.python.org/

digital forensic investigation using social media would be to write, or modify, a Python script that can scrape data from Twitter[7] using the Twitter application programming interface (API). A Python script can be written to extract the tweets and retweets of any Twitter user and extract geolocation data (TJ O'Connor, 2013). An example of a Python script to scape Twitter for tweets amd retweets in a certain location is given in Appendix B.

Web crawlers[8] are also known as web spiders or web scrapers. Clough (2010) explains that a web crawler is a computer program that executes across the Internet performing functions such as searching, retrieving and copying information from various websites. Web crawlers are commonly used by spammers to harvet email addresses from websites. Only 17% of the respondents reported that they use web crawlers when performing a digital forensic investigation.

Internet Examiner Toolkit[9] (IXTK), formally known as Cacheback, is described as a multilingual forensic tool that is able to discover, analyse and report Internet evidence from various Internet browsers. IXTK is able to reassemble the history and caches of various Internet browsers and has the capability of allowing the investigator to view the recreated Internet history and compare it to the live website available on the Internet. Only 17% of the respondents reported that they use IXTK.

Bulk_extractor[10] is a digital forensic tool which can extract information such as email addresses, usernames and URLs. Bulk_extractor can analyse data from images created of hard drives in addition to also being able to directly analyse media connected to a forensic investigator's computer. The power of the application bulk_extractor is that it is able to search multiple drives or images due to bulk extractor utilising multiple scanners (Garfinkel, 2013). Only 11% of the respondents reported that they use bulk_extractor when performing a digital forensic investigation.

[7] https://about.twitter.com/

[8] http://www.webcrawler.com/

[9] http://www.siquest.com/news/products/ixtk/ixtk_v4_brochure.pdf

[10] http://digitalcorpora.org/downloads/bulk_extractor/

Twitter is an open platform that allows its users to build applications around the information through Twitter, provided that the developer of the Twitter search application adheres to Twitter's API terms and conditions. Maltego has the capability of performing searches using the Twitter API. This search feature in Maltego allows an individual to perform searches across Twitter, providing the Maltego user with information such as the location of the tweet and the tweets and retweets of the individual the search was performed on. Any photos uploaded via Twitter by the individual who is being searched for, are also made available to the Maltego user via the Twitter API search function. Python scripts can also be written using the available Twitter API to extract tweets and locations. Only 6% of the respondents reported that they use Twitter search applications when performing investigations using social media.

4.1.4 Social Networking Sites Used Most Frequently

The next question attempted to determine which social media sites are used by investigators when searching for supporting information during an investigation.

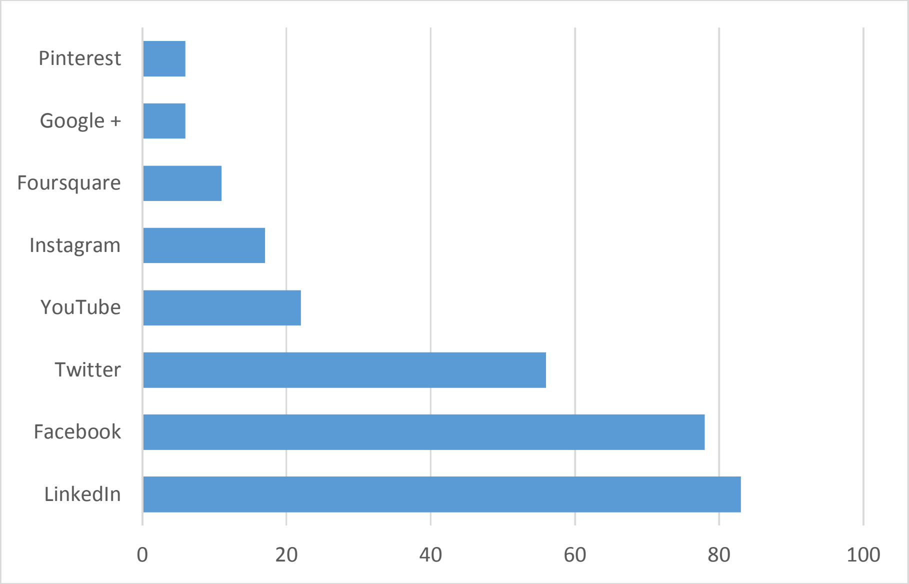


Fig. 4. Responses (n=18) to which social media websites are most frequently used when searching for information as part of an investigation.

The social networking site LinkedIn[11] is used by 83% of the respondents when searching for information as part of an investigation. A LinkedIn profile contains a wealth of information on individuals; apart from the more common information such as email address, geolocation tagging and marital status. Since LinkedIn is predominantly a professional networking tool, it provides information such as an individual's job title, where they last worked and the educational institution they attended. It also indicates people they might know (Bradbury, 2011).

According to Facebook's key facts,[12] there were approximately 1.32 billion Facebook users as at 27th July 2014 with an average of 4.75 billion items being shared daily. A total of 78% of the respondents reported that they use Facebook when conducting investigations. Law enforcement officers who participated in the LexisNexis 2012 report, reported that personalised social media sites such as Facebook and YouTube are used most frequently when conducting investigations (LexisNexis, 2012).

The social networking tool Twitter is used by 56% of the respondents when conducting investigations. Twitter can provide information such as who is following the particular individual associated with the Twitter handle and who the Twitter account is following. This information is valuable especially if an individual is being monitored for communication with other individuals or groups. In 2011, Twitter was used to mobilise a group of Twitter users to request support for the post cleanup of the streets after the 2011 London Riot[13] with the hash tag #riotcleanup.

YouTube[14] is a video sharing platform, which 22% of the respondents reported that they used when performing investigations. One of the law enforcement officers who participated in the LexisNexis 2012 survey reported that a YouTube video assisted them in the successful arrest of gang members. Known gang members had created a video for recruiting new gang members and promoting gang violence (LexisNexis, 2012).

[11] https://www.linkedin.com

[12] http://newsroom.fb.com/company-info/

[13] http://www.theguardian.com/uk/2011/dec/07/twitter-riots-how-news-spread

[14] www.youtube.com

Instagram[15] is a social networking application, which is installed on smartphones via a mobile application allowing users to share movies and photos on Instagram. Users are also able to share photos and movies using Instagram via Facebook and Twitter, 17% of the respondents use Instagram when performing investigations.

Foursquare[16] a social networking application that allows users to perform check-in and real time location sharing with friends, which 11% of the respondents use.

Google Plus[17] is a social network from the search engine Google. Google Plus creates a sharing and storage platform for its users, of which 6% of the respondents use.

Pinterest[18] is a social networking application used by 6% of the respondents. Pinterest enables individuals to create and share visual bookmarks known as Pinboards. A Pinterest account can be registered using an existing Twitter or Facebook account, which enables a user to notify their Facebook and Twitter friends when a new pin is added to their Pinboard.

4.1.5 Performing Social Networking Investigations

The next question attempted to ascertain how comfortable the respondents are performing social media investigations. The majority of the survey respondents (67%) indicated that they were undecided about how comfortable they are with performing social media investigations, 17% of the respondents indicated that they are extremely comfortable, while 11% indicated they are not at all comfortable with social media investigations.

According to the LexisNexis 2012 report, the primary reason why social media is not being used to assist with investigations is the lack of social media skills (LexisNexis, 2012). A third of the 1221 law enforcement officers are uncomfortable using social media; this could be because of a lack of training. Those law enforcement officers who

[15] http://instagram.com/

[16] https://foursquare.com/

[17] https://plus.google.com/

[18] https://about.pinterest.com/en

are comfortable with social media investigations are self-taught. It is possible that the respondents who indicated that they were undecided about how comfortable they are with social media investigations, were undecided owing to their lack of social media skills or lack of training, or are self-taught without having received any formal training, and therefore, may not value their skills as adequate.

4.1.6 Verification of Information Available from Social Media

The following question endeavoured to determine whether the information available on social media platforms should be considered evidentiary or supplementary evidence. Respondents were also asked to explain their answer and offer suggestions relating to methods that could be used to verify the information obtained from social networking sites.

The majority of the respondents (89%) reported that information from social media sites is supplementary, while the minority (11%) of the respondents reported that the information has evidentiary value.

Listed in Table 5 and grouped into common themes are some of the respondents' comments as to why the information obtained from social media websites is considered supplementary.

Table 5: Common Themes for Respondents' Comments on the Supplementary Nature of Information Obtained from Social Media Sites

Theme	Respondents comment
Difficulty in verifying the truth thereof:	• "there are a lot of fake profiles and misleading information on social media sites" • "there is no guarantee that the profile viewed actually is or belongs to a person of interest" • "the information might not be true therefore you will not use it as only evidence you need to get substantial evidence to confirm whatever you see on net"

	• "if the source and validity of the information cannot be corroborated or validated then it pretty much amounts to hearsay" • "it is dependent on individuals voluntarily submitting information in a public domain" • "it's not always easy to verify that the info presented is true" • "important to clarify the source of information so others can evaluate its value"
Providing direction for an ongoing investigation:	• "could be highly important to provide evidence regarding modus operandi" • "it would support the investigation or to direct an investigation" • "it may be used to assist or steer your investigation in the right direction" • "there are varying degrees to which social evidence can be used as primary evidence depending on the nature of the case at hand"
Cyber profiling and intelligence gathering:	• "evidence when building a cyber-profile, or assist with providing or finding other leads" • "it could give you an indicator of the individuals personal life like last checked in places, photos etc…" • "it is generally used for intelligence purposes" • "leaning towards intelligence rather than evidence" • "investigators can also employ social engineering to befriend targets or infiltrate their "groups'" in order to gain the trust of targets"

Below are some of the respondents' comments regarding why they considered the information obtained from social media websites as evidentiary. All of these comments relate to how the evidence is obtained from social media platforms.

One respondent qualified his/her response by including the caveat “If the social media content was accessed lawfully”. Most social media platforms have policies and procedures in place for providing evidence for law enforcement. A subpoena is usually a requirement by the social media platform before the information is disclosed.

Another respondent reported that “the authenticity of the evidence can be assured at a later date”. This comment is interpreted to mean that once evidence is located via other means such as screen scraping using Python scripts, a subpoena can be obtained for the evidence from the social media platform.

Another respondent provided the following comment: “Most social media sites have a law enforcement contact that will supply IP addresses, registration details, etc…” This comment reiterates what was stated above; that is, that most social media platforms have processes in place to be able to provide the evidence required by law enforcement when social media is used to commit a crime, but a subpoena is required before the evidence is released.

A further comment from one of the respondents offers the following insight: “If analysis is being done of a forensic image then it is possible to use the data as evidence”. This means that if a forensically sound image is obtained from the social media platform or the suspect’s PC or mobile phone, the evidence obtained from this image would be evidentiary.

The findings of the LexisNexis 2012 report stated that 60% of the respondents indicated that the information from social media websites is not credible, but that 87% of the time social media evidence holds up in court when used as a probable cause to obtain a search warrant (LexisNexis, 2012).

4.1.7 Using OSINT When Performing Digital Forensic Investigations

The next question attempted to determine whether any of the respondents use OSINT techniques when performing investigations. Respondents were asked to indicate if they had used OSINT techniques and, if they had, whether it was for a reactive or proactive investigation. A reactive investigation is an investigation that takes place as a result of a crime having taken place. A proactive investigation is one performed to prevent a crime from taking place.

By its very nature, digital forensic investigations are reactive as evidence is required, and tends only be collected, after a crime or incident has taken place. An example of a reactive investigation is when a crime has taken place and evidence is required to ascertain if a suspect was at a certain location. In this case, possible OSINT evidence sources are tweets with geolocation information included, or finding associates of the known suspect who may have uploaded photos that include the suspect to social media platforms.

A proactive investigation is an investigation that takes place in order to prevent a crime or incident from taking place. An example of a proactive investigation would be monitoring the tweets of suspects to prevent a crime or incident from taking place such as monitoring the whereabouts of known suspects or gang members.

More than half of the respondents (56%) reported that they have used OSINT techniques for reactive investigations, while 33% of the respondents reported that they have never used OSINT techniques.

The minority of the respondents (6%) reported that they have used OSINT techniques for proactive investigations, while 6% also reported that they used OSINT techniques for both proactive and reactive investigations.

It has been reported (Brunty & Helenek, 2013) that various law enforcement groups are already using social media sites as a tool to assist in the monitoring of known individuals, gangs and criminals. Brunty and Helenek also predicted that as more law enforcement officers realise the value of the information stored in social media sites, more law enforcement officers will turn to the Internet for assistance in investigations.

4.1.8 Intelligence Gathering

The respondents of the survey were asked to state which of the given four statements relating to the information available through OSINT were the most applicable in their opinion.

More than half of the respondents (56%) indicated that they agreed with the statement: “OSINT should be recognised as a valuable source information in its own right and not only as supplement and complement to covertly obtained information“.

A smaller number of the respondents (39%) reported that that they agreed with the statement: “OSINT provides insights and assists with the focus for obtaining covert information”.

Fewer respondents (22%) indicated that they agreed with the following statement: “More information and value is gained from covertly collected intelligence than from OSINT”.

The minority of the respondents (11%) indicated that they agreed with the statement: “The cost and processes that are used when obtaining covert information ensure that the information can be trusted”.

