Personal Information Disclosure and Privacy in Social Networking Sites

MASHAEL AL JOHANI

A thesis submitted to the Faculty of Design and Creative Technologies
Auckland University of Technology
In partial fulfilment of the
requirements for the degree of
Master of Information Security and Digital Forensics

School of Engineering, Computer and Mathematical Sciences

Auckland, New Zealand
2016

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which, to a substantial extent, has been accepted for the qualification of any other degree or diploma of a University or other institution of higher learning, except where due acknowledgment is made in the acknowledgments.

…….....

Mashael Al Johani

(17-October-2016)

Acknowledgements

This thesis was completed at the Faculty of Design and Creative Technologies in the School of Engineering, Computer and Mathematical Sciences at Auckland University of Technology, New Zealand. All praise to Almighty Allah, the most merciful and compassionate, who gave me this opportunity and supplied me with the skills and abilities to complete this thesis successfully.

I must express my gratitude to my parents, who have always been a source of encouragement throughout my life and from the start to the end of this thesis. They were always there for me and I could not have done it without their support. Likewise, I would like to thank my brothers and sister for their limitless support throughout the writing of this thesis.

I would like to thank my supervisor, Dr. Alastair Nisbet, for the patient guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. His commitment and hard work have been a great inspiration for me to complete this thesis successfully. I would also like to thank my second supervisor, Dr. Kelly Blincoe, who gave me valuable advice and many helpful suggestions.

I would like to thank the government of Saudi Arabia for providing me with this great opportunity to pursue my dreams and achieve my goals. In addition, I would like to thank the Ministry of Higher Education in Saudi Arabia, and the Saudi Cultural Mission in New Zealand for their continuous support.

Last but not least, I would like to express my sincere appreciation to my dearest friend, Norah Alomairah, who was always there for me from the start, supporting me, encouraging me, and pushing me to achieve my potential.

Abstract

Social networking sites (SNSs) have revolutionized traditional information-sharing methods. They are mostly built on an individual's offline social circle and they provide users with a wide variety of virtual interaction mechanisms. Such sites and applications have become important communication platforms that are integrated into the daily life of many users. However, they have also blurred the line between users' offline and online lives and created the illusion of familiarity and intimacy over the web, which may have resulted in the sharing of a significant amount of personal information that users might have preferred to remain private. The technology of social networks is a double-edged sword. Although it can provide great benefits to its users, it comes with a huge price and responsibility: users' privacy. SNSs users are not just the source of information; they are also the targets. They leave digital footprints during their visits to those websites and mobile applications, where privacy breaches and identity theft cases are increasing at an alarming rate. Users are vulnerable to privacy breaches from many different entities. They can come from SNSs service providers, third party applications, other users from SNSs users' social networks, or other malicious attackers. However, privacy protection responsibility lies primarily with the individual user and often depends on users' levels of personal information disclosure and knowledge of protection methods. The sharing of personal and identifying details such as gender, age, education, location, address and other personal information such as personal and family photographs can assist in establishing an identity that can be easily stolen and used by criminals. Identity theft criminals exploit the lack of awareness of SNSs users to gather personal information that has been freely supplied by the user.

The purpose of this research is to assess the intensity of this problem by identifying SNSs users' personal information disclosure levels, the kinds of information that they reveal, the degree to which they expose personal information to the public, the privacy settings they apply and their level of knowledge and awareness about how their information is protected by SNSs service providers. In addition, this research studied the effects of gender, age, education, and level of privacy concern on the amount and kind of personal information disclosure and privacy settings applied. Two methods of data collection were used. Firstly and primarily, an online survey was used to collect

information about users' behaviour on SNSs. The secondary method was a social experiment that tested SNSs users' reactions to profile access requests by a stranger. The research focused on four different social networks: Facebook, Twitter, Instagram, and Snapchat.

The survey showed that there was a significant amount of disclosure of personal information; however, it differed from one social network to another. Facebook had the highest level of information disclosure, whereas Twitter had the lowest amount of information disclosure compared to the other networks. The research revealed that gender, age, and education had significant influences on information disclosure and users' privacy settings. In general, males, young people between 16-24, and high school students showed reckless and very identifying behaviour on SNSs that might comprise their privacy and, in the worst cases, their safety, as they become more vulnerable to attacks from identity thieves and other malicious entities.

The findings of the social experiment indicated that the majority of Facebook, Instagram, and Twitter users were likely to accept complete strangers into their personal private profiles. The study concludes by offering recommendations and guidelines that may provide a safer browsing experience for social network users.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

- EXIF Exchangeable image file format
- VPN Virtual Private Networks
- SNSs Social Networking Sites
- PII Personally Identifying Information

Chapter 1: Introduction

1.0 BACKGROUND

The emergence of social networking sites (SNSs) in people's daily lives has transformed the way users communicate and share information. Before SNSs, people's means of communication and information sharing were very limited, especially in terms of interaction, and people mostly communicated with others they knew personally. Currently, individuals use SNSs to share user-generated content online via computers or smartphones in many different formats, depending on the social network of their choice (Ge, Peng, & Chen, 2014). Users share news about their lives effortlessly, whether it is in the form of a video, a photo, a post or a status update. In addition, users nowadays share information with a much larger audience, sometimes larger than they intend.

Advances in technology have enabled SNSs to develop enormously in a way that has created new methods of sharing information. Social networks began as websites where users only had access via a laptop or desktop. However, with the development of smartphones, social networks released mobile application versions and/or developed stand-alone mobile applications. This development made it easier and more convenient for users to access their online profiles, updating them more actively and in real time (Aldhafferi, Watson, & Sajeev, 2013). However, the more accessible the social network and the easier it is to use, the more information users share (Coyle & Vaughn, 2008) due to its constant presence in their lives. SNSs unquestionably have a strong social impact; however, as a result, the lines between individuals' virtual and offline lives have been blurred.

As of August 2016, there were over 2.34 billion social network users globally. This number is expected to increase to 2.95 billion social networks users by 2020, which is approximately a third of the world's entire population (Statista, 2016). Due to the increased use of SNSs, social networks have become rich sources of users' personal information. Users' personal information is very valuable to many different parties and can be exploited for financial gain. Firstly, advertisers can invade users' privacy by accessing their personal information and browsing habits, which is supplied by SNS providers, in order to

recommend products and services; such promotions are referred to as targeted personalised ads. Secondly, sharing of personal information such as full name, age, gender, and other personal information such as family photos leaves users vulnerable to online criminals, who may exploit such information for malicious actions such as identity theft or online stalking. Such actions can affect users' safety and cause not just financial loss but also emotional distress to the victims.

The aim of this research is to shed light on SNS users' personal information disclosure behaviours, their privacy protection settings, privacy policies, and SNS users' privacy knowledge and awareness. It examines the effect of gender, education status, and age on the degree of personal information disclosure and protective privacy settings applied by the user using factor analysis. An online survey was designed in order to answer the following research questions. In addition, an experiment was conducted to test how users react to "friend requests" from strangers. The following are the three main research questions this research aims to answer.

Q1: What are the personal attributes that can have an influence on information disclosure by and the privacy settings of SNS users?

Q2: How do users' levels of privacy concern affect the amount of information they disclose in social networking sites?

Q3: How aware are users of the extent to which their information is protected by SNS providers according to the privacy policies that the users have agreed to?

1.1 MOTIVATION

Section 1.0 briefly discussed the background to this research in order to understand the importance of the chosen research area. This section discusses the motivation for investigating social network users' personal information disclosure and privacy-related issues. The research was motivated by reading previous research published in the area of social network privacy and can be summarised in four main points, described in the following four paragraphs.

Firstly, the importance of researching the privacy of users' information on SNSs is increasing due to the rising popularity of SNSs. In addition, new social networking applications are emerging, providing more sources through which users' personal information can be accessed. One of the reasons why SNSs are becoming part of users' lives is the fact that almost all social networks have a mobile application version or are stand-alone mobile applications, which eases the process of accessing social networks and sharing daily updates. It also eases the process of tracking users' behaviour.

Secondly, this research aims to investigate the type of information users provide in social networks. This study focuses on four social networks: Facebook, Twitter, Instagram and Snapchat. After analysing existing research on social network privacy, it became apparent that there is a lack of research in this area. Some research has focused on just one social network, most commonly Facebook. This research comprehensively analyses not just one, but four popular social networks. Snapchat is the newest social network and is one of the fastest-growing social networks so far, with 100 million active daily users (Statista, 2016). However, very little research has been conducted on Snapchat and its users. Therefore, this research provides significant academic information on the type of personal information disclosed in all four social networks, and the degree of self-disclosure to the public.

Thirdly, an important aspect of this research is to study the way users' personal information is handled by SNS providers according to their privacy policies. The privacy policy statement is one of the most important legal documents on any website as it provides detailed information on how the website uses and protects information collected from visitors. These privacy policies hold important information regarding the privacy of users' personal information, what ownership they have over their data and what actions social network providers perform on users' information. Therefore, this research intensively investigates what those privacy policies indicate regarding users' personal information. The analysis of privacy policies is reported in Chapter 2. In addition, this research measures SNS users' knowledge and awareness of how their information is collected, stored, shared, and displayed.

Lastly, understanding the type of information users disclose, and their awareness levels, assists with initiating protective recommendations and guidelines that may enhance users' security, protect their privacy and most importantly increase their awareness, which is one of the ultimate goals of this research. Therefore, another motivation for conducting this research is to suggest privacy and security guidelines for social network users that can protect them from online crimes such as identity theft or from privacy breaches caused by SNS service providers and third parties.

1.2 STRUCTURE OF THE THESIS

This thesis is composed of six chapters: 1. Introduction; 2. Research Literature Review; 3. Research Methodology; 4. Research Findings and Analysis; 5. Discussion of Results; and 6. Conclusion.

Chapter 1 is an introductory chapter that firstly presents background information on SNSs and privacy concerns in this field, and states the importance of this research topic and the motivation for the research. It then presents the research questions and the approach adopted to undertake this research.

Chapter 2 presents the literature review, in three main sections. Firstly, it presents background information on what social networks really are, with some brief history to show how user information disclosure has developed over the years to reach the state that we have today. This is followed by an analysis of the characteristics of the four chosen social networks for this research: Facebook, Snapchat, Instagram, and Twitter. Each of these four social networks is analysed in terms of what personal information the user displays and the default privacy settings for each attribute. The second part of this chapter analyses privacy issues specifically in SNSs. This section discusses the value of privacy and some privacy issues that can result from using SNSs. Lastly, the chapter analyses social network privacy policies and what they contain in terms of how they handle users' data and how they share it.

Chapter 3 establishes the research methodology that is applied in this research. Three similar approaches from three previous studies in the chosen research field are evaluated as

guidance for developing the methodology for the proposed research. Each of these studies provides helpful insights on the topic of SNSs' information disclosure and privacy concerns. Chapter 3 also identifies the main research questions, sub-questions, and data requirements.

Chapter 4 reports the research findings. The main results are presented in tabular and graphic formats. This presentation is followed by statistical analysis of the data in order to understand the relationships and trends in the results.

Chapter 5 discusses the key findings from the analysis of the results presented in Chapter 4 and answers the main research questions and sub-questions presented in Chapter 3. Chapter 5 provides a critical reflection on the results and recommendations for protecting user privacy.

Chapter 6 summarises the research findings. It concludes this research by providing recommendations for future research in the area of social network privacy and information disclosure. This chapter also presents the limitations of the present study.

Lastly, the appendices are attached at the end of this document as supplementary information. They include the ethical approval document, survey information sheet, and survey questions.

Chapter 2: Literature Review

2.0 INTRODUCTION

Technological advancement has become less focused on connecting computers and more concerned about connecting people. A main contributor to this evolution is the use of social networking sites (SNSs), which have seen explosive growth in use in the last couple of years (Zheleva, Terzi, & Getoor, 2012). As of August 2016, there are more than 2.34 billion users of SNSs (Statista, 2016). Due to the increasing popularity of SNSs and the drive to reach customers, more than 70% of businesses are now using SNSs (McKinsey Global Institute, 2012). Although SNSs provide a powerful tool to engage people over the web, they can be a source of possible threats to users' privacy and security, because users routinely and voluntarily provide personal information (Cross, 2014).

In order to provide the reader with an understanding of how users began engaging and sharing personal information about themselves online, this chapter commences in section 2.1 with a brief background on how SNSs have advanced in the past decade. Section 2.2 discusses what kind of sensitive user attributes can be disclosed in different SNSs. The chapter will then cover two main areas of focus:

The first selected area for this literature review is privacy in SNSs as discussed in section 2.3. This section will critically review and evaluate the meaning and value of Internet privacy to users and will help to identify the main privacy concerns that users may encounter while using SNSs.

The second area of focus is an in-depth analysis of the privacy of four popular social networks: Facebook, Instagram, Twitter, and Snapchat, each of which has its own unique attributes. This section will provide an analysis of their privacy policies, which contain valuable but often neglected information about how users' data is handled and whether or not it is shared with other parties.

2.1 SOCIAL NETWORKING SITES BACKRGOUND

Mark Zuckerberg, the founder of Facebook, states,

> *"I wanted to create an environment where people could share whatever information they wanted, but also have control over whom they shared that information with"* (Zuckerberg, 2006, para. 3)

SNSs are a representation of a virtual community where users are encouraged to connect, communicate and engage with other users in the network. Users of SNSs can engage with others by posting personal information about themselves, sharing information and news, uploading videos and images, and/or having instant and real-time conversations with others in the same network by using the chat functionality (Shin, 2010). According to Boyd and Ellison (2007), SNSs are defined as web-based services and applications that enable users to create a public, semi-public, or private profile within a bounded system. The users have the ability to create a list of online friends that contains other users in the same network with whom they share common interests or connections. Each SNS has a different purpose; hence, the nature and nomenclature can vary depending on the site (Boyd & Ellison, 2007). The uniqueness of SNSs revolves around the fact that they give the user the freedom of not only communicating in a network with new people, but also getting in touch with people from their offline social network, using the Internet.

2.1.1 The history and expansion of Social networks

SNSs have evolved over the years and have gone through many phases of development to reach their current state (Hendricks, 2013). According to the above definitions of SNSs, the first recognizable form of SNSs that encouraged users to include personal information for the purpose of social networking emerged in 1997 with a site called SixDegrees (Boyd & Ellison, 2007). It allowed users to open personal accounts and create a list of friends. SixDegrees attracted over a million subscribers at its peak (Chapman, 2009). However, although SixDegrees managed to become popular and obtain a large number of subscribers, the site was not able to maintain its popularity (Boyd & Ellison, 2007). In 2001, SixDegrees.com was shutdown. According to the founder of SixDegrees, the failure of his site was due to the fact that SixDegrees was ahead of its time: at that time, not many people

had friends who were online and the idea of being online-friends with strangers had not yet gained universal acceptance (Prall, 2010).

Although SixDegrees was shut down, the concept of creating virtual social networks inspired other developers (Prall, 2010). In the early 2000s, more people started to have Internet access; hence, the target audience was much broader than it used to be. This helped the success and increased the popularity of SNSs such as Friendster, which has attracted more than 90 million users. It introduced the ability for users to find friends and then friends-of-friends, and thus expand their networks and share more information with others.

The vast spread of SNSs started to occur at the start of 2003, initially when Myspace was launched, which grew to be the most popular SNS in the world at that time (Boyd & Ellison, 2007). Myspace differentiated itself from other competitors by giving users the freedom to customize the look of their profiles. In 2004, Facebook was launched initially as a Harvard-only social network and became the most popular SNS in 2008, overtaking Myspace. As of the second quarter of 2015, Facebook had 1.49 billion monthly active users (Statista, 2016). Facebook manages to maintain its success by constantly improving the site and by adding new features (Hendricks, 2013). At the present time, hundreds of SNSs have emerged, each designed to serve a different audience or have a different style that distinguishes it from other SNSs. Figure 2.1 illustrates the vast growth of SNSs from 2006 to 2012.