4.2 Discussion of Results

The first objective of the thesis was to ascertain whether social networking and OSINT could indeed assist a digital forensic investigator during their investigations. The results from the survey confirm that most of the respondents agree that social networking sites and OSINT can assist digital forensic investigators in this regard. Secondly, the majority (89%) of the respondents agreed that there is value in the information available via social networking, but only a little more than half (56%) of the respondents agreed that the information available from OSINT sources is valuable. A possible reason for the discrepancy between the two results is that the respondents are not familiar with OSINT, do not comprehend the inherent value of OSINT and do not understand that social networking is in fact a form of OSINT. Another possible reason is that respondents are more comfortable with traditional digital forensic techniques, which involve imaging a physical drive rather than acquiring evidence from an Internet source. If the respondents are not active users of social networking sites, this too can cause resistance and discomfort as it is possible that they are not familiar with how to use social networking platforms and their features and settings.

The third objective of the thesis was to determine the criteria for using social networking sites and identify the types of evidence that could be obtained from these sites. This objective has been achieved by the respondents confirming which social networking sites they use most frequently when conducting investigations. The majority of the respondents indicated that LinkedIn is the most frequently used. A possible reason for this choice is that one does not have to be a valid LinkedIn user to view LinkedIn's publically available user profiles. Although LinkedIn user profiles are publically visible, the information displayed depends on the privacy settings that the profile owner has configured for his/her profile. LinkedIn is marketed as a professional networking social networking tool, and therefore it is possible that profile owners do not secure their profiles as they use LinkedIn for new job opportunities. The types of information available from a LinkedIn profile are a profile photo, employment history, interests based on group memberships, education status, and any personal websites the profile owner has added to the profile.

Some of the respondents stated that they are currently using social networking sites to provide supplementary evidence for their investigations. Additionally, the survey results show that a minority of the respondents are comfortable performing social networking investigations. This low uptake amongst the respondents could be attributed to the lack of formal training and guidelines available for such investigations. There may also be trepidation as a result of social networking information being regarded as supplementary rather than evidentiary as, by its very nature, digital forensics is based upon providing irrefutable facts relating to data being analysed, especially when the data is meant to be presented in a court of law. However, if the information is obtained from the social network following the law enforcement process of obtaining a subpoena, and the evidence is acquired forensically, the evidence can then be viewed as evidentiary as indicated by the respondents.

The majority (89%) of the respondents indicated that it would be useful for an OSINT tool-set to be developed and made available. Additionally 83% of the respondents indicated that it would be useful for an OSINT framework to be developed for application when performing investigations. The proliferation of social networking platforms could be a possible reason why respondents report that an OSINT tool-set and framework would be useful when performing investigations. Additionally, the respondents are likely aware of the wealth of information available from the various OSINT sources but are seeking guidelines or a framework to assist them when performing OSINT investigations. The stated need for a framework or tool-set could also be a result of investigators preferring to opt for the use of tools and techniques that have been tried and tested.

4.3 Summary

It has been observed from the analysis of the survey results presented in this chapter that social networking and OSINT sources contain a wealth of information that can assist investigators when conducting investigations. The second objective of the thesis was to create a framework which could be used when conducting investigations involving obtaining information from social networking sites. Three of the thesis objectives have been met as described in this chapter. The proposed framework is discussed in the next chapter.

Chapter 5 Proposed Framework

Based on the survey results and the most important findings presented in the previous chapter, a framework that can be applied when using OSINT and social networking to assist with a digital forensic investigation is presented in this chapter.

5.1 Design of Framework

In the survey respondents were provided with seven steps. Respondents were asked to place the processes in their preferred order from one to seven to indicate in which order the processes should be carried out. Based on these results, a framework has been developed comprising six steps to be followed when performing an investigation using social networking and OSINT sources. It is important to note that the final framework consists of six rather than seven processes due to the researcher combining two of the processes that she regarded as complementing each other.

The initial seven processes provided in the survey consisted of five of the most common steps used when performing a digital forensic investigation as recommended by Casey (2011). Two extra steps were also included, namely visualise and collaborate, because it is useful to be able to visualise social media connections or chart a timeline of events when performing an investigation that includes OSINT sources. This can also be of value when collaborating with other investigators.

Using the respondents' mean ratings for this question and ordering the processes accordingly, a framework has been developed. The reason for adding the question regarding the ordering of seven processes to the survey was because the researcher wanted an independent consensus for the processes that would form the framework.

The proposed framework consists of the following six key steps: Identify, Retrieve and Collect, Analyse and Process, Visualise, Collaborate and Report, in the order shown in Fig. 5.

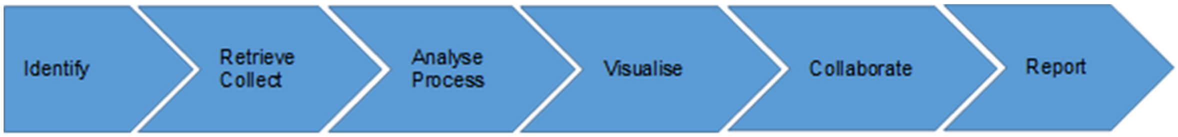


Fig. 5. Key steps in the proposed digital forensic investigation framework.

5.2 Step 1 Identify

The first principle of the framework is Identify, as the key person or persons must be identified before the information gathering can commence. Thus, the aim of this step is to identify possible sources of evidence from various OSINT sources including social networking platforms.

Information obtained may include possible associates, place and date of birth, any photographs affiliated with the key person or persons, possible contact numbers, locations visited, marital status, current employer, previous employers, and a list of schools. There is also value being able to identify the key person or persons' interests such as sports, hobbies or social groups since this information assists with creating a profile of the person or persons of interest. As members of social groups tend to share information freely, photographs of social events which the participants have attended, are particularly useful.

It is also important to perform an assessment of the possible evidence sources with respect to the mandate provided to the digital forensic investigator as this will assist in determining the scope of the digital forensic investigation and which OSINT sources could possibly contain evidence or information. This information is vital for the retrieval and collection principle when the information and evidence must be retrieved, collected and in some instances exploited.

Mckemmish (1999) states that digital forensics consists of four key principles, namely, identification, preservation, analysis and presentation. These four principals are vital in ensuring that the evidence is legally acceptable. Moreover, McKemmish (1999) states that in order to be able to determine which processes are required to retrieve evidence, the digital forensic investigator needs to understand what type of evidence could possibly exist, where the evidence could be stored and how the evidence is stored as this will assist the digital forensic investigator to identify the correct methods to use when extracting the evidence.

Kim & Glassman (2013) suggest that performing a search is the starting point for other fruitful uses of the Internet. This is a new form of OSINT which has been dubbed the Search, Organise and Differentiate (SOD) scheme. One of the Internet's most valuable features is the vast amount of freely available information. This feature is possibly also one of its biggest downfalls as there is a reliance on the ability to be able to organise discovered information and, more importantly, the ability to be able to differentiate between the relevant and irrelevant information. Individuals who use the Internet to improve their knowledge must be able to cope with the large volume of information and also be able to filter worthless and deceptive information efficiently.

Below are the key sub steps that should be completed to fulfil the aims of the Identify process.

- Identify the individual or group of individuals.
- Identify the type of evidence or information that is required for the investigation, such as photographs, confirmation that an individual uses a specific online alias, and so on.
- Identify social networking sites or OSINT sources that can be used when performing searches and which could contain evidence. Examples of these sites and sources include Facebook, LinkedIn and Web Search Engines.

5.3 Step 2 Retrieve and Collect

The second principle is the retrieve and collect process that details how the digital evidence must be extracted and collated. The process of extraction refers to information gathering and the recovery of digital evidence.

Social media is a dynamic platform because, as quickly as information is posted online, it can also be removed. A screencast tool, also known as a video capturing tool, can assist digital forensic investigators as it captures and records exactly what they are viewing on their screen in real time. Some screencast tools also allow for use of a webcam and audio so that a narrative can be provided by the digital forensic investigator who will be able to provide comments while working through the evidence.

It is imperative that the digital forensic investigator remains ethical at all times and does not gather any evidence in a legally questionable manner as this will render the evidence inadmissible in a court of law. The digital forensic investigator must also remain mindful of not violating the terms and conditions of the various social networking sites.

Extracting evidence from social networks is most often a difficult task and legal advice and assistance may be required to obtain the evidence. Various laws governing privacy, access to information, electronic communication and data transmission, to name a few, also need to be thoroughly understood and taken into consideration as the evidence may be located in a foreign country.

The McKemmish principle of preservation should be observed during this phase owing to the likelihood of legal scrutiny. It is imperative that any examination of electronic data is conducted in the least intrusive manner. This means that the least amount of change must be made to the electronically stored data and that any unavoidable changes can be accounted for and justified. In a situation where evidential data is changed, the digital forensic investigator must document and be able to explain what changes have taken place and why (Mckemmish, 1999).

The collection of OSINT can be accomplished in a variety of ways: by performing operator searches using web search engines or by performing searches on various websites that contain their own databases such as government departments and domain registration websites.

Search engines such as the Russian search engine Yandex[19] allow for granular searches when incorporating search operators. An example of a granular search operator is “&” which ensures that key search words must appear in the same line of the search results

[19] https://www.yandex.com/

and “&&”, which specifies the key search words must appear on the same webpage. Another effective search parameter combines the “/” with the “&&” and this returns results where the keywords are within a certain number of sentences from each other. As an example the following key word search, “digital forensics” && /2 OSINT will return web pages where the words “digital forensics” are two sentences from the word “OSINT”. Another useful search operator is “!”, which returns the exact form of the word provided.

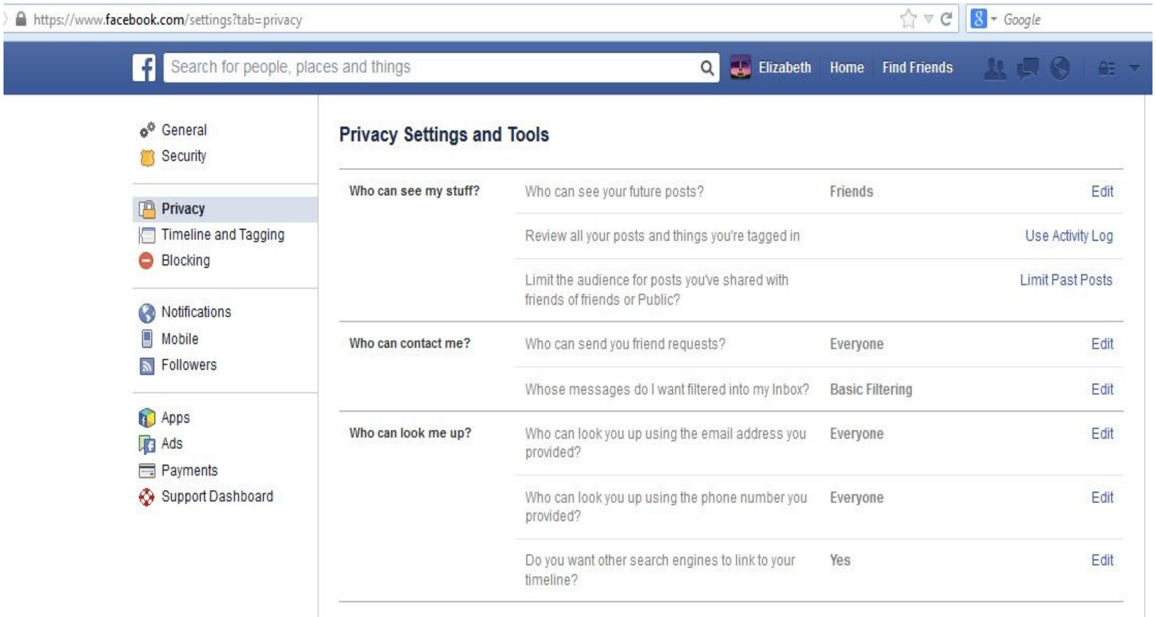
When initiating the process of retrieval or collection, the first step is to identify and gather as much information about the individual or organisation as possible. The information that has been established during the identification process such as a first name, last name, and known location can be used as a starting point to gain further knowledge and gather additional information.

When performing a search on an individual using the individual’s first and last names, it is advisable to first conduct a web search using a web search engine such as Yandex. The search results returned should then be investigated further as there could be useful and revealing information made available from some of the web links gleaned from the web search. A further search can be performed using social networking tools such as the search facilities in LinkedIn and Facebook.

When performing a simple search in Facebook and no results are returned, the advanced search feature in Facebook can be used (see Fig. 8). This feature searches using criteria such as hometown, current city, university and employer. Once the Facebook profile of the individual being searched for has been located, an investigator will be able to view information that the profile owner has made public.

If the privacy settings on the Facebook profile have been secured (see Fig. 6 and Fig. 7), it is recommended to search through some of the friends of the Facebook profile as there is a possibility that one of these Facebook profiles is not as secure and information about the individual may have been shared via friends’ with unsecured Facebook profiles. As illustration of a possible scenario, an individual being investigated could possibly have been tagged in a photograph uploaded by a Facebook friend. If the photograph was uploaded via a computing device with geo-location services enabled, the following information would then be available: the location where the individual was, name tagging information of any other associates and friends in the same photograph or photographs, and the time and date that the photograph was uploaded. Some Facebook users also make use of the Facebook check-in service (see Fig. 9) which uses the location services of a smartphone if the owner has enabled this feature on his/her smartphone.

As of the 22nd May 2014[20], all new Facebook profiles will by default have their privacy settings set to "Friends" instead of "Public", which has been the default for the last few years. Facebook has also stated that over the latter part of 2014 they will be releasing a privacy check-up tool, which will guide Facebook users through a few steps to review their privacy settings, such as what information they are sharing and to whom they are posting.