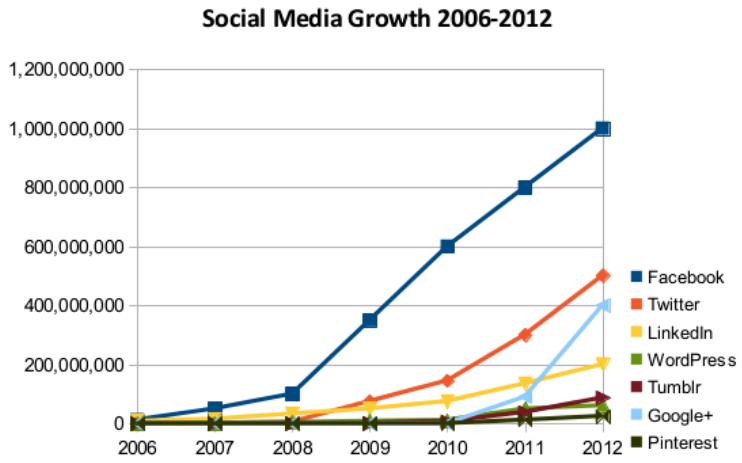

Figure 2. 1. Growth of Online Social Networks, 2006-2012. (Source: White, 2013)

2.1.2 Characteristics of social networks

SNSs can be distinguished from other web-based applications by five core characteristics. Table 2.1 summarises the five core SNSs features according to a research study jointly conducted at Rice University, the University of Maryland, and The Max Planck Institute for Software Systems (Mislove, Marcon, Gummadi, Druschel, & Bhattacharjee, 2007).

Table 2. 1. Characteristics of SNSs

User-based	➢ Users provide the content ➢ The flow of information can be controlled by anyone – not determined by the web-master. ➢ Freeform/unstructured
Interactive	➢ Not just a website with chat rooms or forums ➢ Users have many options: comment on others posts/videos/pictures; play online games provided by the social network (mostly by third party application authorized by SNSs); take quizzes; share the user's own photos, videos, thoughts and ideas with friends.
Community-driven	➢ Members hold common interests or connections ➢ Enables users to be in touch with old friends
Relationships	➢ SNSs are based on the number of friend/follower relationships between their members ➢ If privacy controls are not properly set up on the account, private information may be dispersed to your friends, their friends, and other unknown users.
Emotion over content	➢ Before SNSs, website content was mainly focused on delivering information to visitors. However, with SNSs, people are able to communicate their needs within a community of easy-to-reach friends, which gives SNSs an emotional factor.

2.2 TYPES OF SOCIAL NETWORKING SITES

SNSs can be differentiated based on their purpose and functionalities. For example, Twitter is an appropriate site if the user is seeking to send a short message to a large audience, but it is impractical for trying to communicate with other users if the message is larger than 140 characters. Similarly, if a user wishes to post a video message instead of text, the most

appropriate tool is YouTube. The focus of this research is on four SNSs: Facebook, Twitter, Instagram, and Snapchat. Each of these SNSs has a different purpose, different functionalities, and provides the user with different experiences. Therefore, this section will describe the concept behind each SNSs and what type of personal information it contains and displays. The following information was retrieved from each SNS in May 2016.

2.2.1 Facebook

The concept behind Facebook is that a user can re-create his/her offline social network online by creating a profile that contains personal information. Facebook has many features and functions. Table 2.2 lists the main features and states their default privacy settings. Default settings refer to the settings that are automatically set on the feature. The user needs to change the setting manually in order to achieve a more restrictive view, and vice versa.

Table 2. 2. Facebook features and default privacy settings

Feature	Default privacy setting
Profile picture/ Profile Background picture.	Public
Work and Education: Workplace, professional skills, college, high school.	Public
Places the user has lived in including current city and hometown.	Public
Contact and basic information: email, mobile phone, address, birthdate.	Friends
Other information: gender, language, religious views, political views.	Public
Family members' names and Facebook profiles, relationship status.	Public
Posting thoughts, videos, or pictures to timeline.	Friends
Tagged pictures.	Friends and friends of anyone tagged in the picture

2.2.2 Snapchat

Snapchat is a picture/video messaging mobile application that has gained great popularity since its launch in 2011, with over 100 million daily active users (Tweney, 2016). Snapchat's main functionality is that it allows the user to send to other users an image or a video that can only last up to 10 seconds. The sent and viewed image/video cannot be opened again and it is automatically deleted without the ability to save it directly from the

application. In addition, a “Snapchat Story” is a feature that allows the user to post multiple 10-second videos/pictures that are available for 24 hours. The default setting is friends. However, the user can make it public or restrict it to a customised view for specific friends. In terms of user-generated content, Snapchat is limited to these two main functions.

2.2.3 Twitter

Twitter allows the user to post “tweets”, which are messages that cannot exceed 140 characters. The user profile is public by default but the user can limit their profile and make it private. Twitter users can add the information shown in Table 2.3 to their profiles and this information is public even if the user’s tweets are private:

Table 2. 3 User Profile options and default privacy settings in Twitter

Feature	Default Privacy setting
Username and profile name	Public
Photo (profile and background)	Public
Bio (the user can type any brief information about themselves)	Public
Location (which country or city the user lives in)	Public
Website (if the user has one)	Public

2.2.4 Instagram

Instagram has very limited functionality compared to Facebook. Users do not provide specific personal information as they are able to do on Facebook. Instagram is designed to act as a huge collaborative virtual photo album. The main functionality for users is posting photos, and videos up to 1 minute in length. When a user posts a picture or a video, they can share the exact location of the photo with their followers by adding the location on a map before uploading the image. Users can choose whether they want to make their profiles public or private. Another feature that was introduced to Instagram is direct messages, allowing a user who follows another user to send a private message. The receiver of the message doesn’t have to be a follower of the sender in order to receive the message.

2.3 PRIVACY AND SECURITY IN SNSs

The last decade has witnessed a rapid growth in the number of individuals using SNSs. For instance, as of June 2016, Facebook was regarded as the third most used website globally after Google and YouTube (Alexa, 2016). Although SNSs provide many benefits for individuals such as keeping in touch with friends and family, privacy and security is regarded as a critical issue that can threaten the users of SNSs (Donath, 2007). This is mainly because SNSs encourage their users to reveal a great deal of personal information about themselves by promising them a better user experience if they do so (Luo, Liu, Liu, & Fan, 2009). When users first sign up to Facebook, they will be constantly asked and reminded by Facebook to update their profile with more personal information such as date of birth, hometown, workplace, and/or school in order to find more friends and enjoy the experience more (Lewis, 2015). The growing popularity of SNSs and the fact that they contain enormous amounts of information make these websites an attractive target for malicious hackers. Threats from SNSs can be divided into two different categories; security and privacy. The following section will differentiate between the two types and will discuss the related challenges that concern the users of SNSs; however, the main focus in the following sections will be on privacy-related matters in SNSs.

2.3.1 Privacy vs. Security

The terms 'privacy' and 'security' sometimes overlap and may be used interchangeably by users and researchers. Therefore, in order to provide a clearer conceptualisation, the key terms, privacy and security, will be defined as follows with respect to SNSs:

- Security: In SNSs, security threats result from the technical vulnerabilities of the network (Altshuler, Elovici, Cremers, Aharony, & Pentland, 2013). In 2009, the Secure Enterprise 2.0 Forum identified and listed eight main security threats that may occur when using social networks (Chi, 2011): insufficient authentication controls; cross-site scripting; cross-site request forgery; phishing; information leakage; injection flaws; information integrity; and insufficient anti-automation.

- Privacy: In general, privacy deals with the extent of control over the flow of users' personal information including access to, transfer, storage, management and exchange of that information (Altshuler et al., 2013). When users post personal information, they have the option to control who can see the posted content. However, their privacy is violated when other parties collect their information and use it without their consent, which indicates that the user did not actually have control over their information. Privacy breaches in SNSs can come from four different sources, and are analysed with respect to who can invade a user's privacy:

    1. **SNSs administrators/service providers:** This includes the developers of the SNSs. They store the users' information on their servers and have access to it. Based on the privacy policy that the user has agreed to, they might have ownership and usage rights over the user's data. SNSs' current client–server architecture inherently means that users have to trust SNS providers to protect all the private personal information they've uploaded to their accounts as they do not have any other choice but to trust them. However, SNS service providers can clearly gain many benefits from collecting, examining and sharing users' personal information — for advertising purposes, for instance. Because SNS service providers have the power and the legal right (according to the privacy policy) to use such information however they wish, researchers and privacy advocates have raised serious privacy violation concerns and have attempted to fix this imbalance in power (Gao, Hu, Huang, Wang, & Chen, 2011). Scholars have suggested various alternative SNS architectures as defences. These proposals advocate that SNS users should dictate the fine-grained privacy policies concerning who can access their information. To enforce this proposal of a user-defined policy, the SNSs must store the information with encryption, so that no party, not even the SNS service provider, can access or view the information unless the owner (the user) grants them permission to do so. For instance, this may involve using decentralized storage so that SNS users have the choice of where in the network their information gets stored (Baden, Bender, Spring, Bhattacharjee, & Starin, 2009). However, this has not yet been applied by any of the most commonly used SNSs.

2. **Internet Service providers (ISPs):** Some countries require ISPs to collect data from their users for censorship reasons. For example, this may include what information a user views in SNSs. This information can be collected by governments for purposes such as detecting terrorist activities. However, ISPs can also sell the information they collect to advertising companies. The CEO of NebuAd, an American online advertisement company, stated that his ads are customised to users based on data gathered and bought from sources including ISPs (Narayanan & Shmatikov, 2009).

3. **Third party applications:** Many SNSs offer third-party application services access to their website. For instance, these may include games and quizzes that provide the site with extra functionalities for the user (Cutillo, 2014). These applications are written by third-party developers and usually have minimal security standards (Cheng, Park, & Sandhu, 2013). In some cases, these third-party applications are run by completely separate entities, and they have access to users' information (Cheng et al., 2013). The collected data can then be used for different purposes such as sending tailored and user-targeted advertisements, sending spam to users' contacts, or performing market research studies without the user being aware of it, which violates their privacy. Furthermore, users have to grant these third party applications access to their personal information before they can use them, since such access is required for some applications to execute their functionality. For instance, a horoscope third party application on Facebook must know the user's birthday. Unfortunately, neither the SNS's service provider nor the users are aware of exactly which piece of data is actually needed for the applications. As a result, SNS users have no other choice but to trust the applications to correctly collect just the information they need. Additionally, the mechanism for monitoring how these third party applications manipulate users' personal information is missing. This allows these applications to potentially misuse users' personal information. For example, a famous Facebook application called "Compare Friends" promised to protect users' privacy when they used the application and expressed opinions about their

friends; however, it was later discovered that Facebook offered to sell that information to other users (Singh, Bhola, & Lee, 2009).

4. **Other users in the network:** Other users can be divided into two types:
   1. Trusted users: anyone who is given access to the user's information by the user e.g. followers or friends.
   2. Malicious attackers: anyone who tries to get access to a user's personal profile either by deceiving the user or by launching attacks against the user to gain access.

However, privacy and security in SNSs can overlap because a security breach can result in a privacy violation (Beye, Jeckmans, Erkin, Hartel, Lagendijk, & Tang, 2010).

2.3.2 Privacy

The concept of privacy has attracted enormous attention from academic researchers as well as from the general press (Patil & Kobsa, 2009). Figure 2.2 shows the huge increase in the number of articles and non-fiction books that contain the word "Privacy" in their title, especially from the mid-1990s onwards, which is aligned with the advent of the Internet and later on, with the introduction of SNSs into people's lives.

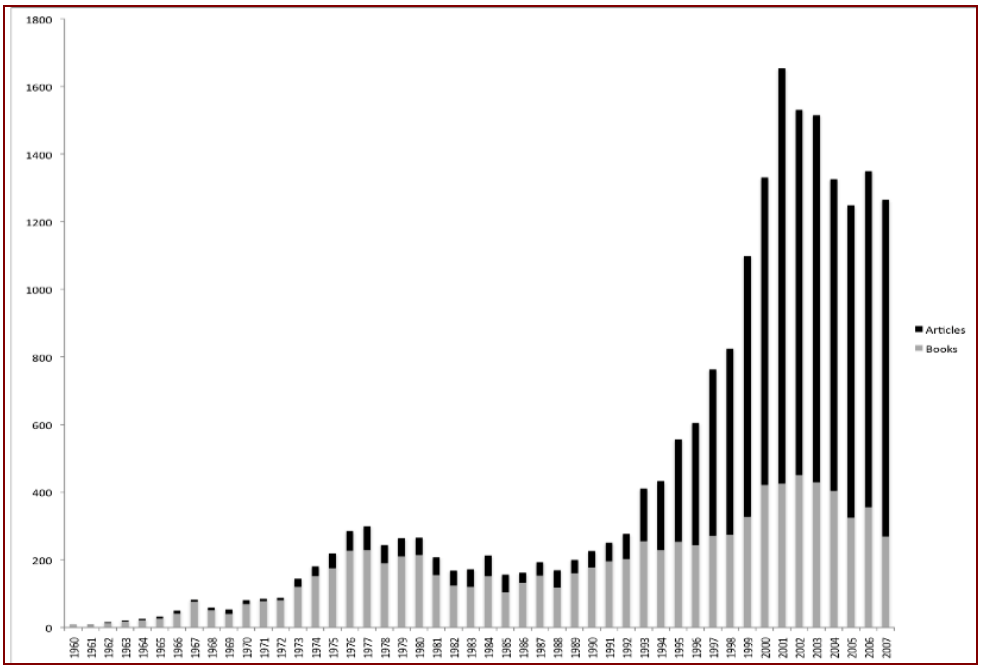

Figure 2. *2* WorldCat illustration of non-fiction books and articles with "privacy" in their titles, 1960-2007 (Source: Patil & Kobsa, 2009)

Privacy has no single clear and straightforward definition (Patil & Kobsa, 2009). Privacy is defined based on the context it is used in; for instance, during the era of non-electronic newspapers, magazines, and photographs, privacy was perceived, according to Warren and Brandeis (1890) as "the right to be left alone." When journalists write about individuals' stories or lives in newspapers and publish photos of them without their consent, this can be considered an invasion of privacy because those people have the right to be left alone and not have their pictures posted in public, especially without their consent (Cutillo, Molva & Onen, 2011).

However, new forms of communication have emerged with the development of technology over the years, giving people more options to explore and exchange information. Communication sources such as SNSs have personal information as their main raw material, which is mostly provided by the users about themselves. Therefore, the definition of privacy has a slightly different meaning in this context, because users usually have a choice about whether or not to upload information about themselves. Privacy in SNSs refers to the ability of a person to have control over how their personal information is obtained, managed, processed, and shared by and to any other entities (Cutillo, Molva & Onen, 2011). In the data science field, Patil and Kobsa (2009) have presented three main ways in which to define and analyse privacy, as shown in Table 2.4.

Table 2. 4 Definition and analysis of privacy with regard to data

Perspective	Concept of privacy	Example	Consequences of privacy violation
Legal (Normative)	The legal aspect views privacy as the ethical and legal right for a person to be left alone. This perspective emphasises protecting individual personal data from governments, corporations, and even other individuals based on existing laws and policies	The European data protection framework	Offenders can face civil and/or criminal penalties.
Social	Privacy aspect that is socially constructed. The social aspect focuses on the interaction that occurs between individuals on a daily basis.	An employee posts personal content that might affect their professionalism if it reached their boss or the clients.	Possible social embarrassment or reputational damage that can affect social and professional life.
Technical	The technical aspect views privacy with regard to the functional characteristics of computerized systems. Privacy is perceived as the need for selective and acceptable control over shared incoming and outgoing data and information.	When a user signs up to a social networking website, they should be able to have appropriate, manageable privacy settings, which give them the ability to control who can view or obtain their data.	Unauthorized data access and collection, Identity theft or fraud, illegal information usage

2.3.3 Privacy concerns related to SNSs:

It has been argued that advances in communication technology have made people more tolerant and more willing to share information about themselves in a way that renounces the value of privacy in order to be more connected and traceable; this applies particularly among younger generations (Tubaro, Casilli & Sarabi, 2014). People now willingly share specific information about their daily life routines, including posting their location to friends or even to complete strangers. With SNSs' intrusions into people's lives, a new hypothesis has emerged claiming the "end of privacy", which is a controversial discussion that needs to be taken seriously (Tubaro et al., 2014). This signifies a huge transformation in our societies' systems of values and behaviour and has changed our political, cultural and regulatory existence. The effect is not only at a social level; it also affects how

businesses create and comply with privacy policies and laws and how they manage business relationships with their stakeholders.

When using SNSs, users choose to share information about themselves by uploading images and posting; however, users might not be aware that they are also revealing other information about themselves. Barnes (2006) indicated that many observers have noted a tendency for SNS users to abandon their privacy in favour of having their digital and online footprint open and traceable. However, what are the consequences of using this technology? The next section will discuss the privacy-related issues that users might face when using SNSs.

2.3.3.4.1 Social networks data mining

There are two methods for collecting data from SNSs, either via a crawling API, or by screen-scraping (Narayanan & Shmatikov, 2009). Mislove et al. (2007) crawled over 11.3 million online profiles in order to examine the structures of multiple SNSs for academic purposes. Most SNSs do not now publish their APIs for security reasons, but screen-scraping is possible using appropriate sources. Web-scraping is prohibited by the European Court of Justice (CJEU) as it is considered to be an intellectual property infringement. In the US, website operators can sue screen-scrapers under copyright laws and trespass to chattels claims (Mclean & Samavi, 2015). Screen-scraping, however, still occurs using illegal technical methods and can result in two privacy and security violations:

1. If malicious hackers are able to screen-scrape SNSs, they can obtain a huge number of names and contact details; hence, they can perform large-scale attacks such as spamming or phishing attacks.

2. Research companies harvest users' data without their consent or the website operator's consent. With the appropriate tools, they can collect millions of users' online information and sell it to other companies, such as marketing companies, which can design targeted ads for their products. It can be argued that this information is already public and the users put this information in the public domain; hence, it is free for anyone to use. However, screen-scraping companies do not ask for users' consent to use their information, which results in

ethical issues and privacy violation. In 2010, Nielsen Co., a media-research company, used highly sophisticated web-scraping software to collect messages that were exchanged between users of a website called PatientsLikeMe (Angwin & Stecklow, 2010). In this network, users post and exchange with others highly personal stories about their physical and mental health, such as their desire to hurt themselves. Nielsen Co. registered on the website as a member and screen-scraped numerous messages that contained people's emotional problems. The company then sold this information (Angwin & Stecklow, 2010). The users of the site believed that it was a safe environment in which to share their suffering but their privacy was violated and their information was collected without their consent.

2.3.3.4.2 Linking users across different domains with location data

Digital traces are easily created during every interaction with technology, from cell towers connecting and routing cell phone calls to businesses recording credit card transactions for purchases; from pictures we take using our phones, to daily status updates on SNSs. The fact that users' digital traces can be merged and connected together across different domains and datasets is fascinating but also unsettling. An example is seen in applications that gather information about user behaviour from multi-domain datasets in order to provide recommendations (Riederer, Kim, Chaintreau, Korula, & Lattanzi, 2016). However, fundamental questions are raised about data privacy when this is done by third parties that the user doesn't have direct interaction with. The majority of smartphone apps and SNSs such as Facebook, Twitter, Snapchat and Twitter capture and record their users' precise geographical locations (Riederer, Zimmeck, Phanord, Chaintreau, & Bellovin, 2016). This substantially increases the number of parties who can collect information about users' whereabouts and exploit that knowledge (Riederer et al., 2016). Although the collected data may be recorded sporadically, these datasets are connected with the user's daily life routine, which makes them very rich and intimate, and they can reflect recognizable patterns of a person's life (Riederer et al., 2016). A recently published study has confirmed that even having a minor amount of location data is sufficient to determine the user's online identity (De Montjoye, Hidalgo, Verleysen, & Blondel, 2013). In addition, it can be used to infer several sociological attributes, including race (Riederer et al., 2016)

friendship, gender, or social status, when aggregated with domain semantic information (Crandall, Backstrom, Cosley, Suri, Huttenlocher, & Kleinberg, 2010)

2.3.3.4.3 Uploaded images' EXIF data

Exif stands for Exchangeable image file format, and it is a standard that states the format and specification for digital images, including smart phone images. When an image is taken, metadata about the image will be stored in the Exif tag of the image. The following is a list of what metadata is stored, according to Kumar, Srikanth, and Sailaja (2016):

- Information about the exact date and time that the picture was taken
- Camera or device settings including model and serial number
- GPS information (altitude, latitude, longitude)
- Creator information if applicable

Most smartphones and cameras save Exif data by default and without the user's knowledge (unless it has been disabled by the user). This can raise privacy and security concerns when these images are uploaded to SNSs, because these networks can be accessed by persons who may be able to extract the Exif data and use it for malicious purposes that may harm the user because the information can reveal the user's physical location. This may be the equivalent of the user publishing their home address online and making it available to the public (Warner, 2014)

IPTC (International Press Telecommunication Council) has published a study of which SNSs retain images' Exif data by testing and checking whether the SNSs display embedded data values. Table 2.5 illustrates which of the most common SNSs strip off metadata and which do not, according to the IPTC Test image (IPTC, 2015)

Table 2. 5 Metadata removal in a range of popular SNSs

SNSs	Exif Display	Save as Exif
Facebook	No embedded metadata shown.	All Exif data is removed.
Twitter	No embedded metadata shown.	All Exif data is removed.
Instagram	No embedded metadata shown.	All Exif data is removed
LinkedIn	No embedded metadata shown.	Embedded Exif fields are preserved in SaveAs files.
Pinterest	No embedded metadata shown.	Embedded metadata preserved in high resolution/original size images, but IIM and XMP metadata is stripped off in downscaled images.
Tumblr	No embedded metadata shown.	Embedded Exif fields are preserved in the SaveAs image files

2.3.3.4.4 Identity theft in social networks

Identity theft is defined as the "misuse of a another person's identity, such as name, social security number, driver's license details, credit card numbers, and bank account numbers" (Denning, 1999, 241). A victim's personal information can be used for both financial gain and to physically misrepresent the victim to people such as law enforcement officials, employers or medical providers.

SNSs' own exposing nature has made them a rich source for identity theft attacks because users often voluntarily place their personal information online. According to the 2014 annual identity fraud report conducted by Javelin Strategy & Research, in 2013, a vast SNS fraud scheme attack was conducted in which attackers cloned users' profiles by stealing their identity and then sending friend requests to other users. Once the request was accepted, the attackers would obtain all the users' private information and send phishing emails to their contacts. The phishing emails contained banking malware. Once opened, the

malware would gather credit card and bank account details from infected computers or smart phones. The devices would then become part of a network of infected devices known as the "Butterfly Botnet". This attack resulted in $850 million in fraud losses. Although some of those responsible for this fraud have been arrested, this stands as an example of the dangers of exposing too much personal information or blindly accepting a friend request thinking that it won't do any harm, when in fact it could greatly affect the security of a person's identity and those of their legitimate contacts (Pascual, 2014).

Recently, Facebook users have been warned about a clever new fraud, in which an attacker steals one person's digital identity and then sends out a desperate request for cash to his or her friends. The scenario of the scam can be similar to the following: "I'm overseas and all of my money and documents have been stolen. Please send me $500 so I can get home and I'll pay you back as soon I get home." Because the profile looks exactly like their friend's profile, concerned and caring friends may be tricked into sending money to the attacker, believing they are helping a friend in need (He, Chen, Su, & Sun, 2014).

This is just an example of how identity theft can be performed in SNSs in a traditional way. These types of traditional attacks, however, are easy to distinguish and identify since the targeted users can see that there are no common grounds such as common friends (Bilge, Strufe, Balzarotti, & Kirda, 2009). However, according to Bilge et al., (2009), identity theft attacks can be more aggressive and are classified into two types: profile cloning attacks and cross-site profile cloning attacks.

- **The profile cloning attack:**

  In this attack, the attacker creates a fake profile in the same social network, which has the victims' name and picture, and sends friend requests to the victim's friends. Names are not unique on social networks, and people may exist who have identical names. Once the attacker's friend request is accepted, he/she will be able to rebuild the victim's online friend's network in order to make the fake identity more believable to others.

- **Cross-site profile cloning attack:**

In this attack, the attacker aims to copy the user's identity from one social network to another social network that the user doesn't have an account with. From there, the attacker can attract the victim's friends from the other social network.

2.3.3.4.5 Factors contributing to identity theft in social networks

According to the Identity Theft Resource Centre (2016), there are seven factors that can facilitate identity theft in SNSs

1. Using low privacy or no privacy settings: Many SNSs' defaults are set to public or have many public features. Not properly setting up SNS privacy can expose the user's personal information to outsiders and attackers (Zheleva et al., 2012)

2. Accepting friend requests from unfamiliar users: SNSs are based on self-selected networks of friends, family, or colleagues. Users will populate their profile with personal and private information assuming that their audience is a trusted audience. As discussed previously, attackers can easily gain access to private information by requesting access to users' private profiles. If the user does not establish that they know the contact personally, a complete stranger can manage to see their personal information.

3. Downloading or using free applications for use on your SNSs profile: Facebook and other SNSs allow third parties to list their applications in the user timeline. Third party applications can be in the form of quizzes, games, or any other kind of application. However, once the user uses these apps, he/she will also give them consent to access his/her private profile.

4. Giving SNS account password details to other people

5. Participating in random untrusted online quizzes from third parties in the network, which may have malicious intentions to acquire personal information by asking the user to divulge personal information.

6. Clicking on URLs that can lead the user to other websites, even if a friend has sent the link. These links may be malicious and can hack and steal the user's profile information.

7. Falling for email scams/phishing that send links to the user so that they can update their SNS profiles. Cybercriminals can create websites that have the same user interface as an SNS's introductory page so that the user can enter their login details.

The consequences of identity theft can be both financially and socially costly, not only for the person whose identity was stolen, but for others who might be deceived by the stolen identity. Attackers can steal a person's identity from social media, including their photos and personal information, to deceive someone else. An example of this is the case of a woman who was a victim to an online scam on Facebook, which cost her NZD $41,000 (Fisher, 2015). The scammer stole another person's identity and pictures, then created an account on Facebook to launch his attack (Fisher, 2015). The attacker contacted the woman via Facebook and started a romantic relationship with her, which lasted 18 months, during which he sent her pictures of the person, whose identity was stolen, claiming it was him. He then told her he was injured and needed money for hospital bills otherwise he would die (Fisher, 2015). This is an example of how scammers can steal information from users' social network accounts if they lack proper privacy settings, and use it to launch fraud attacks.

2.4 ANALYSIS OF FACEBOOK, SNAPCHAT, INSTAGRAM AND TWITTER PRIVACY POLICIES AND PRIVACY VIOLATIONS

This section will review and discuss Facebook, Snapchat, Twitter and Instagram privacy policies. All information regarding privacy policies was retrieved from each SNS's latest published privacy policies on their website/application in May-June 2016. This section will also discuss each network's general privacy settings and related privacy and security concerns.

Firstly, a privacy policy is prepared by individuals or organisations to provide users with information about how their data will be collected, stored, used and shared (Talib, Ismail, Olowolayemo, Naser, Haron, & Yusof, 2016). It may contain any relevant laws that are applicable to the industry. The terms and conditions, however, state what the user must agree to if they want to sign up or use the website. The privacy policy is included as one of the items that the user is agreeing to when they agree to the terms and conditions. Any user

of SNSs is required to agree to the terms and condition of the SNSs they wish to join. Most SNSs have a mandatory checkbox or a button which states that the user agrees to those terms and conditions (Talib et al., 2016). The terms of use contain important information regarding how a user's data is used and viewed and it represents a legal contract between the user and the SNSs. Many people blindly agree to SNSs' privacy policies, usually because it takes a lot of time to read them and the legal language used makes them difficult to understand (White, 2013). According to a study conducted by Carnegie Mellon researchers, the time consumed in reading all the data policies the average American has agreed to would be 76 working days (Madrigal, 2012). The following section will analyse 4 SNSs' terms and conditions specifically in relation to user privacy

2.4.1 Facebook

2.4.1.1 Facebook's privacy analysis

According to Felix (2012), Facebook tracks all other web activities while the user is logged into an active Facebook session. To support her claim, she used a diagnostic tool called "Abine DNT+" to determine that Facebook has over 200 "trackers" that track the web activity of the user. The definition of a tracker, according to Abine, is a request made by a page in a website in an attempt to make the user's browser perform, which will gather data that is intended to collect, record, profile, or share the user's online activity (Felix, 2012). These trackers can be cookies, 1-pixel beacons, Iframes or Javascripts. Facebook uses cookies, which capture a background of user behaviour and interests, for advertising purposes. This act is called spying by critics, but targeting by advertisers. Abine privacy analyst Sarah Downey stated that not only do these trackers invade users' privacy, but they also consume a large amount of data and time to process and transfer (Felix, 2012). As a result, browsing speed will decrease if these tracker cookies are not blocked.

Over the years, Facebook has become more and more permissive with default privacy settings. Figures 2.3 and 2.4 show a significant difference in just 5 years in the availability of users' personal data on Facebook (default settings). As seen from the graph, Facebook used to restrict personal data for all of its functionalities to only Facebook users. Now most of its default features are public and the user has the choice to set his/her own privacy settings.

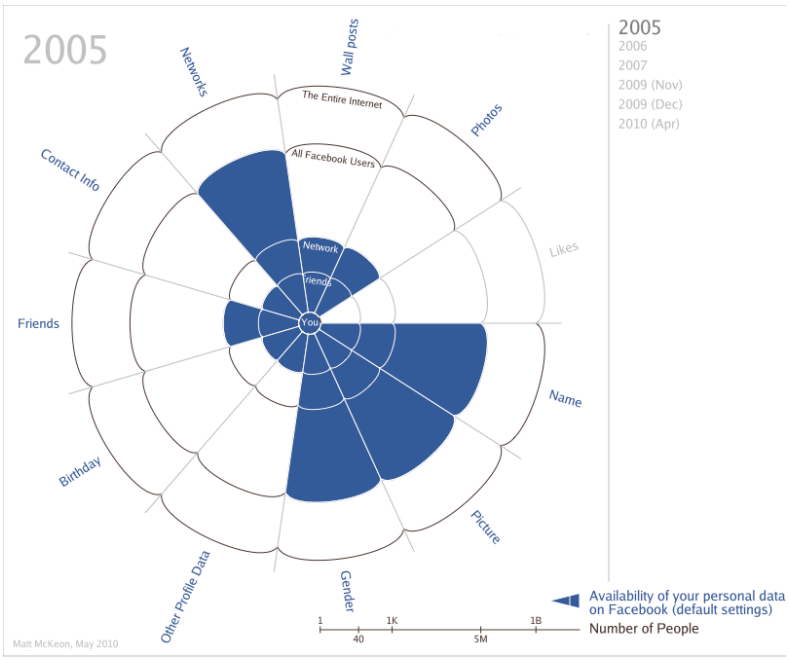

Figure 2. 3. Personal data availability on Facebook in 2005 (Source: McKeon, 2010)

The difference between the amount of information offered in 2005 and 2010 is clear as seen in Figure 2.3 and Figure 2.4. In addition, the availability of information has increased dramatically in just five years. In 2005, personal information was visible only to friends and Facebook users at a maximum. However, Figure 2.4 illustrates that information availability has spread to include all Internet users in many cases.

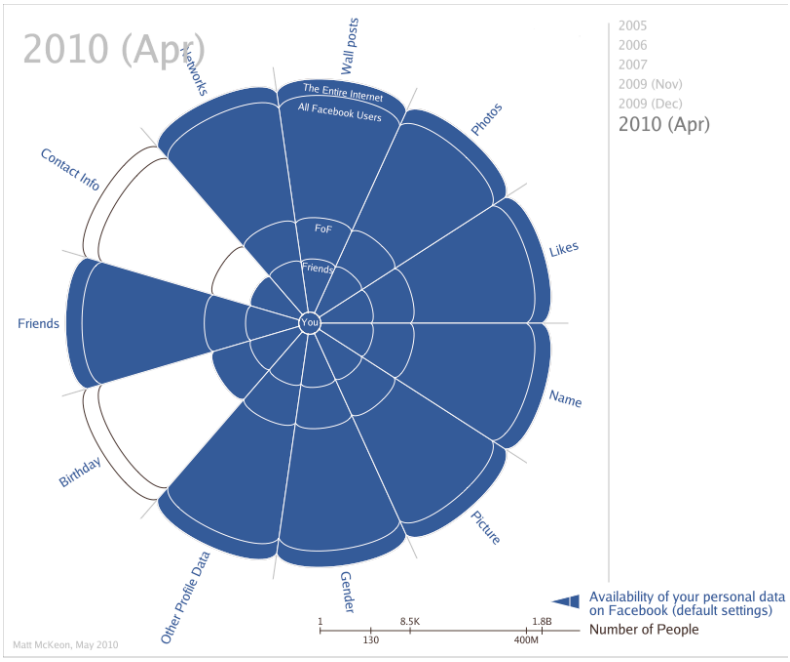

Figure 2. 4. Personal data availability on Facebook in 2010 (Source: McKeon, 2010)

2.4.1.2 Facebook's privacy policy analysis

According to Facebook's Data policy use (Facebook, 2016), it collects the following from users without notifying them and without providing any compensation

- Location of all photos users post on their walls or send to others, including information such as when the photo or the file was created
- Information on users' activities including what content they view and the duration and frequency of users' activity. However, there is no clear indication of whether the viewed content is just on Facebook pages or on all other web activities while the user is logged into Facebook.
- Device information: Facebook collects information about what devices users have used to connect to Facebook, regardless of whether these are phones or computers. This information includes:
    1. Operating system of the device, the hardware model and version, device settings (no clear indication in the policy what kind of settings are collected), file and software names and types, the strength of the signal of the device used, battery, and other device identifiers.
    2. Specific geographical location of the device, either through GPS, Bluetooth, or Wi-Fi signals.
    3. Information about the user's connection such as the name of the phone ISP, IP address, browser type, and phone number

Facebook claims in its data policy that it uses this information to improve the user experience; however, it also states that this information is used to provide effective advertisements to the user. While Facebook provides the user with the option of disabling personalised ads, the default is that Facebook is allowed to sell user information to advertising agencies in order to provide the user with more personalised advertisements. The "disable" button option is not in a clear location; the user has to go through the data policy or click through a number of buttons in order to find its location.