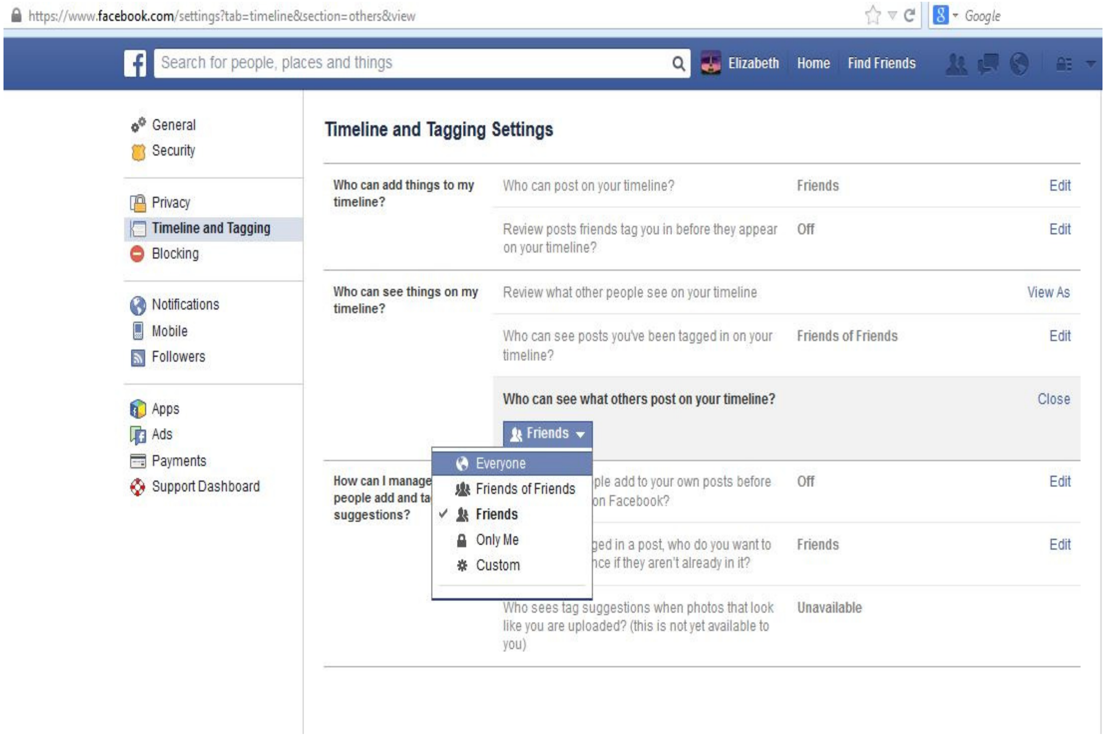


Privacy Settings and Tools

Who can see my stuff?	Who can see your future posts?	Friends	Edit
	Review all your posts and things you're tagged in		Use Activity Log
	Limit the audience for posts you've shared with friends of friends or Public?		Limit Past Posts
Who can contact me?	Who can send you friend requests?	Everyone	Edit
	Whose messages do I want filtered into my Inbox?	Basic Filtering	Edit
Who can look me up?	Who can look you up using the email address you provided?	Everyone	Edit
	Who can look you up using the phone number you provided?	Everyone	Edit
	Do you want other search engines to link to your timeline?	Yes	Edit

Fig. 6. Facebook profile privacy settings.

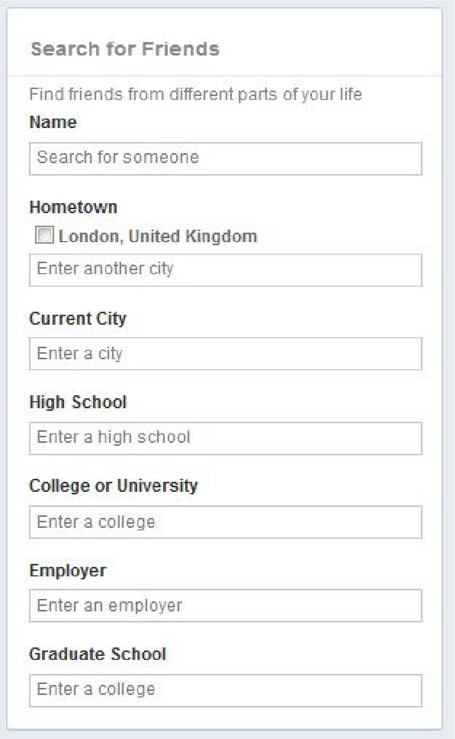
[20] http://newsroom.fb.com/news/2014/05/making-it-easier-to-share-with-who-you-want/



Timeline and Tagging Settings

Who can add things to my timeline?	Who can post on your timeline?	Friends	Edit
	Review posts friends tag you in before they appear on your timeline?	Off	Edit
Who can see things on my timeline?	Review what other people see on your timeline		View As
	Who can see posts you've been tagged in on your timeline?	Friends of Friends	Edit
	Who can see what others post on your timeline? Friends		Close
How can I manage people add and ta… suggestions?	…ple add to your own posts before …on Facebook?	Off	Edit
	…ged in a post, who do you want to …nce if they aren't already in it?	Friends	Edit
	Who sees tag suggestions when photos that look like you are uploaded? (this is not yet available to you)	Unavailable	

Fig. 7. Facebook profile privacy settings for the timeline and tagging of photographs.



Search for Friends

Find friends from different parts of your life

Name

Search for someone

Hometown

□ London, United Kingdom

Enter another city

Current City

Enter a city

High School

Enter a high school

College or University

Enter a college

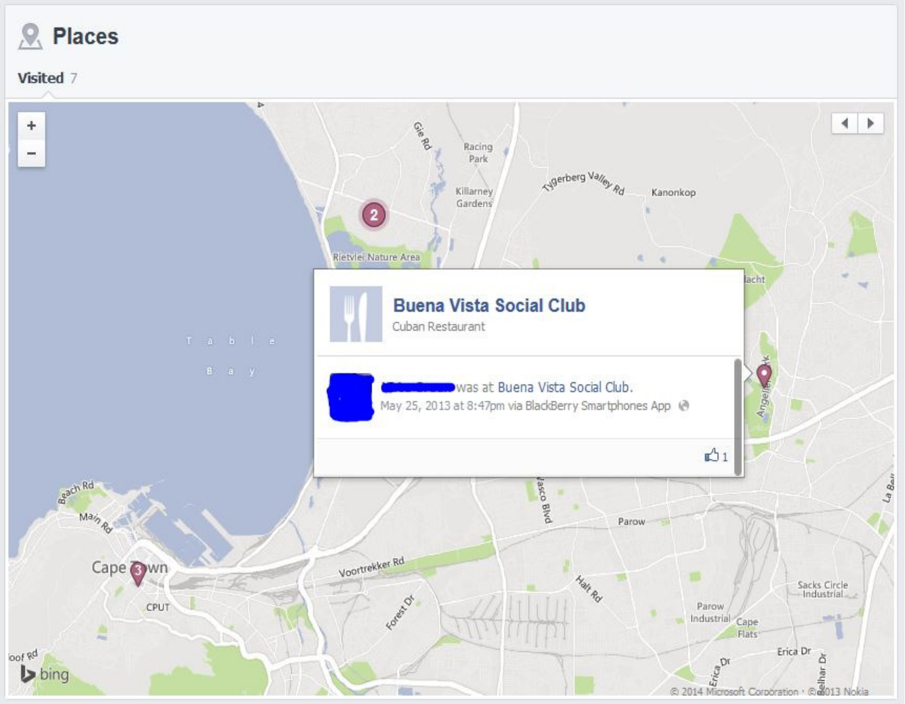
Employer

Enter an employer

Graduate School

Enter a college

Fig. 8. Facebook advanced search criteria.


Places

Visited 7

Buena Vista Social Club

Cuban Restaurant

[redacted] [redacted] was at Buena Vista Social Club.

May 25, 2013 at 8:47pm via BlackBerry Smartphones App

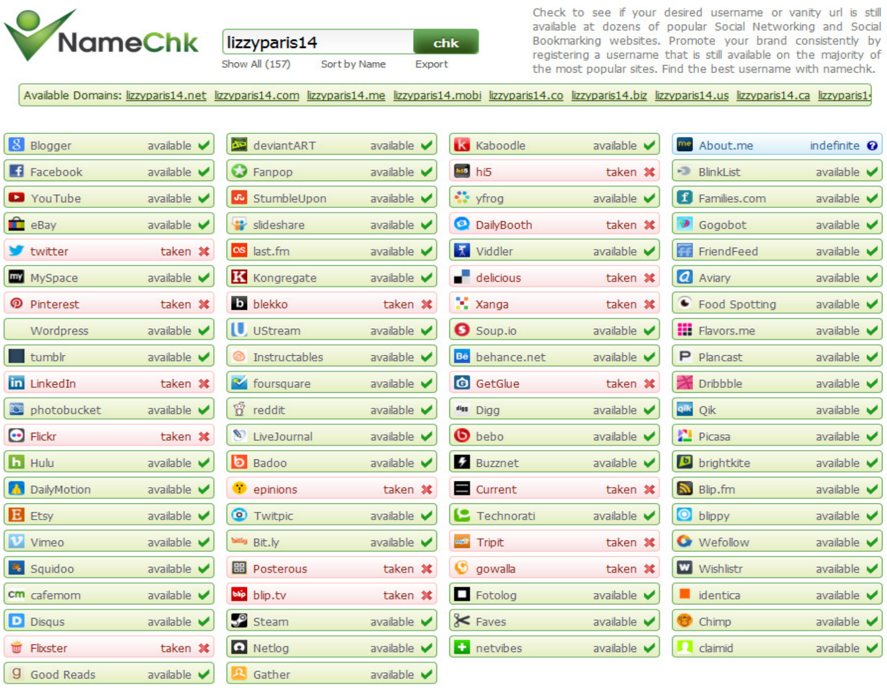
Fig. 9. Facebook check-in feature using a smartphone.

Another useful piece of information to have when conducting searches via the web is the name that an individual chooses to use for his/her online activities, known as the screen name, username, online handle or online alias. It is common practice for social media users to use the same username or online alias across multiple social networking platforms. Once the screen name or online alias has been determined for an individual, an online search using a web search engine should be conducted or the online tool known as "NameChk"[21] can be used in conjunction with the known online alias.

The web tool "NameChk" allows one to enter an online alias to see the availability of a particular online alias across various social networking sites. If the online alias is in use, the link can be selected and it will open to the social networking site that has been registered with that particular online alias, allowing the investigator to determine whether a site does indeed belong the individual being searched. Fig. 10 shows a

[21] http://nameck.com

screenshot of NameChk displaying the results of a search using the online alias “lizzyparis14”.



NameChk

lizzyparis14 chk

Show All (157) Sort by Name Export

Check to see if your desired username or vanity url is still available at dozens of popular Social Networking and Social Bookmarking websites. Promote your brand consistently by registering a username that is still available on the majority of the most popular sites. Find the best username with namechk.

Available Domains: lizzyparis14.net lizzyparis14.com lizzyparis14.me lizzyparis14.mobi lizzyparis14.co lizzyparis14.biz lizzyparis14.us lizzyparis14.ca lizzyparis1

Blogger	available ✓	deviantART	available ✓	Kaboodle	available ✓	About.me	indefinite ?
Facebook	available ✓	Fanpop	available ✓	hi5	taken ✖	BlinkList	available ✓
YouTube	available ✓	StumbleUpon	available ✓	yfrog	available ✓	Families.com	available ✓
eBay	available ✓	slideshare	available ✓	DailyBooth	taken ✖	Gogobot	available ✓
twitter	taken ✖	last.fm	available ✓	Viddler	available ✓	FriendFeed	available ✓
MySpace	available ✓	Kongregate	available ✓	delicious	taken ✖	Aviary	available ✓
Pinterest	taken ✖	blekko	taken ✖	Xanga	taken ✖	Food Spotting	available ✓
Wordpress	available ✓	UStream	available ✓	Soup.io	available ✓	Flavors.me	available ✓
tumblr	available ✓	Instructables	available ✓	behance.net	available ✓	Plancast	available ✓
LinkedIn	taken ✖	foursquare	available ✓	GetGlue	taken ✖	Dribbble	available ✓
photobucket	available ✓	reddit	available ✓	Digg	available ✓	Qik	available ✓
Flickr	taken ✖	LiveJournal	available ✓	bebo	available ✓	Picasa	available ✓
Hulu	available ✓	Badoo	available ✓	Buzznet	available ✓	brightkite	available ✓
DailyMotion	available ✓	epinions	taken ✖	Current	taken ✖	Blip.fm	available ✓
Etsy	available ✓	Twitpic	available ✓	Technorati	available ✓	blippy	available ✓
Vimeo	available ✓	Bit.ly	available ✓	Tripit	taken ✖	Wefollow	available ✓
Squidoo	available ✓	Posterous	taken ✖	gowalla	taken ✖	Wishlistr	available ✓
cafemom	available ✓	blip.tv	taken ✖	Fotolog	available ✓	identica	available ✓
Disqus	available ✓	Steam	available ✓	Faves	available ✓	Chimp	available ✓
Flixster	taken ✖	Netlog	available ✓	netvibes	available ✓	claimid	available ✓
Good Reads	available ✓	Gather	available ✓				

Fig. 10. Results of an online alias search using NameChk.

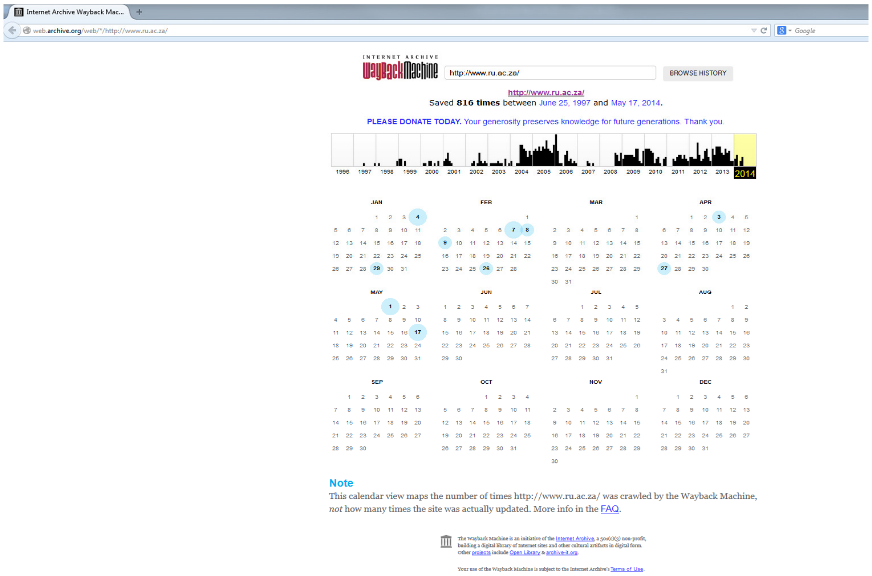
Some additional tools that can be used to collect and gather information are discussed below.

Belkasoft Facebook Profile Saver[22] is a free tool that captures information publically available in Facebook profiles. The Belkasoft Facebook Profile Saver allows one to save any publically available wall contents and photo albums of a user including comments and descriptions. The information is extracted and an HTML report is generated containing the original links in case the photo album or wall post are to be reviewed at a later stage.

[22] http://forensic.belkasoft.com/en/facebook_profile_saver

The digital Internet archiving tool known as the Wayback machine[23] has been building an Internet archive since 1996 and has archived 400 billion web pages as of the 9th March 2014. The Internet archive contains web pages, movies, audio, text, and software, allowing one to search for archived information from 1996. However, not all websites allow archiving; if the website does not allow the *robots.txt* protocol the website will not have been archived by the Wayback machine. An example is Facebook, which does not allow the *robot.txt* protocol to run on its site without written permission. However, if an investigator were to enter the URL of a website that has been archived by the Wayback machine, he/she would be able to access the webpage as it appeared on each of the dates on which the Wayback machine initiated archiving of the website.

Fig. 11 shows an example of the output of the Wayback machine with the number of times the website http://www.ru.ac.za has been archived and the dates on which the archived website can be retrieved. Fig. 12 displays a copy of the archived website http://www.ru.ac.za as at the 12th June 2004.



http://www.ru.ac.za/

Saved **816 times** between June 25, 1997 and May 17, 2014.

PLEASE DONATE TODAY. Your generosity preserves knowledge for future generations. Thank you.

Note

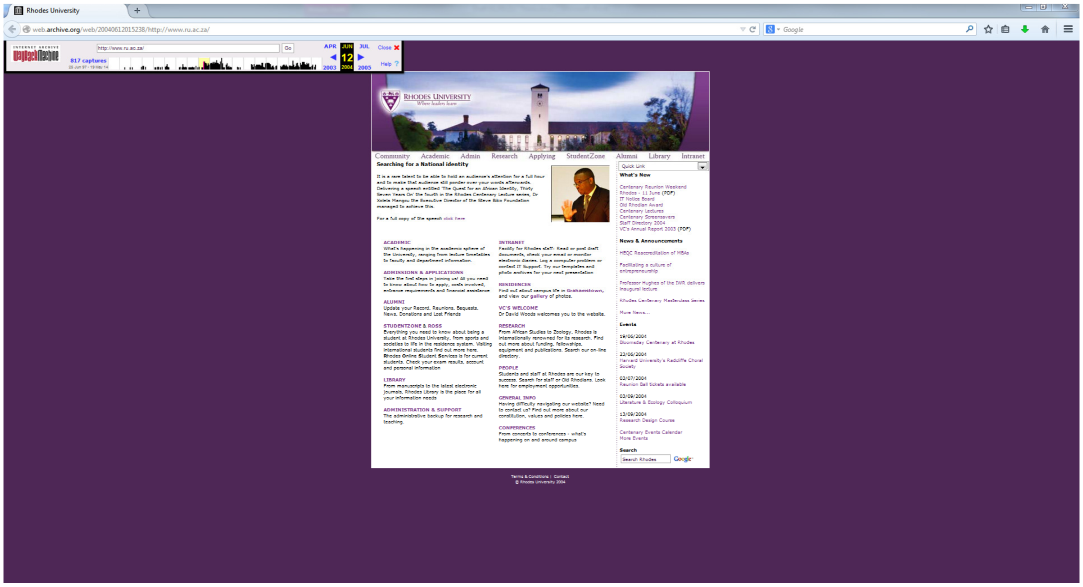
This calendar view maps the number of times http://www.ru.ac.za/ was crawled by the Wayback Machine, *not* how many times the site was actually updated. More info in the FAQ.

The Wayback Machine is an initiative of the Internet Archive, a 501(c)(3) non-profit, building a digital library of Internet sites and other cultural artifacts in digital form. Other projects include Open Library & archive-it.org.

Your use of the Wayback Machine is subject to the Internet Archive's Terms of Use.

Fig. 11. Results of doing a search for ru.ac.za via the Wayback machine.

[23] http://archive.org/web/



Searching for a National identity

It is a rare talent to be able to hold an audience's attention for a full hour and to make that audience still ponder over your words afterwards. Delivering a speech entitled 'The Quest for an African Identity, Thirty Seven Years On' the fourth in the Rhodes Centenary Lecture series, Dr Xolela Mangcu the Executive Director of the Steve Biko Foundation managed to achieve this.

For a full copy of the speech click here

ACADEMIC
What's happening in the academic sphere of the University, ranging from lecture timetables to faculty and department information.

ADMISSIONS & APPLICATIONS
Take the first steps in joining us! All you need to know about how to apply, costs involved, entrance requirements and financial assistance

ALUMNI
Update your Record, Reunions, Bequests, News, Donations and Lost Friends

STUDENTZONE & ROSS
Everything you need to know about being a student at Rhodes University, from sports and societies to life in the residence system. Visiting international students find out more here. Rhodes Online Student Services is for current students. Check your exam results, account and personal information

LIBRARY
From manuscripts to the latest electronic journals, Rhodes Library is the place for all your information needs

ADMINISTRATION & SUPPORT
The administrative backup for research and teaching.

INTRANET
Facility for Rhodes staff: Read or post draft documents, check your email or monitor electronic diaries. Log a computer problem or contact IT Support. Try our templates and photo archives for your next presentation

RESIDENCES
Find out about campus life in Grahamstown, and view our gallery of photos.

VC'S WELCOME
Dr David Woods welcomes you to the website.

RESEARCH
From African Studies to Zoology, Rhodes is internationally renowned for its research. Find out more about funding, fellowships, equipment and publications. Search our on-line directory.

PEOPLE
Students and staff at Rhodes are our key to success. Search for staff or Old Rhodians. Look here for employment opportunities.

GENERAL INFO
Having difficulty navigating our website? Need to contact us? Find out more about our constitution, values and policies here.

CONFERENCES
From concerts to conferences - what's happening on and around campus

What's New

Centenary Reunion Weekend
Rhodos - 11 June (PDF)
IT Notice Board
Old Rhodian Award
Centenary Lectures
Centenary Screensavers
Staff Directory 2004
VC's Annual Report 2003 (PDF)

News & Announcements

HEQC Reaccreditation of MBAs

Facilitating a culture of entrepreneurship

Professor Hughes of the IWR delivers inaugural lecture

Rhodes Centenary Masterclass Series

More News...

Events

19/06/2004
Bloomsday Centenary at Rhodes

23/06/2004
Harvard University's Radcliffe Choral Society

03/07/2004
Reunion Ball tickets available

03/09/2004
Literature & Ecology Colloquium

13/09/2004
Research Design Course

Centenary Events Calendar
More Events

Search
Search Rhodes

Terms & Conditions | Contact
© Rhodes University 2004

Fig. 12. The website http://www.ru.ac.za retrieved from an archive available on the Wayback machine.

Some other useful tools that are also freely available are domain registration websites such as *.co.za*[24] or *whois.net*[25]. Domain registration sites can contain valuable personal information as sometimes, when individuals register a domain, they include their physical address, contact numbers and full name.

Another valuable online tool is Robtex[26], which is a website that is able to provide the same information that can be obtained from a domain registration website but with additional information of the IP address and servers used to host the website.

There is a vast amount of relevant evidence available via OSINT; however, obtaining this information can be challenging for digital forensic investigators who are only familiar with traditional digital forensics making use of evidence obtained from physical devices.

[24] http://co.za/

[25] http://whois.net/

[26] https://www.robtex.com

Below are some relevant points and criteria to consider when performing the step Retrieve and Collect.

- Use a screencast tool to capture the social networking investigation.
- Ensure that the evidence collection is legal and ethical, especially if the evidence is to be presented in a court of law. This would also include adhering to the terms and conditions of the OSINT source to ensure these are not violated.
- Ensure that when the evidence is collected, it is not tainted to ensure that it is legally admissible.
- It is advisable to start the initial investigation using a web search engine and then narrow the search to specific OSINT sources depending on the mandate of the investigation.

5.4 Step 3 Analyse and Process

During the analyse phase, the analysis principle from the McKemmish four principle rule should be considered. According to McKemmish, the extraction, processing and interpretation of the digital evidence forms the primary activity of a digital forensic investigation. McKemmish's analysis principle addresses the process of converting machine data into a format that can be read and understood by humans. (Mckemmish, 1999)

Once the evidence has been collected or exploited, the next step is to analyse and process the evidence. When analysing the evidence, especially if the majority of evidence has been gathered from OSINT sources, it is important to take the evidence certainty of the OSINT source into consideration.

There are two principles that should be followed when processing and analysing OSINT evidence: the first is to make use of the OSINT source checklist described in Chapter 2 and the second is to evaluate the evidence against the evidence certainty scale also described in Chapter 2.

The OSINT checklist provides baselines against which the OSINT source can be evaluated. The following baselines are suggested when evaluating evidence obtained from an OSINT source: the authority of the OSINT source, the accuracy of the OSINT source, the objectivity of the OSINT source, the absence or presence of a time and date stamp for the OSINT source and, lastly, the relevance of the OSINT source.

The evidence certainty scale provides a scale with which to measure the certainty of any evidence that has been obtained. The scale ranges between zero and six, with zero holding the lowest value for the evidence because the evidence is contradictory to known facts, and six reflecting evidence that is tamperproof and holds a high level of assurance.

Once the investigator has determined which OSINT evidence can be used by evaluating the evidence against the OSINT checklist and the evidence certainty scale, the next step is to start building the forensic case. This is achieved by analysing the evidence and searching for multiple evidence sources that provide corroboration: for example, a tweet via Twitter by a suspect placing them at a location and evidence from Facebook such as the suspect being tagged in a photo taken at the same location.

Below is a checklist that can be applied when performing the step Analyse and Process:

- Review the evidence using the evidence certainty scale.
- Always try to obtain evidence that is tamperproof and contains a high level of assurance.
- Check whether the evidence can be corroborated via several independent sources.
- Discard evidence that is contradictory to known facts.
- Use OSINT sources, which command authority, are time and date stamped and are known for their accuracy.

5.5 Step 4 Visualise

An early Chinese proverb states: “One picture is worth more than ten thousand words”. Following this advice the fourth step is visualise. After the processing and analysing stage, there are likely to be numerous evidence items and therefore an important part of a digital forensic investigation is to reconstruct a timeline of events.

A timeline is a valuable tool as it allows a digital forensic investigator to prove or disprove any hypothetical model proposed for the investigation. The timeline can also provide support for the mandate the digital forensic investigator received prior to commencing the investigation (Ieong, 2006).

Providing answers to the "When" aspect of the investigation is easily accomplished using the timeline. For example, a timeline can assist in determining when an incident took place, when the suspect sent his/her first tweet, or when a suspect was tagged in a photo uploaded via Instagram or Facebook.

One of the popular open source timeline tools is *Log2timeline*[27]. This tool is useful for parsing log files from various systems; the output is a CSV (comma separated value) formatted file, which can be opened in Microsoft Excel making it easy to analyse. Unfortunately, if the evidence items are not in the form of a log file, a manual timeline must be created, for which Microsoft Excel can again be used.