2.4.2 Snapchat

2.4.2.1 Snapchat's privacy analysis:

Snapchat has become popular because of its reputation for protecting user privacy. The idea behind the application is as follows:

1. The user can send a snap (either a picture or a video) to specific users, which can only be viewed once, can play for up to 10 seconds on the receiver's phone and is then automatically deleted.
2. The user can send a snap (either a picture or a video) with their story, which can be viewed, depending on the privacy settings the user has established, by anyone, only friends, or only a customized group of friends. The story can only be viewed within a 24 hr period. Then it disappears from both the receiver's and sender's Snapchat application. Users cannot save what they receive from others, as Snapchat doesn't offer a 'save' option. They can screenshot/capture the screens they receive but Snapchat will notify the user that someone has taken a screenshot. Snapchat gives its users a false sense of privacy. This has given Snapchat users the impression that their data is not saved and that no one can keep a track of what they send; therefore, many share very personal pictures and videos of themselves, friends, and family. However, Snapchat does store all videos and images shared in the app, not only in their own storage units, but also with third parties, whether they are service providers or partners, according to their privacy policy as discussed by Ayoub (2015).

The process is as follows. The sender sends a snap (either a picture or a video), which is sent and saved to the Snapchat server. When the receiver receives a snap (either a picture or a video) from the server, it is saved in the mobile device of the receiver and can be retrieved with the appropriate forensic tools. Snapchat communicates with the server through an API; many believe that it is not possible to hack into it and get to the server. In 2013, a hacker managed to get into the Snapchat server and released 4.6 million users' personal information, including their phone numbers and their contact details, and published it on the Internet (Shu, 2013). In addition, there are many third party applications on the market that save what the user has sent without the sender being aware that their picture or video was viewed and saved. These applications access the Snapchat platform, indicating that there are security holes in the Snapchat system that have not yet been

addressed by Snapchat developers (Young, 2013). An example of this is an application called SnapSave, which is available from both the app store and Google store.

A data research centre based in the US, called Decipher Forensics, has recently announced that it has discovered a method to recover Snapchat data and has begun offering public services for monetary compensation (Young, 2013).

Research was conducted by Utah Valley University as part of their Advanced Mobile Forensics course to test whether the snaps that get sent and received really disappear or not (Leavitt, 2014). The experiment was conducted on Android devices, using a Samsung Galaxy Note 2 (sender) and Samsung Galaxy S3 (receiver), where a Snapchat account was created in each device and pictures and videos were sent from the sender to the receiver. Some images were viewed while others were not. The phones were then acquired and images were forensically analysed (Leavitt, 2014). For Android devices, most Snapchat data is kept within a data/data/com.snapchat.android folder. Inside this directory, there is a "received_image_snaps" folder, where every received image is located, including viewed and expired images. This showed that all pictures sent via Snapchat do not really "disappear forever", and by using appropriate tools, these images are indeed recoverable (Leavitt, 2014).

iPhone users are not safe from image recovery either. Using a third party program called iFolder will enable the user to find expired videos received via Snapchat. Other easy and simple ways to find expired media also exist; just plugging the iPhone into the computer and using a third-party file-browsing application called iFunbox enables the user to navigate to a folder called Snapchat/tmp. In this folder, the user will be able to re-watch all the videos that were sent to him/her without notifying the user, who believes that his/her snaps have disappeared.

As for their privacy settings, Snapchat's terms of service state that the user can adjust the privacy settings of some services in regards to who can see the content created by the user. Users have two privacy settings that they can manage. As shown in Figure 2.5, users can manage who can send them snaps and who can view their stories. In terms of who can send them snaps, users can choose between two options; everyone or friends. When a user

chooses ‘everyone,’ this means that even Snapchat users who haven’t been accepted as friends can send media. For story viewing, the user gets three options; everyone, friends or custom. In the custom option, a user can block even certain friends from viewing her/his story to add more privacy.

Figure 2. 5. Snapchat privacy settings

2.4.2.2 Snapchat’s privacy policy

Snapchat’s privacy policy and terms of use are short, clear, and easy to read (Young, 2013). According to Snapchat (2016), their privacy policy starts by stating what data Snapchat collects from its users:

1. Contact information: Snapchat provides an option for its users to enter their mobile number in order to be able to find who is also a Snapchat user among their contact numbers. However, giving Snapchat permission to do that allows Snapchat to collect all your contacts’ details. This means that other users who are not Snapchat users and who did not agree to Snapchat’s terms and conditions are affected by this and their phone numbers are collected involuntarily.

2. Camera and Photos: Snapchat allows users to save their own Snapchat images. However, in order to do so, the user must give Snapchat access to their Photos as shown in Figure 2.6, which clearly shows that the purpose of the access is to save snaps. However, Snapchat’s privacy policy says that it will collect material from the device’s camera and photos, which is misleading. The user thinks that Snapchat will only save their photos, but in reality Snapchat will get access and will be able to collect photos from the device’s photo album.

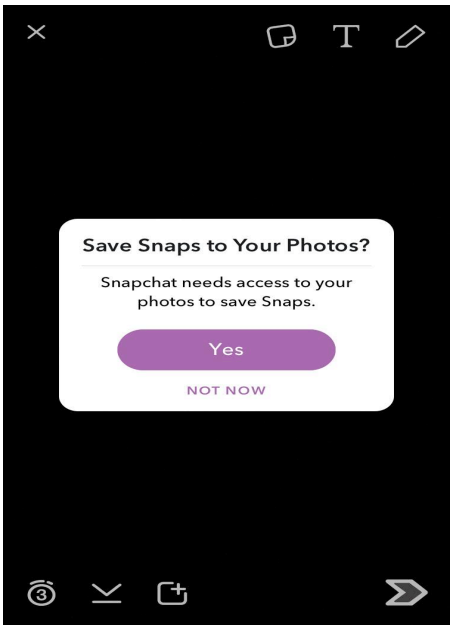

Figure 2. 6. Saving snaps locally

3. Log Information: In its privacy policy, Snapchat indicates that it collects information about how the user uses this service. Snapchat collects device information, pages viewed, IP addresses, cookies that can uniquely identify the user device, and most importantly, it states that it collects information about the pages that the user has previously visited before navigating to the device. Since Snapchat is a mobile application that does not have web browser access, it is not clear what kind of "pages" it monitors.

4. Location information: Snapchat collects users' precise locations through GPS, Wi-Fi access towers and cell towers. Snapchat asks for access permission for the user's location when the user wishes to use features such as GeoFilters.

Furthermore, according to Snapchat's privacy policy, "you grant Snapchat a worldwide, perpetual, royalty-free, sub-licensable, and transferable license to host, store, use, display, reproduce, modify, adapt, edit, publish, create derivative works from, publicly perform, broadcast, distribute, syndicate, promote, exhibit, and publicly display that content in any form and in any and all media or distribution methods (now known or later developed)." (Snapchat, 2016)

In addition, Snapchat's privacy policy insists that the default is to delete all exchanged private snaps once they are viewed and expired, because it claims that the server is

programmed to delete viewed or expired snaps from the recipient's device. However, as previously discussed, the images are not fully deleted and can be recovered using forensic tools and third party programs.

2.4.3 Instagram

2.4.3.1 Instagram's privacy analysis

An Instagram profile can be either all public or all private. Users who may wish to post private pictures and videos about themselves can adjust the privacy settings. This can give the user a feeling of security, and a belief that no one but his/her choice of followers can see his/her private posts. However, Instagram is not immune to security holes that can threaten the user's online privacy. In 2015, a privacy hole was discovered that exposed an enormous number of pictures and videos that users believed would remain private (Yanofsky, 2015). The hole in the system occurred with pictures/videos that were posted publicly but then made private by the user. Anyone who had a link to the photo/video was able to view it even if it was then made private. This can also happen with content that is posted in a private account. In addition, Instagram allows users to share their posted content to other SNSs such as Facebook, Twitter, Flickr and Foursquare, from inside the Instagram app. If a user with a private account shares a photo/video to other SNS services at the same time they are posting it to Instagram, their photo/video is viewable by anyone with the link. After the Quartz report, Instagram stated that it had fixed the problem with a new update, which makes photos private even if they had previously been public, (Yanofsky, 2015).

2.4.3.2 Instagram's privacy policy analysis

According to Instagram's privacy policy (Instagram, 2013), the following attributes are what Instagram collects from users:

- Sign-up information: Username, password, email address.
- Profile Information: first and last name, profile picture, phone number
- User content: this includes all posted information such as photos, videos, comments and other material. It is not specified what is meant by other "material".

- The user's contact list, which includes all the information about the user's contacts in their mobile device. However, the user has to give Instagram permission to access their contact list.
- Web pages visited by the user.
- Instagram uses cookies and other similar technologies in order collect information about how users use Instagram, which is used by Instagram and other third parties.
- Log information about the user's usage of the service, such as which IP was used to access the service, domain names, pages view, and any other similar information.
- Device identifiers are minor data files that are stored in the device to uniquely identify it. Instagram may access users' devices that store these data structures. Third parties use these device identifiers in order to personalize ads. The privacy policy does not state whether the user has the choice of disabling such technologies that are used by third parties to track the user's mobile behaviour.

Instagram states clearly in its privacy policy that deactivating, terminating, or deleting the account won't result in the disappearance of the user's content; it might still be viewable or searchable in archived copies or by other means. In addition, Instagram clarifies that the user owns their photos or videos, but posting such material to Instagram gives Instagram a non-exclusive license to use the content that the user posts. If the user has made his/her account private, Instagram will keep the content private. However, if it is public, Instagram has a license to use this content.

2.4.4 Twitter

2.4.4.1 Twitter's privacy analysis

Compared to other SNSs profiles such as Facebook, Twitter profiles contain fewer information categories. The user profile contains their name, header photo (optional), profile photo, a 160-character bio (optional), location (optional), link to their website (optional), birthday (optional), following and followers, and the user's tweets, which can include videos and photos. Figure 2.7 displays an example of a Twitter profile. The focus of Twitter is not the specific information about the user, it is the tweets themselves. According to Humphreys, Gill, and Krishnamurthy (2011), tweets rarely contain personally

identifiable information such as telephone numbers, email or home addresses. However, they do contain information about the activities the user is engaged with in their non-digital world. Leaking sensitive information via tweets can have an effect on the user's offline life. For instance, tweeting about going away on a vacation makes the user vulnerable to theft. As for security breaches, Twitter has had its fair share. In 2013, Twitter announced that more than 250,000 accounts had been hacked (Jones, 2013). In a sophisticated operation, an anonymous attacker was able to obtain more than 250,000 accounts' usernames, emails, and passwords. The attack affected more than a quarter of a million Twitter users at that time (Jones, 2013).

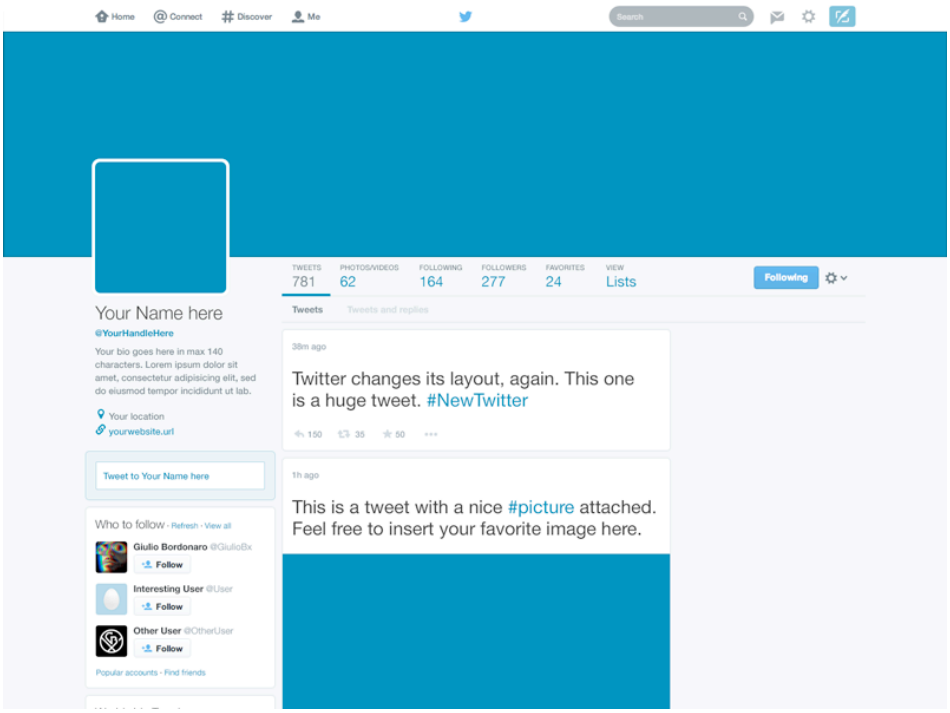

Figure 2. 7. An example of a Twitter profile

2.4.4.2 Twitter's privacy policy analysis

Twitter has two separate companies that collect users' information. If the user is based in the US, an American company, Twitter Inc., collects the user data. However if the user is anywhere else in the world, an Irish company called Twitter International Company collects the data. According to Twitter (2016), the type of data collected includes:

1. Basic Account Information: the information that the user provides when logging in, such as name, username, and email address. Twitter lists the name and the username publicly, which makes them searchable in search engines.
2. Profile additional information: The user's biography, location, website, date of birth, or a profile and cover picture.
4. Tweets, Following, Lists and other Public Information: This includes the metadata of the tweets such as when the user tweets, what client the user uses, the location, time zone and language that is associated with the tweet.
5. Location Information:
6. Links: Twitter tracks the links the user clicks and where it redirects them.
7. Cookies: Twitter uses both session cookies and persistent cookies in order to collect usage data.
8. Log Data: This can include IP address, browser and operating system type, pages visited, precise location information, mobile service provider, mobile device type and ID, search terms, and cookie information.
9. Widget Data: Third party websites that contains Twitter buttons or widgets automatically send log data to Twitter. At first, the information provided is personally identifiable. However, after 10 days, Twitter starts the process of removing or de-identifying widget data, which might take up to a week. This information is mostly used for tailoring ads.
10. Commerce Services: Some Twitter services require the user to provide payment information such as their credit card details and shipping address, which are stored privately.

How Twitter uses the information:

The default is always making information public and collectable by Twitter. Twitter tracks what other websites the user has recently visited. Twitter uses the collected information to provide tailored ads and suggestions for the user by sharing it with Twitter's ad partners, which is agreed to by default when users sign up to Twitter. However, after a maximum of 10 days, Twitter claims that it removes all account identifiers from the data that was collected from website pages the user visited. In order to provide users more privacy, Twitter supports a no-tracking privacy preference, which can be set in web browsers and on Twitter accounts. Do Not Track (DNT) enables users to control how their information is

used by Twitter in order to personalize their experience. However, because the tracking is automatically done, users need to visit their account and browser settings and disable these functions by unchecking the boxes. In addition, Twitter gets information about the user from ad partners and matches it with the user's Twitter account. Based on this information, the user will receive ads. Twitter can receive email lists from businesses and may be asked to promote products or services on Twitter by matching the email lists to Twitter accounts. Alternatively, the business can share with Twitter information about the users who visit the business website so that Twitter can match their account and these users receive the ad. Users need to uncheck the automatically checked buttons if they do not wish their account to be matched with information provided about them by Twitter's ad partners.

When a user chooses to deactivate his/her account, Twitter will still own the rights to the user's content, except for the copyright license, which survives termination. After 30 days, Twitter starts the process of deleting the user's account. However, even after deactivating the user account or deleting, the user's public tweets will not disappear from search engines and other third parties.

Furthermore, Twitter can amend its terms and conditions policy any time it sees fit. It will notify the user, via their official twitter account or via email. In September 2009, Twitter gave very short notice before it substantially altered the terms and conditions policy for the copyright license, just a few hours before it applied the changes to its policy (Tosdr, 2012). Even if the users didn't check their email or their Twitter account, if they continued to access Twitter after the changes to the policy they were still bound by the revised policy.

In its privacy policy, Twitter indicates that users should carefully think about the public content that they post because as soon as a user posts something in their open profile, the information is immediately sent to Twitter's partners and third parties via SMS and APIs. This includes developers, search engines and publishers, in order for them to integrate Twitter content into their services. Additionally, Twitter's information sharing includes organizations such as universities and public health institutions, which use and analyse the information in order to produce trends and insights. The use of the word "immediately" in its privacy policy indicates that all the information is transferred automatically, which may also indicate that even if the user deletes the information they posted after some little time,

the information will still be there with Twitter's partners. There is no indication by Twitter whether the information shared with their third party partners is anonymised.

2.5 CONCLUSION

Chapter 2 delivered a comprehensive literature review on the privacy of user's information on SNSs. Firstly, this chapter provided background information on how SNSs started and how they acquired the forms that we know today. The idea of SNSs has been around for a long time; however, it did not become immediately successful because not many people were initially connected to the Internet. This indicates that online users originally did not value the idea of joining SNSs until their offline social networks of friends and family joinedl. The chapter then defined social networks by listing the characteristics that differentiate SNSs from other websites and applications. Section 2.2 highlighted the differences between existing SNSs. The chapter then reviewed literature on privacy in SNSs, discussing the difference between security and privacy and outlining the possible sources of privacy breaches. The chapter then discussed the value of privacy and the possible privacy issues that may result from using SNSs. Lastly, the chapter provided a detailed analysis of Facebook, Twitter, Instagram, and Snapchat privacy and security.

Chapter 3, the research methodology chapter, will formulate the research plan and outline the research methodology by reviewing studies related to the chosen research area. The main research questions and associated sub questions will be established and presented.

Chapter 3: Research Methodology

3.0 INTRODUCTION

Chapter 2 reviewed the literature relevant to SNSs' history, features, privacy concerns, and privacy policies. This chapter describes the methodology used to identify the level of privacy awareness and the degree of information disclosure among social network users. Chapter 4 presents the findings of the applied methodology, and Chapter 5 discusses the findings and establishes guidelines and instructions to protect users' security and privacy from malicious attackers and privacy violations committed by service providers and other third parties.

In order to develop the research methodology, similar studies were reviewed and analysed. In section 3.1, three studies selected from the reviewed sources serve as a guideline for the design and implementation of the research methodology. After analysing related work in the area of SNS privacy, based on the literature that has been reviewed, the research questions for this study are presented in section 3.2. An appropriate research methodology is described in section 3.3 in order to answer the research questions presented in section 3.2.

Section 3.4 discusses the techniques used in this research to collect relevant data, and discusses the reliability and validity of the collected data. Section 3.5 deals with data processing and analysis, while section 3.6 discusses the ethical concerns faced in this research and how they were addressed.

3.1 REVIEW OF RELATED WORK

Three studies were selected from the available literature and analysed to assist with the design of the research method for this thesis. These studies illustrate the many different approaches to studying how users behave in social networks and how this can affect their privacy and security. The results of these studies are also used in Chapter 5 to provide

comparison with the results from the present study and thus help to answer the research questions.

3.1.1 From Privacy Concerns to Uses of Social Network Sites: A Cultural Comparison via User Survey

This study was conducted by Ho Keung Tsoi and Li Chen from the department of Computer Science at Hong Kong Baptist University in 2011. In this study, Tsoi and Chen (2011) examined the effect of cultural variables on users' privacy concerns and trust in SNSs, and how this affected users' motivation to use such sites, their actual usage, their attitudes and likely future behaviour. The paper focused on the differences between Hong Kong and French SNS users with respect to a number of measures. The purpose of the survey was to identify whether the two cultural groups had different levels of privacy concerns with regard to SNS use. In addition, Tsoi and Chen investigated whether the two groups' differences regarding privacy would influence their trust in SNSs and their motivation regarding SNS use. For instance, if an SNS user is very concerned about the possibility that their personal information will be used by the site owners for purposes other than merely displaying the information, will that user be less likely to trust the site and thus less motivated to share information? Finally, the researchers wanted to see whether users' privacy concerns, trust and motivation would influence their actual usage of SNSs, their overall attitudes, and any other future behavioural intentions. The researchers applied the Theory of Planned Behaviour in the SNS context to identify associations between the variables. This theory is a predictive and persuasive type, where the subjective norm (in the case of this research, trust, privacy, and enjoyability), is connected to individual behaviours such as attitude and behaviour intentions. Figure 3.1 illustrates the relationship between the measured variables and the research hypotheses.

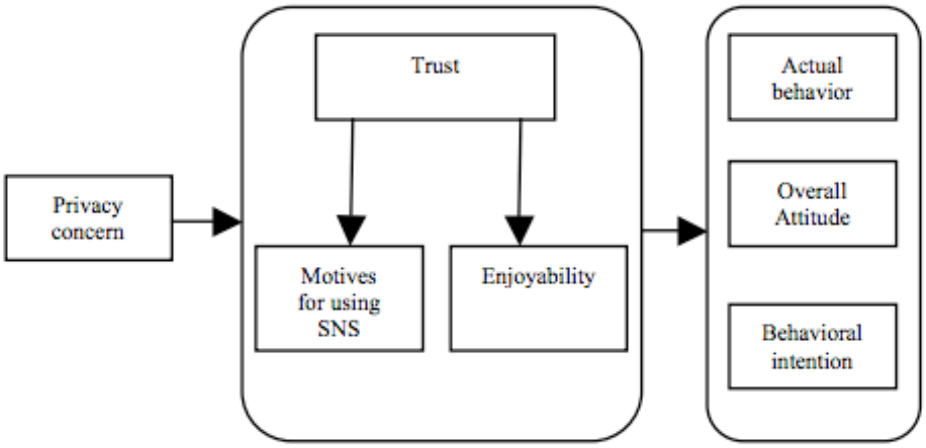

Figure 3. 1. The relationship between measured variables and the research hypotheses. Source: Tsoi and Chen (2011)

An online survey was used to collect the data: the researchers obtained 154 participants. The survey was distributed through French and Hong Kong public messaging boards and popular forums. Gender, age, educational background and profession were used to classify the survey participants. The survey's main focus was on privacy and trust concerns. The results were analysed using multivariate analysis of covariance. According to Tsoi and Chen, this tool was used because of its ability to adjust mean values and because it is able to identify any differences that can be attributed to nationality or other possible factors such as gender. The survey used a 5-point Likert scaling method, For instance, some questions had answers that ranged from 'very seldom' to 'very often.'

The survey was composed of three sections. The first section contained questions associated with privacy and disclosure of personal information; the degree of comfort the users felt when giving personal information in their SNSs profile; their control over their profile when specifying/updating information; and the overall privacy protection that the user perceived for the SNS. Tsoi and Chen (2011) also asked questions about users' general privacy concerns when using the Internet; their specific and current privacy settings in SNSs; the types of personal information that users provided in their profiles; and the type of posts that they often posted on the site. The second section had questions relating to users' level of trust in SNSs; their main motives for using SNSs; and the degree of enjoyment they got from using SNSs. The third section contained questions about SNS users' overall attitudes towards SNSs, such as whether they considered the use of SNSs to be part of their

everyday activity. The third section also contained questions about users' behavioural intentions; for instance, measuring the level of intention to keep using SNSs more frequently and on a regular basis, and intention to invite friends to use and join them in the network. The privacy questions were used as dependent variables, whereas nationality and gender were used as two covariates. The results showed that nationality had a significant effect on the differences, but gender did not have a significant statistical effect. After determining how the survey participants perceived their SNS privacy, Tsoi and Chen evaluated the kind of personal information that was disclosed in user SNS profiles of French and Hong Kong (HK) participants to determine the differences. The results showed the HK users tended to share more identifying information than the French users. Further results, combined with users' privacy concerns, implied that French SNS users' higher privacy worries probably resulted in a lower level of disclosure of personal information. The lower sharing rate also indicated that they were having a less enjoyable experience, resulting in less motivation to share. Multiple regression analyses were conducted in order to correlate users' privacy concerns with trust, motives, and enjoyability. The latter factors were correlated to users' actual SNS usage, overall attitudes when using SNSs, and behavioural intentions. The results showed a difference in causal factors between the two cultural groups. The French users' results showed the three privacy factors all had effects on one or more of their motives. For HK users, the only factor that played a significant role in influencing the users' motives was their degree of control in updating their profiles. The results also showed that French users' visiting frequency, attitudes, and intentions to using SNSs were genuinely affected by privacy factors. Tsoi and Chen concluded that the French users' higher privacy concerns affected their use of SNSs differently to those of HK users, who were more active due to their lower privacy concerns.

3.1.2 A Survey of the degree of online self-disclosure (DOSD)

In this research, conducted in China, Ge et al. (2014) studied privacy leakage issues in Pengyou, one of the most widely used SNSs in China. They explored how users responded to friendship invitations from strangers and therefore their willingness to give access to private information and lists of friends. The researchers created eight fake user profiles, each of which then gained a number of online friends by sending daily friend requests to random strangers. They gained access to 2761 profiles. They then proceeded to

quantitatively analyse and examine the degree of online self-disclosure (DOSD), the age distribution of the fake users' friends, and photographic information leakage from the fake users' friends. The researchers used field leakage and DOSD to quantify users' tendency to disclose their own personal information. Figure 3.2 illustrates the average DOSD in different ages and genders based on the researchers' calculation of the profiles categorised. The calculation used four formulae, which are listed in the paper. However, Ge et al. (2014) did not explain how they derived the mathematical formulae. Personal information was analysed based on occupation, residential address, education, email address, hobbies, birthday information, mailing address, and telephone.

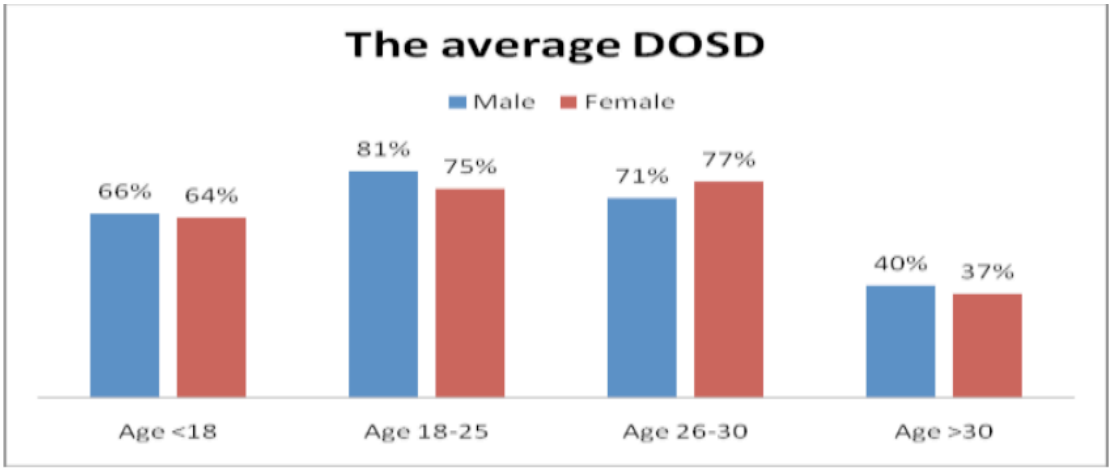

Figure 3. 2. The average DOSD by gender and age. Source: Ge et al. (2014)

The survey results revealed some serious privacy threats for SNS users, particularly for minors and young people. The groups most willing to disclose their personal information in that particular SNS were men (81%) aged 18-25 and women (77%) in the same age group.

3.1.3 Social Networking and Online Privacy: Facebook Users' Perceptions

This study examined the perception of SNS online privacy among Facebook users in regard to their current level of awareness of privacy issues and how it influenced their behaviour online. Torres (2012) began by assessing and evaluating the relationship between trust and privacy offline, online, and in the SNS environment. According to Torres, there is a correlated relationship between trust and information disclosure; the more a person trusts an environment, the more they disclose information about themselves. In this study, Torres

examined Facebook business evolution and the impact it had on privacy and user information disclosure. The chosen research methodology was a case study that involved a number of data collection methods and different analysis techniques, including both quantitative and qualitative methods. Firstly, Facebook privacy was comprehensively discussed and analysed from many different angles, including privacy control, privacy policies, and privacy settings. In addition, the level of trust among users was discussed based on input from other literature. In order to collect data, the research conducted focus groups to assist the development of the main data collection method, which was a survey. Torres used focus groups to gain an insight into users' behaviour on SNSs and to examine users' attitudes in regard to privacy and trust in the SNS environment. Torres qualitatively analysed the data and the result played a major role in developing the survey. The focus groups consisted of three active Facebook members from three different age groups, in order to observe their different behaviour and their concerns regarding privacy and trust in SNSs. The age groups were 18-21, 22-29 and 30+ years of age respectively. The results revealed that there was an element of uncertainty regarding privacy issues in Facebook and that there was a contradiction between users' privacy concerns and their behaviour in the network. From the focus groups, Torres developed a survey aimed at Facebook users with the purpose of evaluating their levels of awareness specifically in regards to privacy issues on SNSs. She also aimed to determine the level of influence user awareness had on attitudes and trust levels. The data collection instrument used was Survey Monkey, a web-based tool. The survey contained 23 questions that were developed by reviewing literature, similar successful studies, privacy policies, and statements from the focus group discussion. Data analysis involved descriptive statistics, filtering of survey answers, cross-tabulations, and analysis across age groups. A survey pilot test was conducted to ensure the reliability of the study and to verify that the questions were worded appropriately for the target group. Torres then used the results to provide privacy recommendations for both Facebook as a company and for users on how to ensure that users are fully aware of the consequences of their actions when using the network. The study showed that Facebook users believed that Facebook was obligated to protect their privacy and their information.

3.2 RESEARCH QUESTIONS

The research questions were established based on the literature review in Chapter 2 and similar research that has been conducted in the area of SNS privacy, as described in section 3.1. Based on the literature, it is evident that SNS privacy and security issues are widespread issues and a newly developing area in the academic literature, particularly since popular social network apps and websites have emerged recently and relatively little research has been conducted in that field. Therefore, the aim of this research was to identify the extent of privacy awareness among SNS users; how that can affect issues such as online identity theft; and how users can protect their security and privacy when using online social networks. In order to achieve the purpose of this research, three research questions were identified and are presented below.

The aim of the first main research question is to understand whether the three factors of age, education and gender have an impact on the way users behave on SNSs regarding their privacy awareness. In order to answer the main research question and identify the effect of these factors on levels of self-disclosure, the actions of the users needed to be assessed based on the data collected.

Q1: What are the personal attributes that can have an influence on information disclosure and privacy settings of SNSs users?

The following sub-questions were derived in order to answer the first main research question stated above:

Sub-Question 1 (SQ1):
What is the influence of gender on SNS users' information disclosure and privacy settings for Facebook, Snapchat, Instagram and Twitter?

Sub-Question 2 (SQ2):

What is the influence of age on SNS users' information disclosure and privacy settings for Facebook, Snapchat, Instagram and Twitter?

Sub-Question 3 (SQ3):

What is the influence of education on SNS users' information disclosure and privacy settings for Facebook, Snapchat, Instagram and Twitter?

The aim of the second research question was to identify whether users' levels of privacy concern, (identified and categorised from the collected data) have any impact on what they reveal in terms of personal information.

Q2: How do users' levels of privacy concern affect the amount of information they disclose in social networking sites?

As for the third main research question, a significant part of this research focuses on privacy policies and users' knowledge. Chapter 2 analysed four privacy policies of Facebook, Twitter, Instagram, and Snapchat and interesting findings were made. The aim of this question is to survey users with the results of the analysis and identify their level of their awareness of the privacy policies of each SNS.