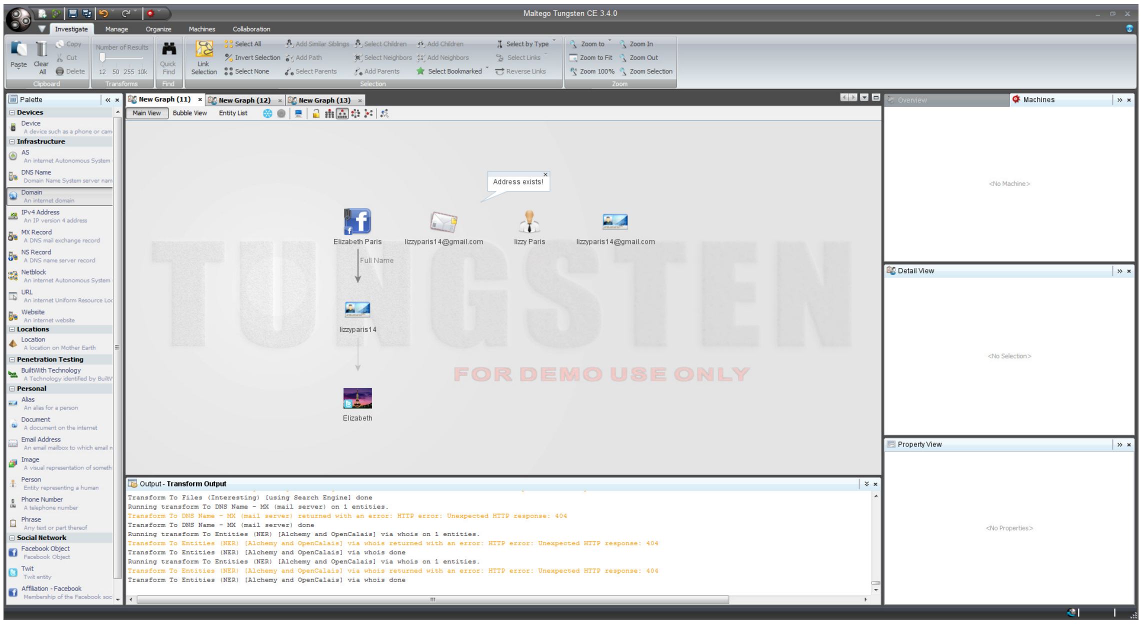
Bradbury (2011) confirms the importance of visualisation in social networks owing to the large amount of data and relationships available in these networks. An example of the type of valuable information that a visualisation tool can provide, is the number of people that are indirectly connected.

A useful visualisation tool is Maltego, which is an OSINT and forensic tool. Maltego provides a mechanism for mining and gathering information using OSINT via the Internet and is able to illustrate the information in an easily understandable format.

Maltego, which consists of search criteria code known as transforms, searches for links between people, social networks, companies and various Internet infrastructures. It then displays the search results graphically. For example, if an online alias, email address, or company name can be determined from the evidence obtained during the processing and analysing stage, this can be input into Maltego to perform a search. Fig. 13 below is a screenshot of a Maltego graph for searches conducted for "Lizzy Paris". The results of the search may corroborate the evidence that the digital forensic

[27] www.log2timeline.net

investigator has already obtained. Maltego offers a community version, but this version only allows for a maximum of twelve results per transform. Commonly, when searching social networks the number of returned results usually exceed this limitation.



```
Transform To Files (Interesting) [using Search Engine] done
Running transform To DNS Name - MX (mail server) on 1 entities.
Transform To DNS Name - MX (mail server) returned with an error: HTTP error: Unexpected HTTP response: 404
Transform To DNS Name - MX (mail server) done
Running transform To Entities (NER) [Alchemy and OpenCalais] via whois on 1 entities.
Transform To Entities (NER) [Alchemy and OpenCalais] via whois returned with an error: HTTP error: Unexpected HTTP response: 404
Transform To Entities (NER) [Alchemy and OpenCalais] via whois done
Running transform To Entities (NER) [Alchemy and OpenCalais] via whois on 1 entities.
Transform To Entities (NER) [Alchemy and OpenCalais] via whois returned with an error: HTTP error: Unexpected HTTP response: 404
Transform To Entities (NER) [Alchemy and OpenCalais] via whois done
```

Fig. 13. A screenshot of a Maltego graph for searches conducted for “Lizzy Paris”.

Below is a summary of the key sub steps to be carried out in the step Visualise:

- Create a timeline for the investigation, as this helps order the events as they occurred.
- Use a visualisation tool such as Maltego to create a visual map of the link between Twitter accounts or alias used across social networking sites.

5.6 Step 5 Collaborate

Collaborate is the fifth phase of the framework. The principle of collaboration is to work together with other investigators or other investigative teams to provide support to each other and collaborate with respect to information and evidence to achieve the goal of corroboration.

Collaboration between different investigators or investigation teams could provide the investigation team with a better understanding of the crime from the corroboration of information and evidence. The supporting evidence can also assist with a probable cause for a search warrant for evidence as well as being able to assist other investigators in identifying other possible criminal activities.

Collaboration also allows investigators to build a stronger case by piecing together all the various evidence items.

Below are the key points that should be considered in the step Collaborate:

- Whenever possible collaboration with other investigators, is preferable as this provides reassurance that all aspects of the investigation have been covered.
- Collaboration between investigators can allow for the corroboration of evidence.

5.7 Step 6 Report

It is the digital forensic investigator's responsibility to document all actions and observations throughout the digital forensic investigation. All documentation should be complete, accurate, factual and comprehensive resulting in a report being written for the intended audience. Report writing is a vital skill although it is often met with antipathy.

When writing the report it is important to state the facts regarding the evidence discovered during the digital forensic investigation. On the other hand, if the evidence was not present, it is equally important not to state otherwise. When writing the conclusion of the report, the facts must be stated and not the opinion of the digital forensic investigator. A recommendation section can be added allowing the forensic investigator to make recommendations based on his/her opinion. These opinions should be based on best practice as far as possible to reduce any potential subjectivity.

The last principle regarding the handling of digital evidence according to Mckemmish (1999) is the presentation of digital evidence in a court of law. This includes the manner in which the evidence is presented, the expertise and qualifications of the expert witness, and the credibility of the processes used to obtain the evidence.

The final report should include at the very least the following information:

- name of the investigation company
- unique case identifier
- date of the report
- identity and signature of the digital forensic investigator
- identity and signature of the digital forensic investigator who peer reviewed the final report
- descriptive list of items submitted for examination, including all hardware and software used during the forensic investigation
- brief description of steps taken during the examination, for example key word searches, graphic image searches, operators used to perform web searches, online alias name searches and OSINT sources
- the results or conclusion of the digital forensic investigation; depending on the scope of the digital forensic investigation, details of the finding can be included or just a summary of the findings
- any recommendations
- exhibits or list of supporting materials such as printouts of particular evidence items, digital copies of the evidence items, and most importantly the chain of custody documentation
- an optional glossary, which can assist the reader of the report to understand any technical terms used, thereby preventing misunderstanding due to ambiguity

Once the report has been drafted it is imperative that it is peer reviewed. The peer review can take place either during the collaboration phase or after the report has been drafted. A peer review of the evidence findings also contributes towards ensuring that objectivity is maintained as a correctly conducted peer review process assists with assessing a digital forensic investigator's findings for bias or other possible weaknesses (Casey, 2011).

Below are the criteria which can be used a checklist for the step report:

- From the start of the investigation keep notes relating to all steps that are followed during the investigation.

- When drafting a report, state the facts and never provide opinion.
- If possible, ensure that the report is peer reviewed, as this ensures objectivity is maintained.

5.8 Summary

This chapter described the six steps in the proposed framework and provided some examples of how each step could be achieved, as well as a checklist to ensure that the framework has been correctly and rigorously applied.

Chapter 6 Conclusion and Future Work

This final chapter summarises the main findings and highlights the contribution of the research. The chapter closes by recommending future research work that will build on this research.

6.1 Summary of Findings

The explosion of the Internet, the increase in the use of mobile devices and the need to always be connected to the Internet provide ample opportunity for cybercrime making it imperative that the traditional digital forensic investigation frameworks are updated.

Digital forensics and the traditional digital forensic investigation process were discussed at length in the literature summary. It was concluded from the literature summary and the survey responses that the current traditional digital forensic investigation process and tools do not support the use of OSINT evidence because the foundation thereof is based on acquiring physical evidence (Richard & Roussev, 2006).

Due to the volume of personal information which social networking users make publically available, the findings of the research confirmed that a framework for utilising OSINT when conducting a digital forensic investigation would be of value. This finding is also supported by Taylor, Haggerty, Gresty, Almond, & Berry (2014), who confirmed that, at present, there are no guidelines for organisations who need to perform digital forensic investigations of social networking applications.

The research confirmed that there is value in the information available from social networking sites and that this information can assist when performing digital forensic investigations and that the development of a framework would be useful for digital forensic investigators.

Based on the survey results a framework was proposed, which can be used as a guide when performing a digital forensic investigation involving OSINT sources. Each of the principles underpinning the framework was defined. Reasons relating to the importance and inclusion of each principle were provided. Lastly, a sequence was applied to each principle detailing the actions and possible tools to be used.

The researcher provided the forensic framework for digital forensic investigators to use as a guide when conducting OSINT forensic investigations. In addition the researcher also provided criteria for each step of the framework. The aim of the proposed framework was to provide digital forensic investigators with a guide that would ensure that a rigorous process is followed and that in future, evidence from OSINT sources can be seen as evidentiary rather than supplementary.

6.2 Contribution of the Research

The answer to the research question, “Can OSINT be used as a digital forensic investigation tool?” has been answered in this thesis. The literature summary, the analysis and results of the survey, and the creation of a framework for OSINT investigations confirm that OSINT can be used a digital forensic investigative tool.

Moreover, the information obtained from the respondents validates the hypothesis that there is value in the information available from OSINT, which can assist digital forensic investigators.

All the research objectives outlined in Chapter 1 have been realised in this thesis.

The first objective was to ascertain whether the information available via social networking and OSINT could indeed assist a digital forensic investigator during their investigation. This was achieved in Chapter 2 by providing a literature review of digital forensics, including the traditional digital forensic framework and how its current structure does not support OSINT investigations. Social networking was explained and a summary of information available freely on some of the social networking platforms was outlined. Examples were also provided describing how social networking is currently being used to conduct criminal activities. Lastly, the value of the information available through OSINT sources was also discussed.
The definition of the framework for digital forensic investigators to apply when using OSINT to assist with a digital forensic examination (given in Chapter 5) addressed the second objective.

The third objective was to determine the criteria for websites that can be considered social networking sites, and identifying the type of information or evidence items that can be obtained from social networking sites. Various OSINT sources and scenarios relating to the way in which these have been used in digital forensic investigations were also presented in Chapters 2 and 5. In addition, information available from some of the OSINT sources, and how to gain access to it, was also discussed in Chapter 5.

The fourth objective, that is, to outline some of the shortcomings of OSINT and factors that may serve to mitigate these shortcomings, was covered in Chapters 2 and 5.

6.3 Future Research and Recommendations

There is much scope for further research in the areas of digital forensic investigations using OSINT and, more specifically, social networks. For example, Casey's evidence certainty scale should be expanded to include OSINT sources. A starting point for research in this regard could be to incorporate the OSINT source checklist into the existing evidence certainty scale.

References

Adelstein, F. (2006). Live forensics: diagnosing your system without killing it first. *Communications of the ACM*, *49*(2), 63–66.

Al Mutawa, N., Baggili, I., & Marrington, A. (2012). Forensic analysis of social networking applications on mobile devices. *Digital Investigation*, *9*, S24–S33. doi:10.1016/j.diin.2012.05.007

Altheide, C., & Carvey, H. (2011). *Digital Forensics with Open Source Tools*. Waltham, MA: Elsevier Inc.

Appel, E. (2011). *Internet Searches for Vetting, Investigations, and Open-Source Intelligence*. CRC Press. doi:10.1201/b10523

Arthur, K. (2010). *Considerations Towards the Development of a Forensic Evidence Management System*. *Master's Thesis*. University of Pretoria. Retrieved from http://upetd.up.ac.za/thesis/available/etd-07232010-192957/unrestricted/dissertation.pdf

Beebe, N. (2009). Digital Forensic Research: The Good, the Bad and the Unaddressed. In *Fifth IFIP WG 11.9 International Conference on Digital Forensics* (pp. 17–36). Orlando.

Best, R. A., & Cumming, A. (2007). Open Source Intelligence (OSINT): Issues for Congress Specialist in National Defense Foreign Affairs. *Report for the Congress*.

Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. doi:10.1111/j.1083-6101.2007.00393.x

Bradbury, D. (2011). Data mining with LinkedIn. *Computer Fraud & Security*, *2011*(10), 5–8. doi:10.1016/S1361-3723(11)70101-4

Brunty, J., & Helenek, K. (2013). *Social Media Investigation for Law Enforcement* (First., pp. 1 –112). Elsevier.

Carrier, B. (2002). Open Source Digital Forensics Tools. Retrieved July 07, 2013, from http://www.digital-evidence.org/papers/opensrc_legal.pdf

Casey, E. (2002). Error, Uncertainty, and Loss in Digital Evidence. *International Journal of Digital Evidence*, *1*(2).

Casey, E. (2004). *Digital Evidence and Computer Crime* (second., p. 690). Academic Press, Elsevier.

Casey, E. (2011). *Digital Evidence and Computer Crime* (3rd ed., p. 837). Baltimore, Maryland, USA: Academic Press, Elsevier.

Choo, K. R., Smith, R., & McCusker, R. (2007). Future directions in technology-enabled crime. *Australian Institute of Criminology 2007*, (78).

Clough, J. (2010). *Principles of Cybercrime*. Cambridge: Cambridge University Press.

Cohen, M. I., Bilby, D., & Caronni, G. (2011). Distributed forensics and incident response in the enterprise. *Digital Investigation*, *8*, S101–S110. doi:10.1016/j.diin.2011.05.012

Cuijpers, C. (2013). Legal aspects of open source intelligence – Results of the VIRTUOSO project. *Computer Law & Security Review*, *29*(6), 642–653. doi:10.1016/j.clsr.2013.09.002

Dixon, P. (2005). An overview of computer forensics meddling with or exploitation of data. *IEEE Potentials*, (December), 7–10.

Duncan, S. H. (2008). Myspace Is Also Their Space: Ideas for Keeping Children Safe from Sexual Predators on Social-Networking Sites. *Kentucky Law Journal*, *96*, 527 – 577.