Q3: How aware are users of the extent to which their information is protected by SNS providers according to the privacy policies that the users have agreed to?

3.3 RESEARCH DESIGN:

In order to answer the research questions, data was collected from two different sources: an online survey and a social experiment. The focus of both the survey and the social experiment was on four different social networks: Facebook, Twitter, Snapchat and Instagram. Each one of these SNSs has its own unique purpose and functionality, and displays different kinds of personal information, as explained in section 2.3. Due to the uniqueness and popularity of each of the chosen networks and the fact that each one requests and presents different kinds of information, it was decided that data would be collected for each one, rather than generalizing the survey questions and the social experiment on all social networks as a single entity.

Research can be a combination of different kinds of approaches that are used together in order to answer the research questions. The research in this thesis is categorised into the following basic types. Firstly, this was descriptive research since it included a survey as the primary source of the research data. It therefore sought to describe the current state of

affairs, as it existed in the current environment, without any interference from the researcher, who exerted no control over it (Thomas, 2003). The social experiment was also categorised as descriptive rather than analytical because its purpose was to measure the current level of awareness of SNS users. Secondly, because this research was data-based, aiming to reach conclusions that were verified by both the survey and the social experiment, it was considered to be empirical research since it derived knowledge from measured responses rather than a theory or a belief (Thomas, 2003). In this research, two different methods were chosen to collect the data. For the purposes of this study, the research did not follow a specific framework to develop the methodology. The chosen methods were developed and improved based upon on the study of previous related research, which was described in section 3.1, and other previous literature. The two methods are described below in section 3.3.1 and 3.3.2:

3.3.1 Survey Description:

The survey was designed to provide insight into the following areas of interest. Firstly, SNS users voluntarily enter personal information about themselves, their family members and/or their friends. The survey questions identified the type and amount of information disclosed by users and what type of privacy settings were applied with regard to whether such information was made available to the public or not. Prior to developing the survey material, a literature review was conducted, SNS privacy policies were reviewed and analysed, and SNS features and privacy settings were studied accordingly.

The survey was divided into six sections (pages) and contained 31 questions, which were a mixture of dichotomous questions, nominal/measurement questions, multiple choice questions, and filter/contingency questions. The number of questions was kept to the minimum possible since participants can lose interest in completing surveys if too many questions are asked.

The first two sections were general; the questions were designed to provide research data giving a comprehensive overview of how SNS users used SNSs, their motivations, and how active they were. It also included questions about how they perceived the privacy of their information in SNSs, and their level of trust towards SNS service owners. The purpose of

identifying users' degree of privacy perception in the earlier sections was to compare their perceptions with their actions in the network. The following are the two first general sections of the survey:

1. Demographics: In order to categorise the users into statistical groups from the sample, required information included age, gender, and education.

2. Use of SNSs and main privacy concerns: This section aimed to provide a general insight about users' trusts levels and privacy perceptions of SNSs as a whole. The section started with general questions such as the user's motive behind joining SNSs and the amount of time spent using SNSs, then proceeded to ask the user more detailed questions such as whether they read privacy policies. The hypothesis is that very few users actually read privacy policies. The follow-up question asked about their reasons for not reading such policies; whether it was because they took time, involved complicated legal language, and/or users simply did not care. The participants were provided with an "other option" box to provide other reasons that had not been listed as options. The participants were then required to answer a main question for this survey, which was whether privacy was an actual concern for them. The answers to this question were then compared in later sections with users' actual practices in specific SNSs.

The later four sections are more specific: each is concerned with one social network from the group of Facebook, Twitter, Snapchat, and Instagram. Participants were able to skip any of these pages if they were not users of that particular social network. Each page's questions aimed to acquire information about the user's behaviour and actions in the network. The structure of the questions was similar for each social network page, but differed in detail because each SNS has its own data policy and displays different information about its users. The common structure of the questions in the last four sections was as follows:

1. The survey asked the participants what kind of name they provided in each social network page: real name (first and last); part of the real name; or a fake name. The aim of this question was to measure how users protect their identity online from identity thieves or others. The result of this question was combined with the results of other questions to

compare whether users provided more public and accurate information about themselves and whether there was an ability to link users' identities across different networks, which can increase the amount of personal information they provide about themselves online.

2. The second type of question in the survey involved multiple questions with yes/no answers that were specific to each SNS. For instance, Figure 3.2 shows the Facebook survey page questions. The aim of these questions was to understand the social network user behaviours that can substantially affect privacy and security. For instance, Facebook recently published a new feature called "Nearby friends". This feature informs the user about other users' locations and how far away they are from you. Enabling this feature can affect the user's security and privacy, particularly when they have added people they do not know personally in real life. This has the potential to compromise the user's physical security.

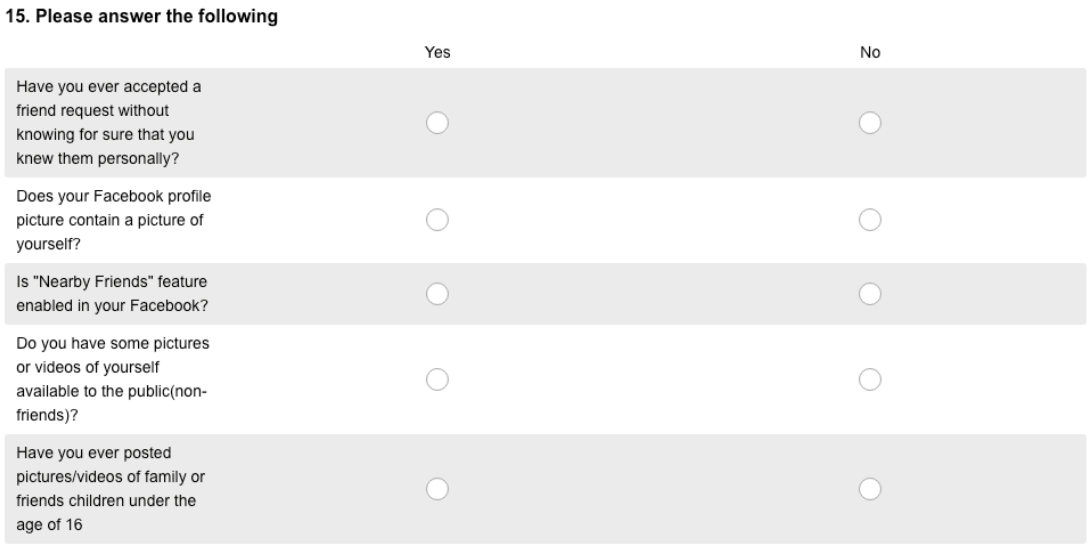

Figure 3. 3. An example of a question from the survey

3. Privacy settings: Each SNS survey section contained questions about the privacy settings users implemented in their accounts. As discussed in Chapter 2, there are many settings that are set up by default that might not necessarily serve the best interests of the user. The aim of these types of questions was to identify and measure the level of privacy protection users had. For instance, in the Facebook survey page, the user was asked to review their privacy settings for displaying the following information: profile picture, hometown, current city, family members, relationship status, birthday, education, events, locations visited, friends

list, and contact information such as emails and phone number. This information can be displayed to the public, friends, a customised group of friends, or not shared. According to Lewis (2015), displaying such personal information publicly can facilitate identity theft and fraud. The survey results identified what personally sensitive information users publicly shared and what they kept private in different SNSs.

4. Privacy policies questions: as discussed and analysed in Chapter 2, SNS privacy policies contain critical information about how the user data is treated. The assumption is that many users are not aware of exactly what they're agreeing to when joining SNSs. Therefore; the users were presented with a list of actions in each SNS survey page. Those actions were retrieved from the analysis conducted in Chapter 2, and for each action the user could answer yes or no. In order to measure the level of awareness among SNS users, this question asked the survey participants which of those actions the social network has the legal right to do based on the terms and conditions that the user had agreed to. The purpose of this question was to measure user awareness of how their data is handled and processed.

3.3.2 Social Experiment description:

The purpose of this experiment was to test the level of people's awareness by evaluating how they responded to strangers requesting access to their private information in a real-life situation. SNS users might not share information publicly; hence, they might feel confident that no one but authorized and known friends/followers can access their information. The purpose of this experiment was to test the ease with which strangers, who might be malicious, could access users' private profiles in the four selected social networks in this research: Facebook, Twitter, Snapchat, and Instagram. Four fake profiles were created using fake names and very minimal displayed information. The purpose was to make the profiles as vague as possible, which indicates that the user has no likelihood of detecting the person behind the invitation. Then, friend requests were sent to random people over the social network. If the user accepted the friend request, they gave access to their private information that wouldn't be viewable otherwise. For each social network used in this experiment, there is an option for users to send a private message before accepting an invitation. The data collected was only whether the user accepted the friendship request or

not and whether they had attempted to identify the sender's identity before accepting. No personal data was collected.

3.4 DATA COLLECTION:

For the survey, the link was distributed in the four social networks in order to ensure that targeted participants were reached. In Twitter, for instance, the link for the survey was tweeted with trending hash-tags in order to ensure it had wide exposure. In the post, there was a brief description of the survey in order to encourage users to take part in it. For the social experiment, users were selected randomly from their participation in public pages such as newspapers or public figures' pages by either liking a post or commenting on a post.

3.4.1 Target population and sample size

The survey and the social experiment targeted users on four social networks: Facebook, Twitter, Instagram, and Snapchat. In order to calculate the sample size, Survey Monkey provides a tool to calculate the required number of respondents. Three attributes needed to be determined in order to calculate a representative sample size. According to Statista (2016), these four networks have over 2.34 billion users: Facebook has 1.59 billion; Twitter has 313 million; Snapchat has 100 million; and Instagram has 400 million. There are, as stated above, different population sizes for each network. However, due to the fact the sample size requirement does not change significantly for population sizes larger than 100,000, the sample size was calculated in the same way for all networks. Therefore, with a margin of error of 5%, a confidence level of 95%, and population size of 1.5 billion, the required sample size was 385.

3.4.2 Validity and reliability

It is extremely important to create a research methodology that yields results that are valid and reliable. According to Bryman and Bell (2011), validity refers to whether an instrument used to gauge an idea indeed measures that concept. In order to maintain the validity of this research, the literature on SNS privacy was carefully examined in order to formulate the

research questions. Other questionnaires from other studies were extensively analysed in order to formulate the questions for this study and to adopt and adapt material from previous studies. Privacy policies of SNSs were also fully read and studied to support the purpose of this research and extract their statements. To ensure the reliability of this research, a pilot test was conducted in order to verify questions were worded correctly and to get feedback from the participants about whether there was anything unclear that needed further clarification. Survey Monkey was chosen as the tool to deliver the survey questions because of its clear layout and clear instructions to participants on how to navigate between pages and questions, which increased the reliability of this study.

3.5 DATA ANALYSIS

Two stages of analysis were used in this research to derive the main findings:

1. Exploratory data analysis (EDA): In this stage, the data files were viewed before completion of the data collection in order to get some ideas about the initial results. The purpose of this stage is that it may indicate further data are required: for instance, there may be more female responses than male responses, which could affect the accuracy of the results. This preliminary stage ensured that any imbalances and limitations in the data were resolved before the end of the data collection period. This stage overlaps with data cleaning because anomalies can become evident. Therefore, in an optimal situation, before the end of this stage, there should be a clean dataset that is ready for the next stage of analysis.

2. Deriving the main findings: This stage generates a summary of the findings, relationships, trends, interpretations and narratives. When analysing the data, the type of questions dictate the type of analysis. However, in general, two tools were used together to analyse the data. The first tool was filtering, which is provided by Survey Monkey to help break down the results in order to focus on a specific data subset. It allows viewing specific respondents' answers to specific questions. For instance, it allows viewing of all the answers of male respondents who are between the ages of 20-24 years and who answered that they do not trust SNS providers with their information. Secondly, the information is transferred into SPSS in order to analyse it statistically. Factor analysis was conducted. Separate chi-square tests of contingencies were conducted in order to understand and determine the differences in user privacy-setting behaviours and personal information

disclosure variables with gender, age, education, and privacy rating for each of the four social networks. All chi-squares were interpreted at a conservative alpha of .01 to control for multiple tests. The Chi-square analysis helps to determine whether two discrete variables have any statistical association and whether there is a statistical significance between the variables.

The social experiment was only analysed statistically as it involved no direct discussion with the users. The analysis involved only two options: the user accepted the request/the user didn't accept the request. This data was added to the research to enable discussion of the proportion of people in the sample who were willing to expose their personal information online to complete strangers. When the participant accepted the friendship request or follow request, their response was manually recorded in an Excel spread sheet by the primary researcher.

3.6 ETHICAL APPROVAL:

Both the survey and the social experiment were approved by the AUT Ethics Committee (Ethics application number 15/429). The research collected data from two sources and studied Facebook, Instagram, Snapchat and Twitter.

The first source of data collection was an online-based survey that was distributed via the above networks. The survey was completely anonymous and did not require the participants to provide any personal information. The link to the survey was posted in different social media accounts with no direct invitations; no specific users were asked to take the survey. It was completely voluntary with no pressure on the user to complete it. Survey results were anonymous, de-identified data. Data was not shared with third parties and no IP addresses were tracked.

The second data source was the social experiment, which allowed for testing user security from a different angle than the survey. Users might not share much information "publicly", but how easy it is to get into their private accounts where everything is available? The experiment was conducted by creating a completely fictional profile then sending a friend request or following/adding request to random people who had closed profiles. Once the

respondent accepted the friendship request, they were sent a private message informing them that this was part of an academic research project that aimed to understand security awareness among social network users and asking them whether their response could be recorded or not. If they accepted, they were provided with an information sheet about the research and asked for formal consent.

3.7 CONCLUSION

Chapter 3 has presented the proposed research framework, which includes the research methodology, the three main research questions and related sub-questions, data collection methods, and data analysis methods. Relevant academic studies in the area of SNS privacy have been reviewed and analysed in order to provide guidance while developing a methodology suitable for the purpose of this research. Chapter 4 presents the research findings of the survey and the experimental scenario produced by applying the methodology discussed in this chapter.

Chapter 4: Research Findings and Analysis

4.0 INTRODUCTION

The privacy of users' information in SNSs is becoming one of the most discussed topics in both the industry and academia (Li, Lin, & Wang, 2015). The review of literature in Chapter 2 discussed the types of privacy issues that users may encounter while using SNSs. It also analysed how SNS owners collect, process and share users' data according to their privacy policies and other sources. Chapter 3 established and presented the methodology that was used in this research to investigate users' privacy concerns, settings protection, awareness, and personal information disclosure in four different SNSs: Facebook, Twitter, Instagram and Snapchat. This chapter presents the findings of the investigation. The chapter is divided into two main sections: survey findings and social experiment findings.

Firstly, section 4.1 presents the survey findings, starting with the response rate in section 4.1. A summary of the main findings of the survey questions is then presented in section 4.1.2. Lastly, the statistical analysis results are presented in section 4.1.3. The survey results are analysed both qualitatively and quantitatively based on the nature of the questions and their outcomes. For the quantitative analysis, statistical analysis using SPSS was performed as described in section 4.1.3 in order to derive the main findings for this research and identify any trends, associations or relationships found in the data set that could assist in answering the research questions. Secondly, section 4.2 reports the findings of the secondary data collection method, the social experiment, which was conducted to support the data acquired from the survey. As previously explained in Chapter 3, this method involved an experiment to test the ease with which users' SNS profiles can be accessed. Overall, the results are reported throughout this chapter in a descriptive and visual manner.

4.1 THE SURVEY

The survey section presents the findings of the survey questions and the data collection process. Firstly, section 4.1.1 describes the response rate to the survey and the general

reaction to it. The following two sections present a summary of the survey question findings and the statistical analysis of the data. In the summary of findings section, each question asked in the survey is presented, with its findings, in either a graph or a table. In the second section, statistical analysis is performed on the main survey questions in order to identify trends and relationships that may assist with understanding the data and maximising insight into the collected data set, to help in answering the research questions.

4.1.1 Response rate

The survey was distributed to SNS users in multiple social network platforms and was open for participation from January 2016 to March 2016. During that period, the responses were checked regularly to remove any incomplete responses. The reason for doing so is that incomplete responses create a false impression that sufficient responses have been collected; therefore, checking responses in advance provides a clear idea of how many complete responses are still needed to provide a clean dataset and thus a legitimate and representative sample. In addition, incomplete responses affect the data analysis since the data set would have many missing values. Therefore, removing any incomplete responses before beginning analysis improves efficiency. As discussed in section 3.5, this phase was the exploratory data analysis stage. In total, 415 responses were gathered, containing 385 completed responses, meaning a completed response rate of 93%. The number of completed responses met the target that was calculated in Chapter 3 as being necessary to have a representative sample.