Dykstra, J., & Sherman, A. T. (2012). Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques. *Digital Investigation*, *9*, S90–S98. doi:10.1016/j.diin.2012.05.001

Edgar-Nevill, D., & Qi, M. (2011). Social networking searching and privacy issues. *Information Security Technical Report*, *16*(2), 74–78. doi:10.1016/j.istr.2011.09.005

Ellickson, J., & Lynch, J. (2010). Obtaining and Using Evidence from Social Networking Sites. Retrieved May 27, 2013, from https://www.eff.org/files/filenode/social_network/20100303__crim_socialnetworking.pdf

Garfinkel, S. L, Farrell, P., Roussev, V., & Dinolt, G. (2009). Bringing science to digital forensics with standardized forensic corpora. *Digital Investigation*, *6*, S2–S11. doi:10.1016/j.diin.2009.06.016

Garfinkel, S. L. (2010). Digital forensics research: The next 10 years. *Digital Investigation*, *7*, S64–S73. doi:10.1016/j.diin.2010.05.009

Garfinkel, S. L. (2013). Digital media triage with bulk data analysis and bulk_extractor. *Computers & Security*, *32*, 56–72. doi:10.1016/j.cose.2012.09.011

Gibson, S. (2004). An Intelligence Lifeline. Retrieved October 06, 2013, from http://www.rusi.org/downloads/assets/JA00365.pdf

Glassman, M. (2012). An era of webs: Technique, technology and the new cognitive (r)evolution. *New Ideas in Psychology*, *30*(3), 308–318. doi:10.1016/j.newideapsych.2012.05.002

Glassman, M., & Kang, M. J. (2012). Intelligence in the Internet age: The emergence and evolution of Open Source Intelligence (OSINT). *Computers in Human Behavior*, *28*(2), 673–682. doi:10.1016/j.chb.2011.11.014

Huber, M., Mulazzani, M., Leithner, M., Schrittwieser, S., Wondracek, G., Weippl, E., & Fl, O. (2011). Social Snapshots Digital Forensics for Online Social Networks. In *Annual Computer Security Applications Conference (ACSAC)*.

Hulnick, A. S. (2010). The Downside of Open Source Intelligence. *International Journal of Intelligence and CounterIntelligence*, 37–41.

Ieong, R. S. C. (2006). FORZA – Digital forensics investigation framework that incorporate legal issues. *Digital Investigation*, *3*, 29–36. doi:10.1016/j.diin.2006.06.004

Jordaan, J. (2012). A Sample of Digital Forensic Quality Assurance in the South African Criminal Justice System. In *Information Security for South Africa ISSA* (pp. 1–7).

Kim, Y., & Glassman, M. (2013). Beyond search and communication: Development and validation of the Internet Self-efficacy Scale (ISS). *Computers in Human Behavior*, *29*(4), 1421–1429. doi:10.1016/j.chb.2013.01.018

Koen, R. (2009). *The development of an open-source forensics platform*. University of Pretoria. Retrieved from http://upetd.up.ac.za/thesis/available/etd-02172009-014722/unrestricted/dissertation.pdf

Kohn, M., Eloff, J., & Olivier, M. (2006). Framework for a Digital Forensic Investigation. Retrieved July 30, 2014, from http://icsa.cs.up.ac.za/issa/2006/Proceedings/Full/101_Paper.pdf

Koops, B.-J., Hoepman, J.-H., & Leenes, R. (2013). Open-source intelligence and privacy by design. *Computer Law & Security Review*, *29*(6), 676–688. doi:10.1016/j.clsr.2013.09.005

Lang, J. C., & De Sterck, H. (2014). The Arab Spring: A simple compartmental model for the dynamics of a revolution. *Mathematical Social Sciences*, *69*, 12–21. doi:10.1016/j.mathsocsci.2014.01.004

LexisNexis. (2012). Law Enforcement Personnel Use of Social Media in Investigations : Summary of Findings. Retrieved November 08, 2013, from http://www.lexisnexis.com/government/investigations/

Long, J. (2008). Google Hacking for Penetration Testers Volume 2. Retrieved June 16, 2013, from http://mkbonlinereputatie.nl/wp-content/uploads/2014/02/Google-advanced-search.pdf

Marsico, E. M. (2010). Social Networking Websites: Are Myspace and Facebook the Fingerprints of the Twenty-First Century? *Edward M. Widener Law Journal*, *19*(3), 967–976.

Mckemmish, R. (1999). What is Forensic Computing? Retrieved November 17, 2013, from http://www.aic.gov.au/documents/9/C/A/{9CA41AE8-EADB-4BBF-9894-64E0DF87BDF7}ti118.pdf

Mulazzani, M., Huber, M., & Weippl, E. (2012). Social Network Forensics : Tapping the Data Pool of Social Networks. Retrieved November 04, 2012, from https://www.sba-research.org/wp-content/uploads/publications/socialForensics_preprint.pdf

Neri, F., Geraci, P., & Pettoni, M. (2011). Stalker: overcoming linguistic barriers in open source intelligence. *International Journal of Networking and Virtual Organisations*, *8*(1/2), 37. doi:10.1504/IJNVO.2011.037160

NW3C. (2013). Criminal Use of Social Media (2013). Retrieved August 02, 2014, from http://www.nw3c.org/docs/whitepapers/criminal-use-of-social-media.pdf

O'Connor, T. (2010). *About Face : Defending Your Organization Against Penetration Testing Teams*.

O'Connor, T. (2013). *Violent Python A Cookbook for Hackers, Forensic Analysts, Violent Python A Cookbook for Hackers, Forensic Analysts.* Elsevier Inc.

Palmer, G. (2001). *A Road Map for Digital Forensic Research*. *Digital Forensic Research Workshop (DFRWS)* (pp. 1 – 48). Retrieved from http://www.dfrws.org/2001/dfrws-rm-final.pdf

Peisert, S., Sishop, M., & Marzullo, K. (2008). Computer Forensics in Forensics. Systematic Approaches to Digital Forensic Engineering. *IEEE*, 102 – 122.

Politt, M. (2004). Six blind men from Indostan. In *Digital Forensic Research Workshop (DFRWS)*.

Pouchard, L. C., Dobson, J. M., & Trien, J. P. (2006). A Framework for the Systematic Collection of Open Source Intelligence, 102–107.

Quayle, E., & Taylor, M. (2011). Social networking as a nexus for engagement and exploitation of young people. *Information Security Technical Report*, *16*(2), 44–50. doi:10.1016/j.istr.2011.09.006

Qureshi, P. A. R., & Memon, N. (2012). Hybrid model of content extraction. *Journal of Computer and System Sciences*, *78*(4), 1248–1257. doi:10.1016/j.jcss.2011.10.012

Richard, G. G., & Roussev, V. (2006). Digital Forensic Tools: The Next Generation. In *Digital Forensic Tools: The Next Generation* (pp. 76–91). Idea Group Inc.

Saunders, M., Philip, L., & Thornhill, A. (2009). *Research methods for business students* (5th editio., pp. 1 – 604). Harlow, Essex: Pearson Education Ltd.

Starin, D., Baden, R., Bender, A., Spring, N., & Bhattacharjee, B. (2009). Persona : An Online Social Network with User-Defined Privacy Categories and Subject Descriptors. In *Proceedings of the ACM SIGCOMM 2009 conference on Data communication* (pp. 135–146).

Steele, R. D. (2006). The Handbook of Intelligence Studies: Open Source Intelligence (OSINT). Retrieved July 20, 2013, from http://www.oss.net/dynamaster/file_archive/060409/5432a5e19def62b82684a111fe03f899/STEELE OSINT FOR HANDBOOK 3.3 Chapter.doc

Taylor, M., Haggerty, J., Gresty, D., Almond, P., & Berry, T. (2014). Forensic investigation of social networking applications. *Network Security*, *2014*(11), 9–16. doi:10.1016/S1353-4858(14)70112-6

US Dept of Homeland Security, & US Secret Service. (2007). Best Practices For Seizing Electronic Evidence v. 3: A Pocket Guide for First Responders. Retrieved August 23, 2013, from https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=239359

Van Manen, M. (2010). The pedagogy of Momus technologes: Facebook, privacy and online intimacy. *Qualitative Health Research*, *20*(8), 1023–1032.

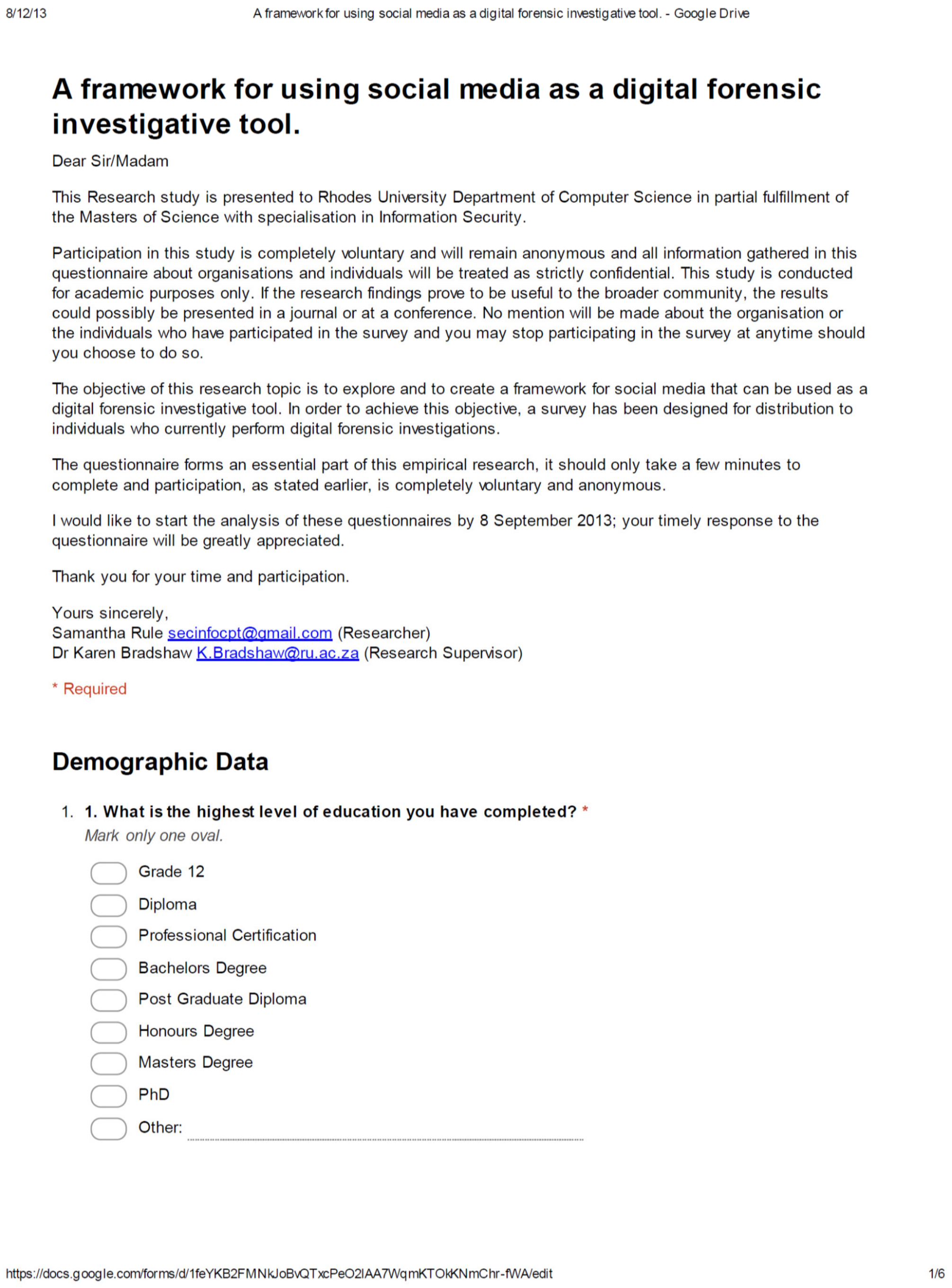
Von Solms, S., Lourens, C., Reekie, C., & Grobler, T. (2006). A Control Framework for Digital Forensics. In *Advances in Digital Forensics II* (pp. 345 –355).

Xu, C., Ryan, S., Prybutok, V., & Wen, C. (2012). It is not for fun: An examination of social network site usage. *Information & Management*, *49*(5), 210–217. doi:10.1016/j.im.2012.05.001

Zainudin, N. M., Merabti, M., & Llewellyn-Jones, D. (2011). A Digital Forensic Investigation Model for Online Social Networking. In *Research and Innovation in Information Systems (ICRIIS)* (pp. 1 – 6).

Appendices

Appendix A: Electronic Survey Questionnaire



8/12/13 A framework for using social media as a digital forensic investigative tool. - Google Drive

A framework for using social media as a digital forensic investigative tool.

Dear Sir/Madam

This Research study is presented to Rhodes University Department of Computer Science in partial fulfillment of the Masters of Science with specialisation in Information Security.

Participation in this study is completely voluntary and will remain anonymous and all information gathered in this questionnaire about organisations and individuals will be treated as strictly confidential. This study is conducted for academic purposes only. If the research findings prove to be useful to the broader community, the results could possibly be presented in a journal or at a conference. No mention will be made about the organisation or the individuals who have participated in the survey and you may stop participating in the survey at anytime should you choose to do so.

The objective of this research topic is to explore and to create a framework for social media that can be used as a digital forensic investigative tool. In order to achieve this objective, a survey has been designed for distribution to individuals who currently perform digital forensic investigations.

The questionnaire forms an essential part of this empirical research, it should only take a few minutes to complete and participation, as stated earlier, is completely voluntary and anonymous.

I would like to start the analysis of these questionnaires by 8 September 2013; your timely response to the questionnaire will be greatly appreciated.

Thank you for your time and participation.

Yours sincerely,
Samantha Rule secinfocpt@gmail.com (Researcher)
Dr Karen Bradshaw K.Bradshaw@ru.ac.za (Research Supervisor)

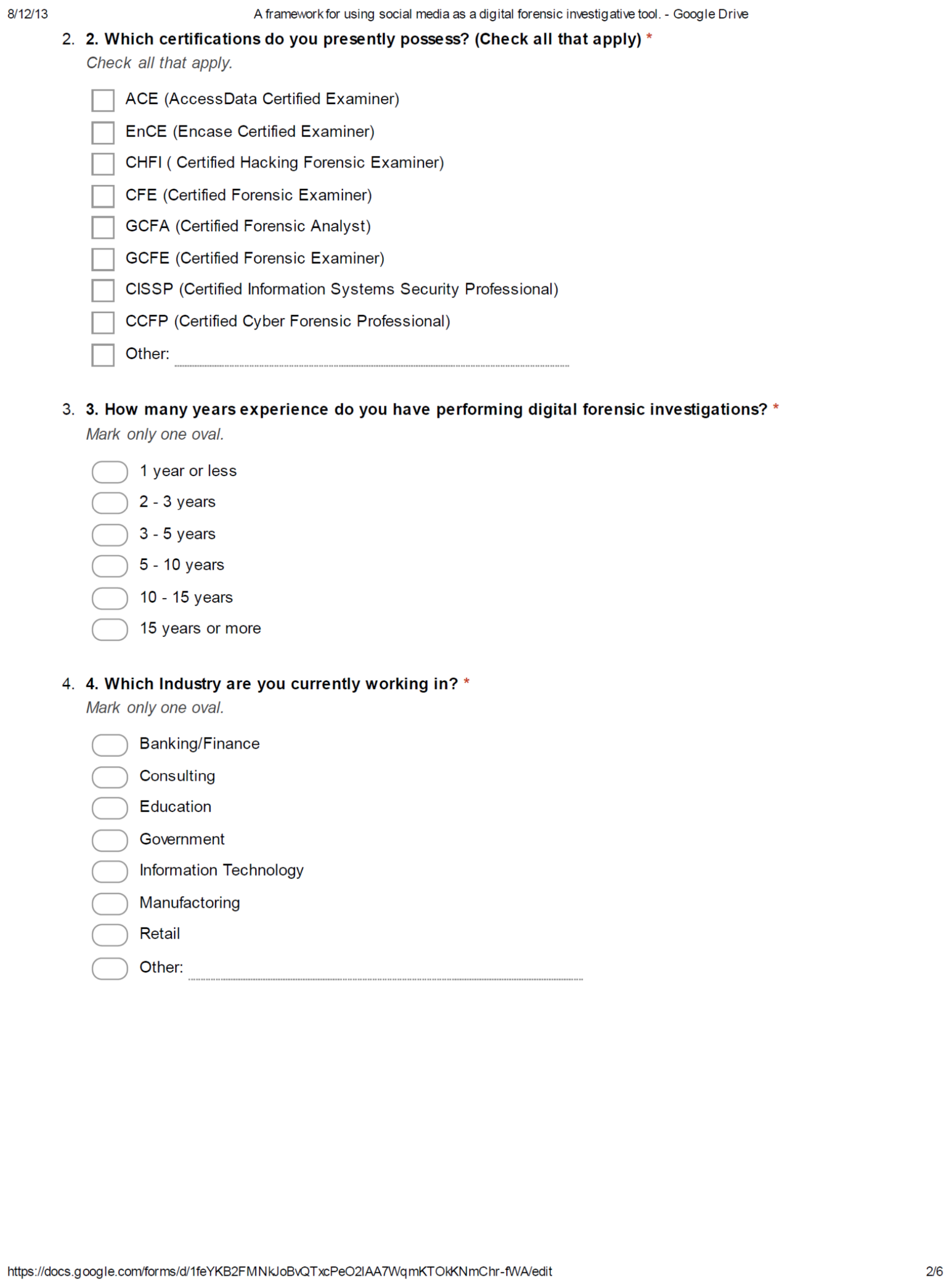
* Required

Demographic Data

1. **1. What is the highest level of education you have completed? ***
Mark only one oval.

○ Grade 12
○ Diploma
○ Professional Certification
○ Bachelors Degree
○ Post Graduate Diploma
○ Honours Degree
○ Masters Degree
○ PhD
○ Other: ______________________

https://docs.google.com/forms/d/1feYKB2FMNkJoBvQTxcPeO2lAA7WqmKTOkKNmChr-fWA/edit 1/6



2. **2. Which certifications do you presently possess? (Check all that apply)** *

Check all that apply.

- □ ACE (AccessData Certified Examiner)
- □ EnCE (Encase Certified Examiner)
- □ CHFI (Certified Hacking Forensic Examiner)
- □ CFE (Certified Forensic Examiner)
- □ GCFA (Certified Forensic Analyst)
- □ GCFE (Certified Forensic Examiner)
- □ CISSP (Certified Information Systems Security Professional)
- □ CCFP (Certified Cyber Forensic Professional)
- □ Other: ____________________

3. **3. How many years experience do you have performing digital forensic investigations?** *

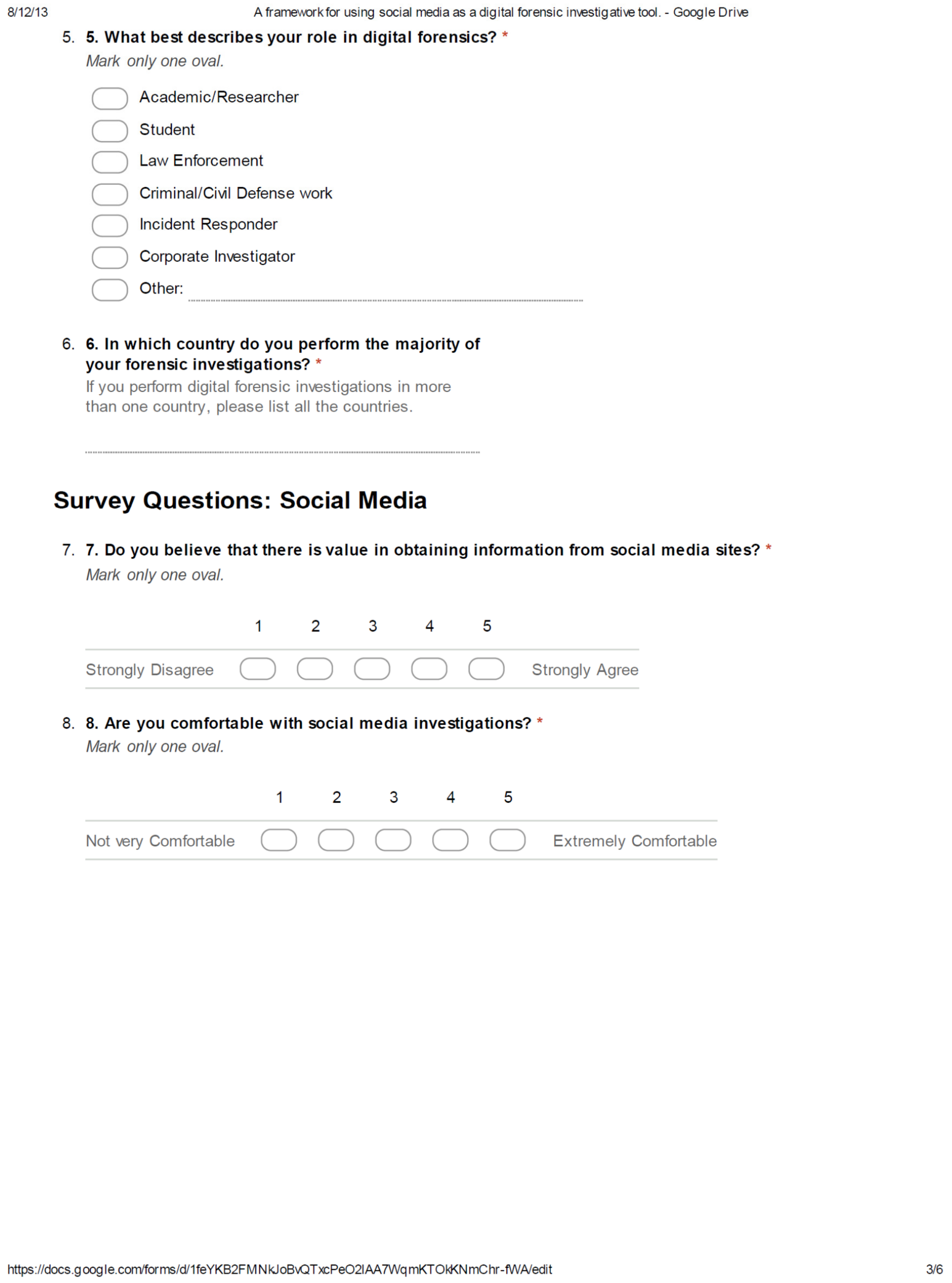
Mark only one oval.

- ○ 1 year or less
- ○ 2 - 3 years
- ○ 3 - 5 years
- ○ 5 - 10 years
- ○ 10 - 15 years
- ○ 15 years or more

4. **4. Which Industry are you currently working in?** *

Mark only one oval.

- ○ Banking/Finance
- ○ Consulting
- ○ Education
- ○ Government
- ○ Information Technology
- ○ Manufactoring
- ○ Retail
- ○ Other: ____________________



8/12/13 A framework for using social media as a digital forensic investigative tool. - Google Drive

5. **5. What best describes your role in digital forensics?** *

   Mark only one oval.

   ○ Academic/Researcher
   ○ Student
   ○ Law Enforcement
   ○ Criminal/Civil Defense work
   ○ Incident Responder
   ○ Corporate Investigator
   ○ Other: ______________________

6. **6. In which country do you perform the majority of your forensic investigations?** *

   If you perform digital forensic investigations in more than one country, please list all the countries.

   ____

Survey Questions: Social Media

7. **7. Do you believe that there is value in obtaining information from social media sites?** *

   Mark only one oval.

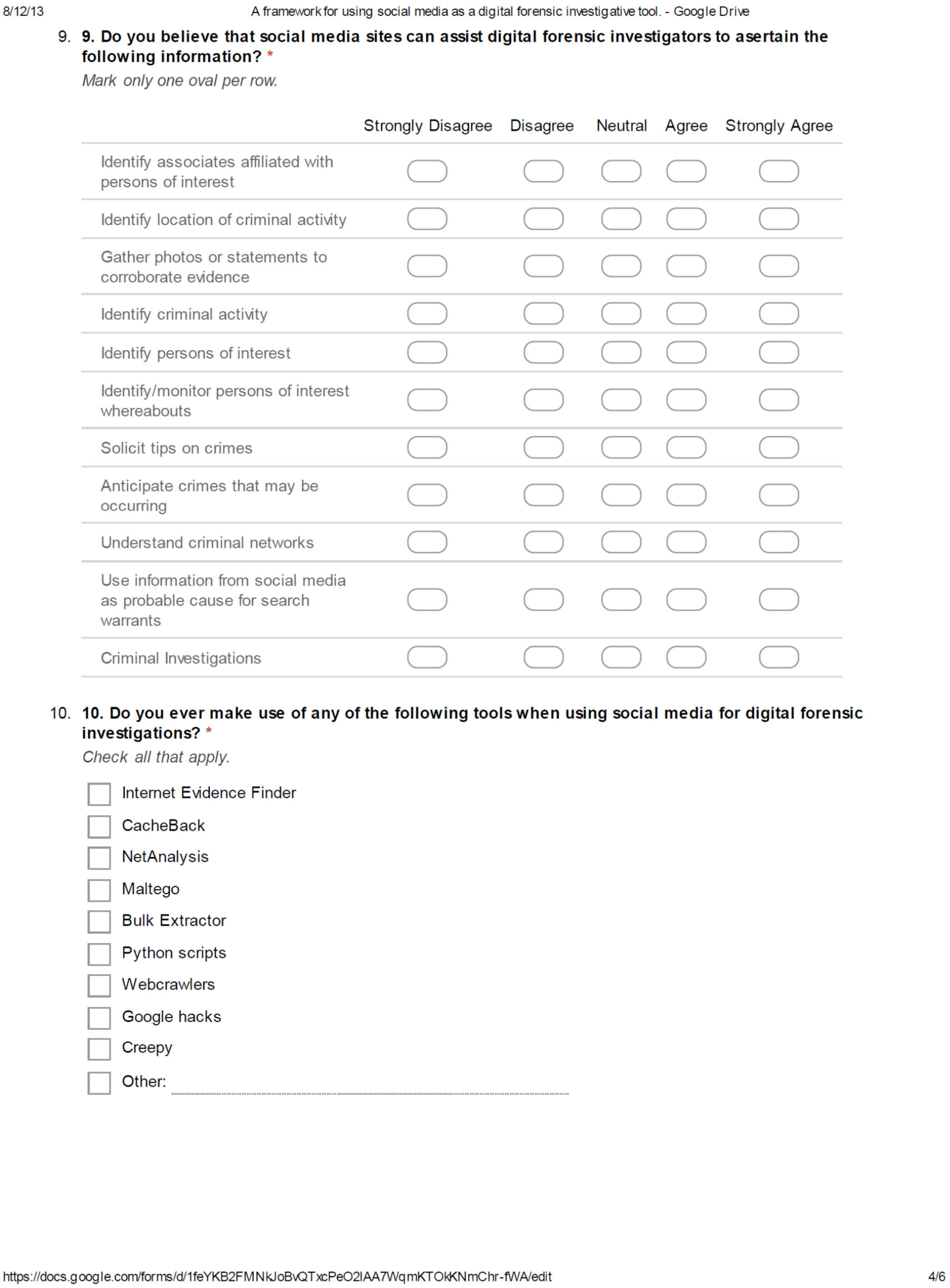
	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

8. **8. Are you comfortable with social media investigations?** *

   Mark only one oval.