4.1.2 Summary of Findings

The following section reports the main findings of the survey questions in either graphical or tabular formats, and describes them in a narrative fashion.

4.1.2.1 Types of SNS users surveyed

The first question in the survey is: Which of the following Social Networking sites do you currently have an active account with and use? (Check all that apply). The purpose of having this question at the start was to disqualify any non-SNS users and to identify what SNSs the survey participant was currently using. The results disqualified 14 respondents who chose the 'none' option and were excluded from the survey. The results revealed that Snapchat was the dominant SNS among the four networks, with a response rate of 69.6%. Snapchat is the newest social network compared to the other three networks. Facebook, which is one of the oldest SNSs, had the lowest percentage of users in this survey at 55.9%. Table 4.1 presents the findings and the rankings of the SNSs by the survey participants.

Table 4. 1 SNSs selected by the users in the sample

Answer Choices	Responses
Facebook	55.90% N=*232*
Twitter	56.87% N=*236*
Instagram	60.96% N=*253*
Snapchat	69.64% N=*289*
None	3.37% N=*14*
Total Respondents: 415	

The results for this question also showed that there was a difference between male and females in the choice of SNSs. Figure 4.1 shows that a majority of males (75.52%) in this sample used Facebook; however, females used Facebook the least and Snapchat the most with 79.02%.

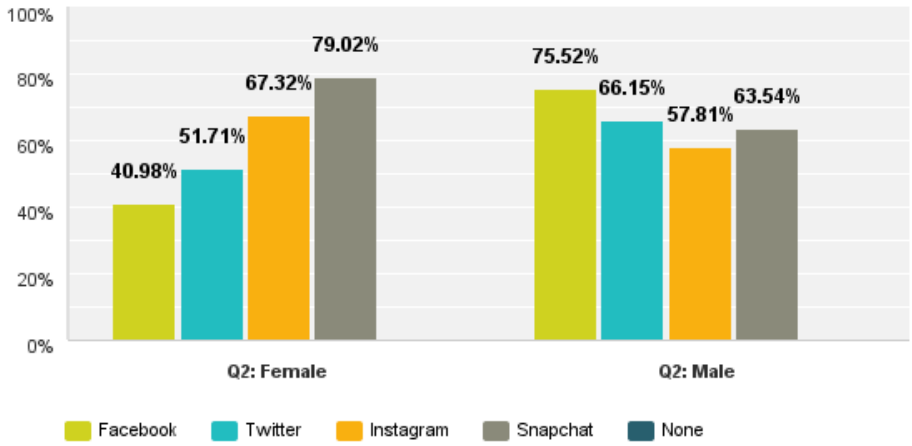

Figure 4. 1. Male vs. Female Choice of SNSs

4.1.2.2 Demographics of the research sample

The participants were asked to provide three items of demographic information about themselves. These demographics were used later to analyse the data in order to find differences in the answers within these three demographic groups. The first question related to gender, as seen in Figure 4.2, which shows that more females (N=204) responded to the survey than males (N=192).

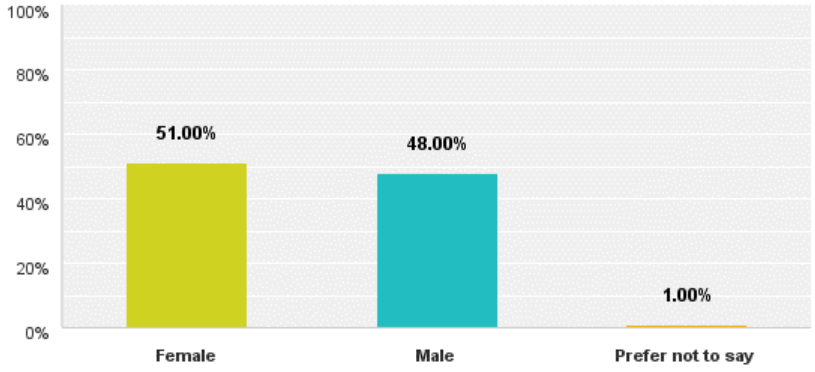

Figure 4. 2. Demographics of the sample: Gender

Secondly, age, as shown in Figure 4.3, was categorised into 8 groups. Due to ethical concerns, those under 16 years of age could not participate; hence, the age groups start at 16 years of age. Figure 4.3 shows that participants aged 20-24 years formed the dominant group in this survey at 35%.

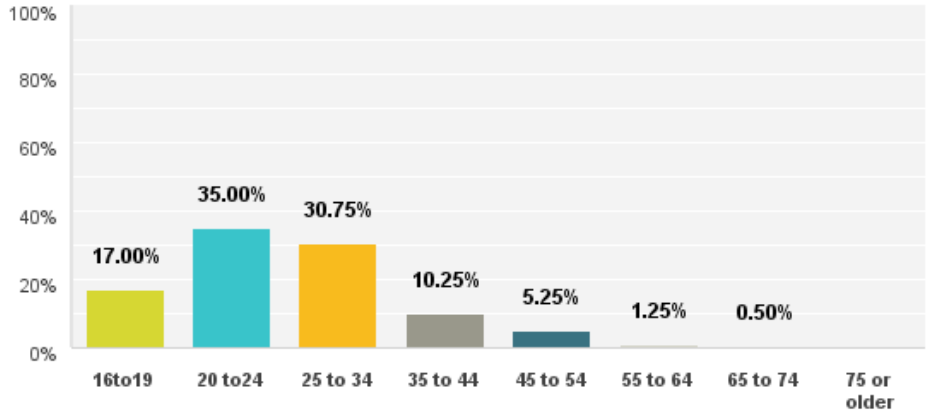

Figure 4. 3. Demographics of the sample: Age

The last demographic question that assisted understanding of the sample population was the education question, which indicated the educational background of the participant. Figure 4.4 shows that most of the survey sample had completed a bachelor's degree. The results are displayed in Figure 4.4 and Table 4.2.

Table 4. 2. Demographics of the sample: Education

Answer Choice	Responses
Less than High School	1.50% N=6
High School	14.75% N=59
Working towards completing a Bachelor's degree	11.50% 46
Bachelor's degree	43.25% N=173
Working towards completing a Master's degree	8.50% N=34
Master's degree	14.50% N=58
Working towards completing a Doctoral degree	2.75% N=11
Doctoral degree	3.25% N=13
Total Respondents=400	

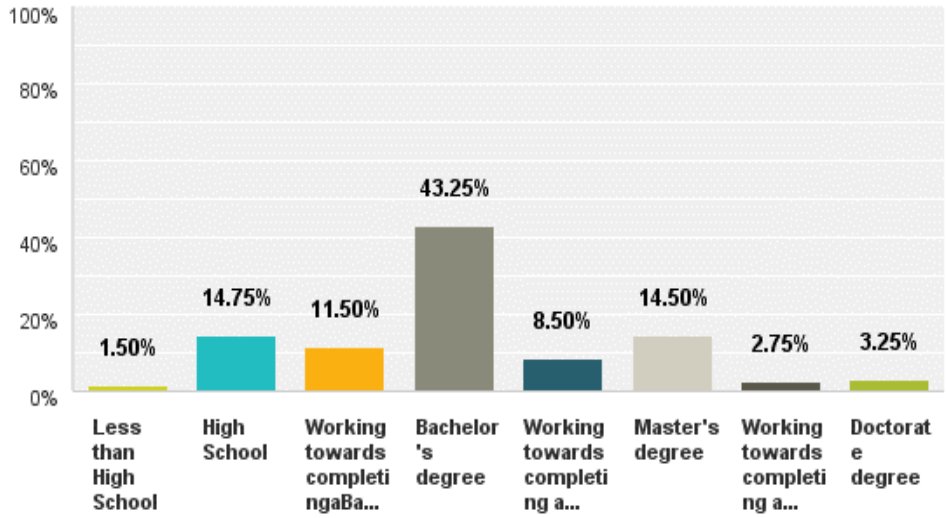

Figure 4. 4 Demographics of the sample: Education

4.1.2.3 Usage of SNSs and main privacy concerns

This section presents the findings for the third page of the survey, which investigated users' motives for using SNSs, frequency of use, and main privacy-related questions. The questions in this section aimed to identify users' SNS usage habits and their main privacy perceptions and were used as a comparison with later questions in order to see how users' privacy perceptions compared with their actual usage. In addition, these findings provided a better understanding of the sample, which allowed more effective qualitative analysis.

The first question addressed the main reason for using SNSs. Keeping in touch with family and friends (51%) and checking news and staying updated (35.6%) were the most prevalent reasons for using SNSs, as displayed in Figure 4.5.

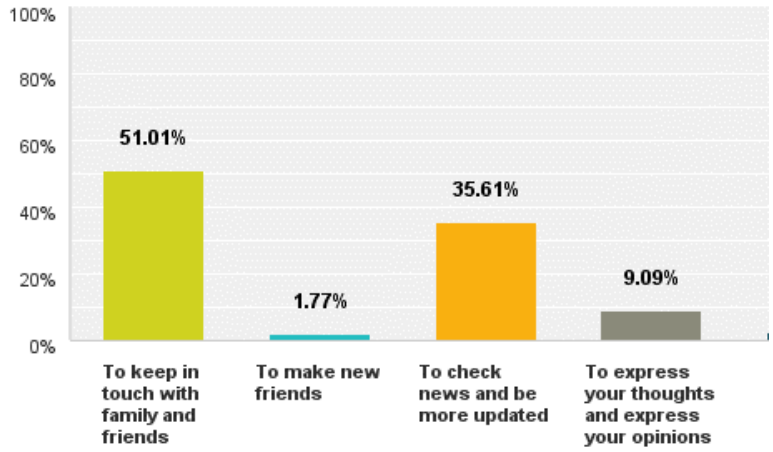

Figure 4. 5 Motives for using SNSs

Figure 4.6 displays the frequency of SNSs use by the survey participants. It shows that most of the sample were frequent users of SNSs, with 82.9% being daily users.

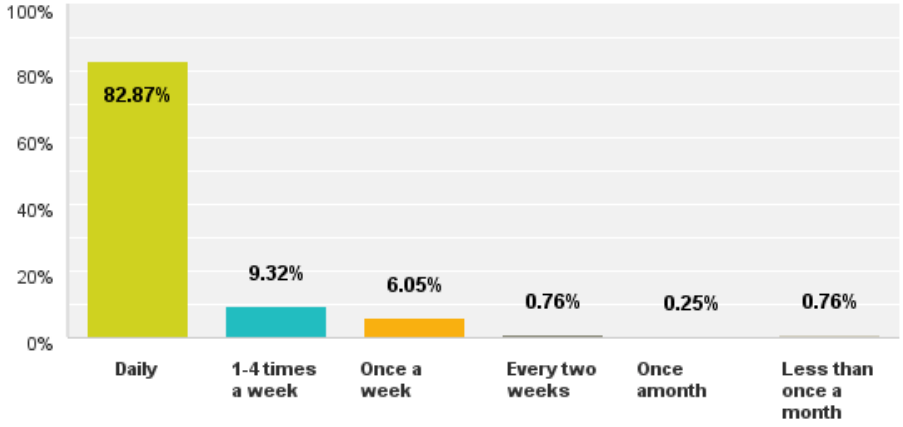

Figure 4. 6 Frequency of SNSs use

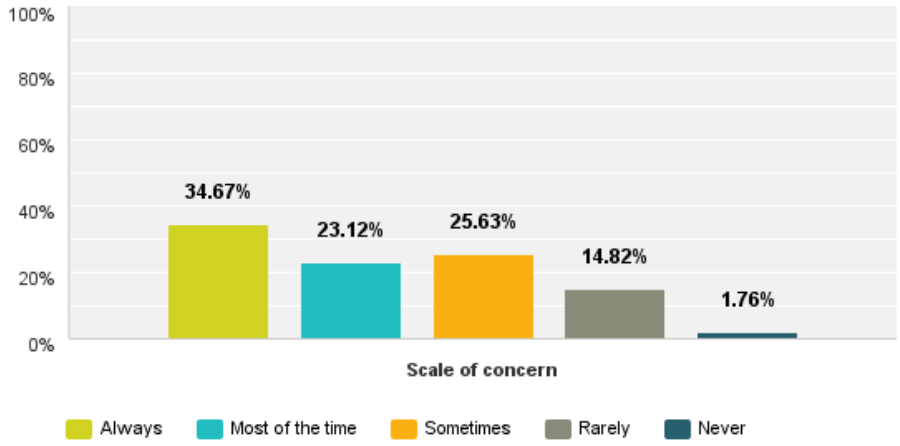

Figure 4. 7 User's privacy concern level

Figure 4.7 presents the findings of the question "Is the privacy of your information on Social Networking sites a major concern for you?" The purpose of this question was to establish the value of online privacy for the user, which can affect their answers to other questions. For instance, if someone is not very concerned about the privacy of their information online, they will probably not apply protective privacy and security settings to avoid leakage of information. In addition, people who value their privacy and are more concerned about their information will probably not share as much personal information compared to those who are less worried about privacy. This hypothesis is tested and discussed later by comparing the level of concern with the amount of information disclosure and privacy settings applied. The results of this question, as shown in Figure 4.7, indicate that the majority of the respondents were concerned about their privacy, although the degree of concern varied. In addition, the survey results showed that there was a lack of trust in SNS providers with regard to storage and protection of users' information, as 66.3% of the survey respondents answered that they did not trust their providers with their information. These findings are used later in this chapter to compare users' actual actions with their levels of personal information disclosure and examine the ways they apply privacy settings to protect their information and online identity.

4.1.2.4 Personal Information Disclosure and privacy settings:

SNSs are rich sources of personal information. Section 2.2 discussed the different types of personal information revealed in SNSs, the default privacy settings and whether it is obligatory for SNSs to provide them or not. Therefore, the following section will present the type of personal information disclosed by users and the tendency to disclose in each SNS. This section will also present the findings on user privacy settings.

4.1.2.4.1 Name used

This section begins with the findings on name disclosure. A username is a compulsory attribute that every user must provide in order to be able to sign up to an SNS. Figure 4.8 illustrates whether users disclosed their real, full name (first and last), part of their name, or a completely fake name in each social network chosen for this research. The results demonstrate that that vast majority of users tended to reveal their full, identifiable name. Facebook has the highest percentage of users sharing their full name at 88.9%. Snapchat users were the least likely to reveal their full name.

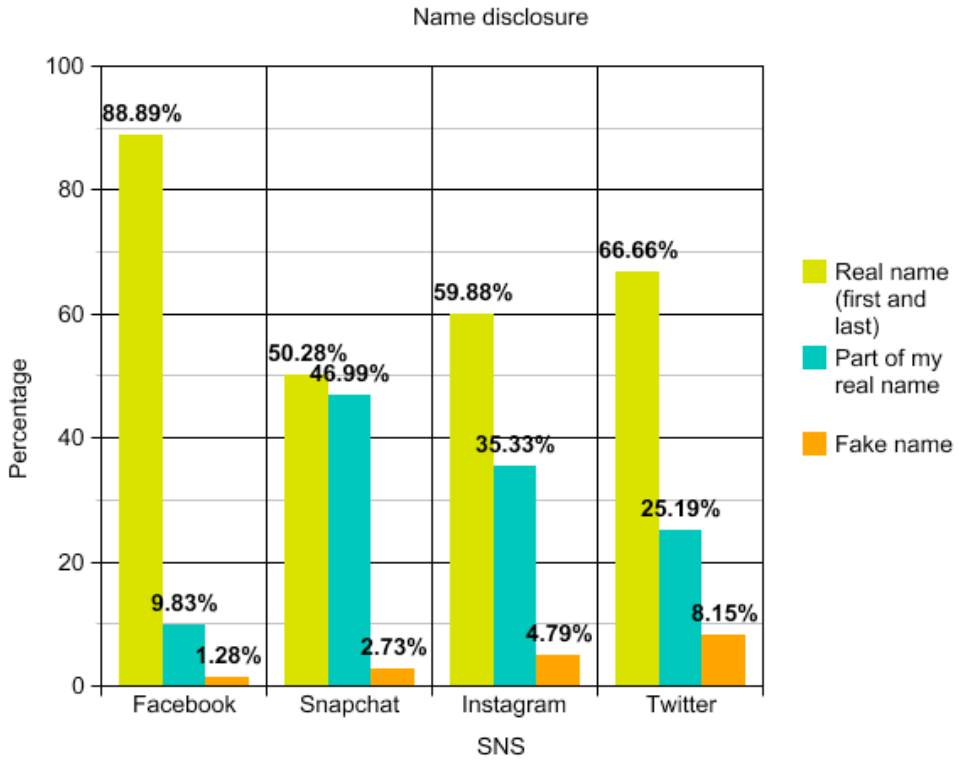

Figure 4. 8. Name disclosure in Facebook, Snapchat, Instagram and Twitter

4.1.2.4.2 Personal Profile information and privacy settings

As discussed in section 2.2, each SNS has its own functionalities; hence, the type of personal information required is different. Respondents in this survey were asked network-specific questions on the type of information they provided in their profile and the results are presented for each of the four SNSs investigated.