	1	2	3	4	5	
Not very Comfortable	○	○	○	○	○	Extremely Comfortable

https://docs.google.com/forms/d/1feYKB2FMNkJoBvQTxcPeO2lAA7WqmKTOkKNmChr-fWA/edit 3/6



9. **9. Do you believe that social media sites can assist digital forensic investigators to asertain the following information?** *

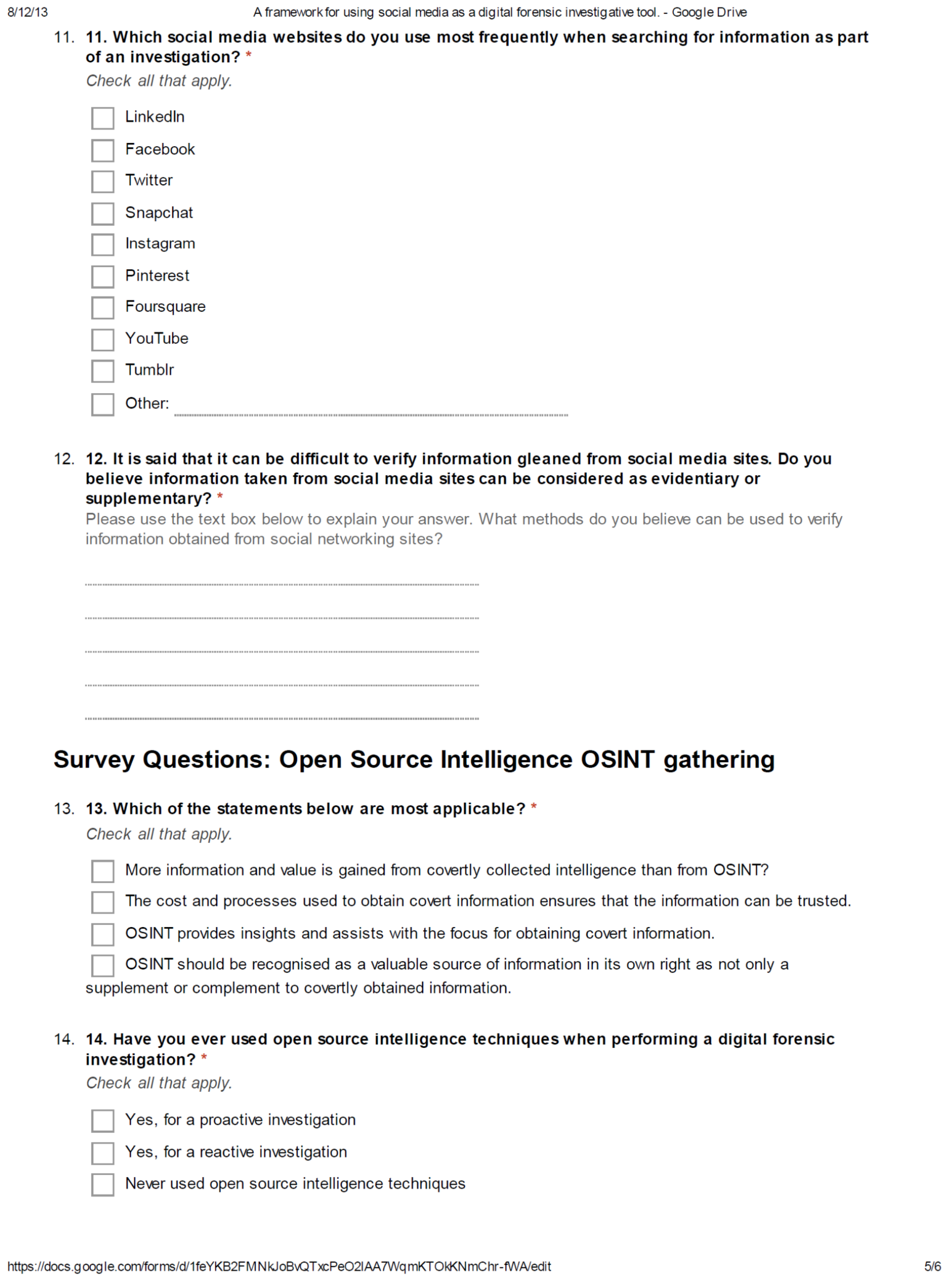
Mark only one oval per row.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Identify associates affiliated with persons of interest	○	○	○	○	○
Identify location of criminal activity	○	○	○	○	○
Gather photos or statements to corroborate evidence	○	○	○	○	○
Identify criminal activity	○	○	○	○	○
Identify persons of interest	○	○	○	○	○
Identify/monitor persons of interest whereabouts	○	○	○	○	○
Solicit tips on crimes	○	○	○	○	○
Anticipate crimes that may be occurring	○	○	○	○	○
Understand criminal networks	○	○	○	○	○
Use information from social media as probable cause for search warrants	○	○	○	○	○
Criminal Investigations	○	○	○	○	○

10. **10. Do you ever make use of any of the following tools when using social media for digital forensic investigations?** *

Check all that apply.

□ Internet Evidence Finder
□ CacheBack
□ NetAnalysis
□ Maltego
□ Bulk Extractor
□ Python scripts
□ Webcrawlers
□ Google hacks
□ Creepy
□ Other: ____________________



11. **11. Which social media websites do you use most frequently when searching for information as part of an investigation? ***

Check all that apply.

- □ LinkedIn
- □ Facebook
- □ Twitter
- □ Snapchat
- □ Instagram
- □ Pinterest
- □ Foursquare
- □ YouTube
- □ Tumblr
- □ Other:

12. **12. It is said that it can be difficult to verify information gleaned from social media sites. Do you believe information taken from social media sites can be considered as evidentiary or supplementary? ***

Please use the text box below to explain your answer. What methods do you believe can be used to verify information obtained from social networking sites?

Survey Questions: Open Source Intelligence OSINT gathering

13. **13. Which of the statements below are most applicable? ***

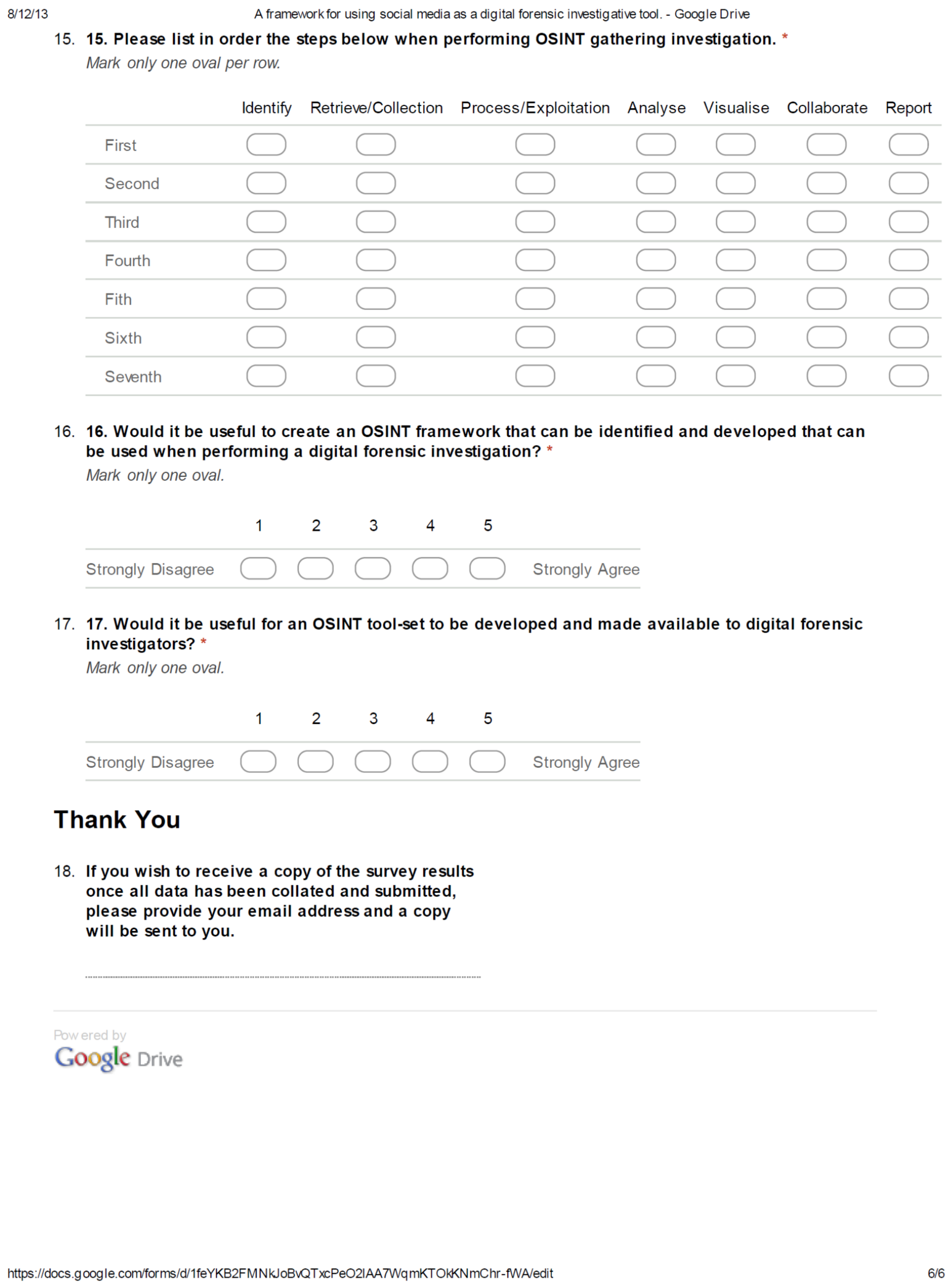
Check all that apply.

- □ More information and value is gained from covertly collected intelligence than from OSINT?
- □ The cost and processes used to obtain covert information ensures that the information can be trusted.
- □ OSINT provides insights and assists with the focus for obtaining covert information.
- □ OSINT should be recognised as a valuable source of information in its own right as not only a supplement or complement to covertly obtained information.

14. **14. Have you ever used open source intelligence techniques when performing a digital forensic investigation? ***

Check all that apply.

- □ Yes, for a proactive investigation
- □ Yes, for a reactive investigation
- □ Never used open source intelligence techniques



15. **15. Please list in order the steps below when performing OSINT gathering investigation.** *

Mark only one oval per row.

	Identify	Retrieve/Collection	Process/Exploitation	Analyse	Visualise	Collaborate	Report
First	○	○	○	○	○	○	○
Second	○	○	○	○	○	○	○
Third	○	○	○	○	○	○	○
Fourth	○	○	○	○	○	○	○
Fith	○	○	○	○	○	○	○
Sixth	○	○	○	○	○	○	○
Seventh	○	○	○	○	○	○	○

16. **16. Would it be useful to create an OSINT framework that can be identified and developed that can be used when performing a digital forensic investigation?** *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

17. **17. Would it be useful for an OSINT tool-set to be developed and made available to digital forensic investigators?** *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Thank You

18. **If you wish to receive a copy of the survey results once all data has been collated and submitted, please provide your email address and a copy will be sent to you.**

....

Powered by

Google Drive

Appendix B: Python Script Twitter-Search-Geo

Python script Twitter-search-geo[28] to scrape data from Twitter for tweets sent close to the Rondebosch Common located in Cape Town. The tweets are then exported to CSV file.

```
#---------------------------------------------------------------------------
# twitter-search-geo
# - performs a search for tweets close to Rondebosch Common, and outputs
# them to a CSV file.
#---------------------------------------------------------------------------

from twitter import *

import sys
import csv

# create twitter API object
twitter = Twitter()

# open a file to write (mode "w"), and create a CSV writer object
csvfile = file("output.csv", "w")
csvwriter = csv.writer(csvfile)

# add headings to our CSV file
row = [ "user", "text", "latitude", "longitude" ]
csvwriter.writerow(row)

# the twitter API only allows us to query up to 100 tweets at a time.
# to search for more, we will break our search up into 10 "pages", each
# of which will include 100 matching tweets.
for pagenum in range(1, 11):

# perform a search based on latitude and longitude
# twitter API docs: https://dev.twitter.com/docs/api/1/get/search
query = twitter.search(q = "", geocode = "-33.952124,18.485330, 1km", rpp = 100, page =
pagenum)
```

[28] https://github.com/ideoforms/python-twitter-examples/blob/master/twitter-search-geo.py

```
        for result in query["results"]:
                # only process a result if it has a geolocation
                if result["geo"]:
                        user = result["from_user"]
                        text = result["text"]
                        text = text.encode('ascii', 'replace')
                        latitude = result["geo"]["coordinates"][0]
                        longitude = result["geo"]["coordinates"][1]

                        # now write this row to our CSV file
                        row = [ user, text, latitude, longitude ]
                        csvwriter.writerow(row)

        # let the user know where we're up to
        print "done page: %d" % (pagenum)

# we're all finished, clean up and go home.
csvfile.close()
```