Facebook:

Out of the four SNSs studied in this research, Facebook can reveal the highest number of personal information characteristics such as profile picture, hometown, current city, education, date of birth, relationship status, friends lists, events, family members, locations visited, and contact information. The user has the freedom to set this information to be seen by the public, friends, customized groups of friends, or not share it at all. Table 4.3 shows what information users share and whether they share it publicly or privately. The results showed that all attributes were shared with either the public or friends. The majority of the sample displayed their hometown and current city to the public together with their full

name. The rest of the attributes were shared with friends. A small percentage of the sample shared their information with a customised group of friends. Customizing friends means changing privacy settings into a more restricted mode of display, in which the user filters their friends, usually based on how well they know them, adding them to customised groups that can only view certain areas of the user's posts. This option allows the user to have more privacy because the content they post cannot be viewed immediately by any friend they accept, but only by specific filtered users.

Table 4. 3 Facebook: Personal information disclosure and privacy settings

Personal Attribute	Public	Friends	Customised group of friends	I don't share this Information with others	Total
Hometown	53.9% 123	36.0% 82	3.9% 9	6.1% 14	228
Current city	52.8% 121	36.7% 84	4.8% 11	5.7% 13	229
Family members	24.9% 57	52.8% 121	7.4% 17	14.8% 34	229
Relationship status	29.3% 67	50.7% 116	4.8% 11	15.3% 35	229
Birthday	41.3% 95	47.8% 110	3.5% 8	7.4% 17	230
Education	44.5% 101	45.4% 103	3.5% 8	6.6% 15	227
Events	26.3% 60	59.2% 135	5.7% 13	8.8% 20	228
Locations visited (check in's)	24.3% 56	55.7% 128	4.3% 10	15.7% 36	230
Friends List	26.8% 61	54.8% 125	6.6% 15	11.8% 27	228
Contact Information (emails, address, phone number ...)	22.2% 51	50.0% 115	7.0% 16	20.9% 48	230

In addition to identifying what profile information users provide in Facebook, the survey investigated other aspects of users' privacy. Firstly, it was found that 74.4% of users have a picture of themselves in their profile picture. In Facebook, profile pictures are set to public by default and the user cannot adjust this. Secondly, the survey investigated whether they have enabled the 'Nearby Friends' feature. Not only does this feature enable the user to be informed about their friends' location, it can also allow the user to track their friends' locations in real time. In addition, enabling this feature gives Facebook the right to collect your precise location, which can be submitted to third parties for advertising purposes. The findings showed that 45.5% of the respondents said that they had this feature enabled. To get more insight into what type of data the user posts, the survey asked whether they had any pictures or videos of themselves available to the public (non-friends); 49.6% of the survey takers said that they did.

Snapchat

Snapchat displays information about the user in a different way from Facebook and other SNSs. It doesn't have a user personal information profile; it publishes what the user decides to send, which is only done by either a photo or a video that is available for viewing for a maximum of 10 seconds. Therefore, in order to get more insight about user privacy in Snapchat, the survey asked users about what kind of information they sent to their viewers. The results are presented in Figure 4.9. The results shows that users frequently sent personal content as 79.8% sent pictures or videos about themselves. Of the survey takers, 68.1% said that they also posted videos or photos of their family members. If users enable location services for Snapchat, they can include in their videos/pictures the name of their city, neighbourhood, or even street in the form of a geographical sticker. The survey results show that 76% of users included location information when posting pictures.

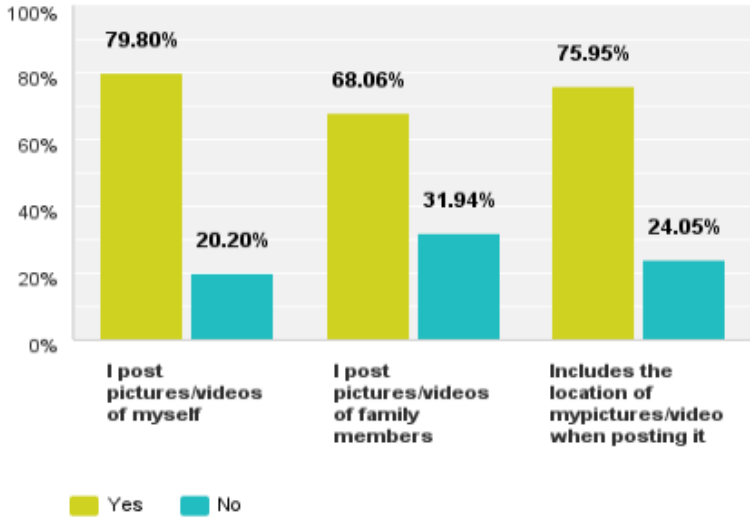

Figure 4. 9. Snapchat: Types of personal information posted

In terms of privacy settings, the survey showed that 48.3% of the survey participants had their Snapchat set for friends only and 31.8% had it customised, which is the most restricted setting.

Instagram

Similar to Snapchat and unlike Facebook, Instagram gives the user more freedom about what to post. Table 4.4 provides users' responses with regard to the type of information they provide. The first question asked whether the participant posted personal photos/videos of themselves. The result shows that the majority provided personal photos as well as photos of family and friends. As explained in Chapter 2, section 2.3, Instagram allows users to provide the geographical location of where the photo was taken by using EXIF data or by asking the user to enable GPS so that their current location is posted with the photo. In this survey, the participants were asked whether they had revealed their location to their followers by including the real location of the photo; 69.4% responded yes. The survey participants were also asked whether they had ever included the location of their home when they posted a photo or a video: 39.8% responded that they had. The questions in this section also asked if the survey participants had provided any other form of contact information in their public profile, and 59.6% responded that they had. Lastly, 60% of the respondents' profile pictures contained a picture of themselves; these pictures are public by default.

Table 4. 4. Instagram: Types of personal information posted

	Yes	No	Total
I post pictures/videos of myself	55.7% 151	44.3% 120	271
I post pictures/videos of family members/friends	58.7% 158	41.3% 111	269
I include the real location of my pictures/videos	69.4% 188	30.6% 83	271
Sometimes I post a photo with my house location in the map	39.8% 107	60.2% 162	269
I include contact information in my profile	59.6% 161	40.4% 109	270
Does your profile picture contain a picture of yourself?	60.0% 96	40.0% 64	160

With regards to account privacy settings, users can set their profiles as public or private. The results showed that there was an even split; 50% set it as public, and 50% set it as private.

Twitter

Twitter contains the fewest information categories for users to complete. In addition, unlike the other SNSs researched, it is not an application designed for photo or video sharing. It is mostly for written communications within a limited number of characters. Three questions were asked to determine the amount and type of information disclosed. Firstly, the respondents were asked whether they had a profile picture that contained a photo of them. The majority replied no (53.6%). The second question asked the respondents whether they indicated their location in their Twitter profile: the majority answered that they did not (51.9%). Lastly, the respondents were asked whether they had ever posted pictures of videos of themselves in their tweets, the vast majority answered that they had not (65.4%). The results revealed that 82.9% of Twitter users had their account set to public. Figure 4.10 provides a bar graph of the data.

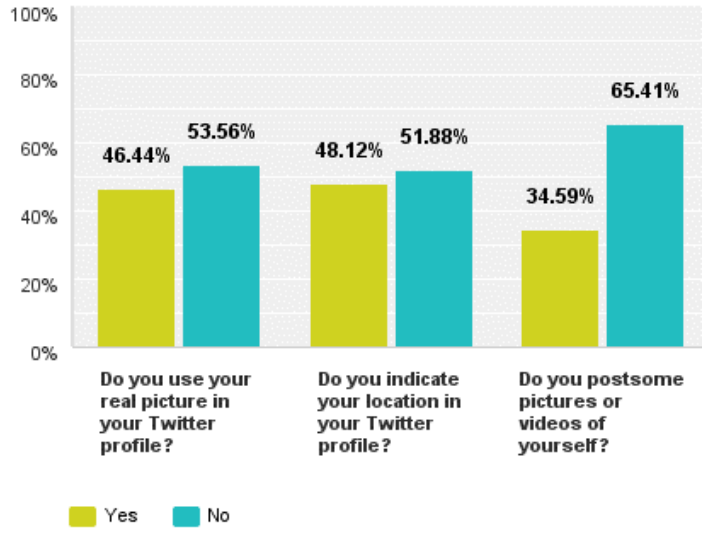

Figure 4. 10. Twitter: Type of personal information posted

4.1.2.4 Adding or accepting strangers into SNS profiles

For each SNS page in the survey, participants were asked whether they had people in their friends list that they did not know in real life. Figure 4.11 shows that there are differences between types of SNSs. Twitter users had the highest incidence of accepting/adding strangers (66%). In contrast, Snapchat users were reluctant to accept/add strangers, with only 36.9% of the Snapchat users in the survey saying that they had strangers in their friends list.

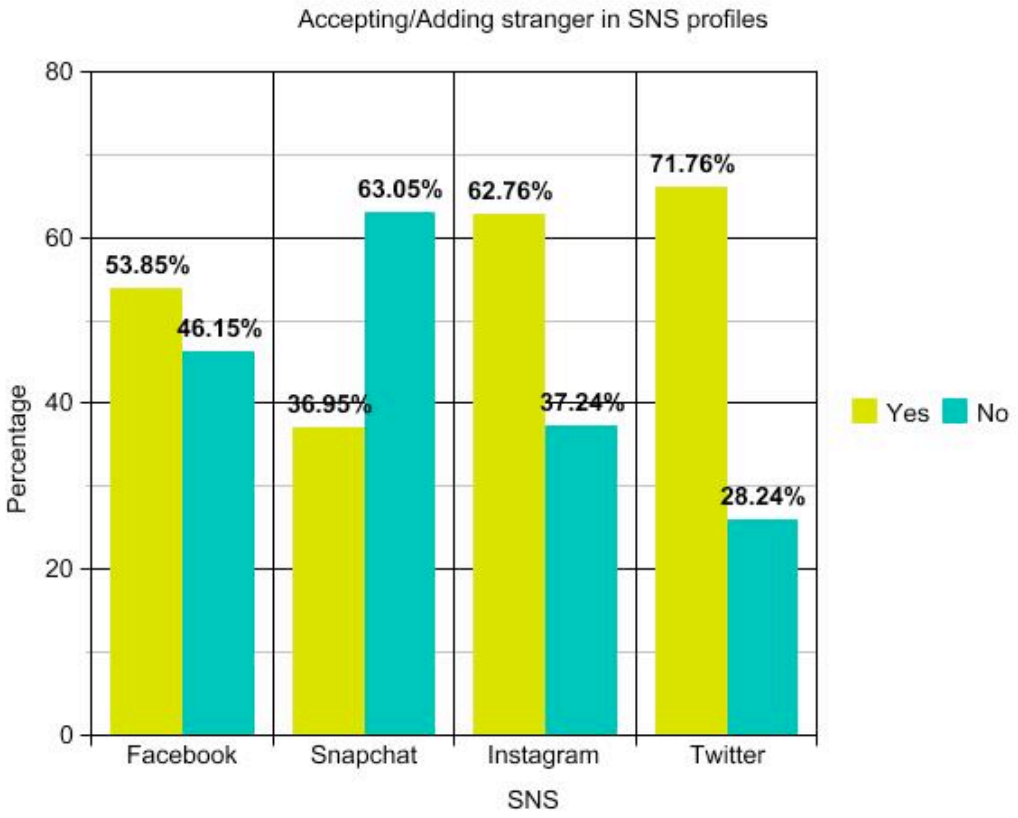

Figure 4. 11 SNS users' acceptance of other users they do not know personally.

.

4.1.2.5 SNSs users' privacy policy awareness

A significant part of this research focuses on SNSs' privacy policies and what they state in regards to how the SNS concerned handles users' information and what rights it legally has over users' information. This section analyses users' privacy awareness regarding how SNS providers/owners handle users' information. Section 2.6 intensively analysed what rights SNSs have over users' data and personal information. In this survey, we tested users' awareness and knowledge of what SNS providers can legally do with their data. The questions presented users with statements that were taken from the privacy policies of the four SNSs and asked whether they believed that the networks had the legal right to perform those actions, taking into consideration that they were listed in the terms and conditions when the users signed up. The results were analysed using frequency analysis in SPSS.

Four main statements were selected for Facebook, Snapchat, Instagram, and Twitter. All selected statements were true and were selected from each SNS's terms and conditions or

privacy policies tables. Users could tick whatever statements they believed the social network had the legal right to do, based on the terms and conditions of the network that the user had already agreed to. In addition, the user could check the "none of the above" box, which indicated that the user did not think that these statements were legally doable by the network. The results for each SNS are listed in Tables 4.5 to 4.8, displaying the statement and the responses of users in terms of whether they believed that the SNS had the right to do it. The results show that the majority of the survey participants did not agree that these statements were true.

Facebook privacy policy

Table 4. 5. Facebook privacy policy awareness question: response frequency

Statement	Proportion of survey participants who do not believe that this statement is true
Collect and use all the information they receive about you to suggest advertisements for you	78(33.6%)
Track your web surfing anytime you're logged into the site	43(18.5%)
Use your public information, such as your profile picture, in ads without asking you first and without any compensation to you	45(19.4%)
Collect information about your device locations, including specific geographic locations, through GPS, Bluetooth, or WiFi signals	59(25.4%)
None of the above	132(56.9%)

Snapchat privacy policy

Table 4. 6. Snapchat privacy policy awareness question: response frequency

Statement	Proportion of survey participants who do not believe that this statement is true
Collect Information about the pages you have visited before using Snapchat	65(22.73%)
Access and collect your device camera and photos (once you enable the option of saving your snaps to your photo Album)	84(29.37%)
Collect information from your device's phonebook including all your contacts (once you enable the option of adding friends from you address book)	86(30.07%)
Collect information about your precise location while using the application (once you enable location services for Snapchat)	103(36.01%)
None of the above	136(47.55%)

Instagram privacy policy

Table 4. 7. Instagram privacy policy awareness question: response frequency

Statement	Proportion of survey participants who do not believe that this statement is true
Use all your public information including pictures and posts for purpose such as advertising or to supply to third party providers	60(22.81%)
Keep hold of all your Information even when you delete a post or delete your account	58(22.05%)
Track your web surfing anytime you're logged into the site	56(21.29%)
Make your posts viewable to others even if you remove information you have posted to the Service through cached and archived pages of the Service	42(15.97%)
None of the above	167(63.50%)

Twitter privacy policy

Table 4. 8. Twitter privacy policy awareness questions: response frequency

Statement	Proportion of survey participants who do not believe that this statement is true
To use and post your public tweets anywhere, with no requirement to pay you for them	69(27.49%)
Collect "log data" from you including your IP address, browser type, operating system, the referring web page, pages visited, location, your mobile carrier, device information (including device and application IDs), search terms, and cookie information.	70(127.89%)
Use all your public information including pictures and posts to customise advertisements for you	65(25.90%)
Use all your public information including pictures and posts with Twitter partners and other third parties, including search engines, developers, and publishers that integrate Twitter content into their services, and insights without your permission.	58(23.11%)
None of the above	144(57.37%)

The results indicate that the majority of survey takers were not aware of how their information was handled, as illustrated by the tables above for each SNS. The survey results also show that 84.1% of the respondents did not read the privacy policies/terms of condition before signing up. The “None of the above” option for each SNS’s privacy policy statement was the highest percentage in each case. All the statements are correct, however, which indicates a lack of awareness among survey takers on how their information is handled by SNSs. The main reason given for not reading the information, as given by 73.1% of the respondents, was that it took too much time.

4.1.3 Statistical analysis: Gender, age, education, and privacy concern impacts on Information disclosure and privacy settings

Separate chi-square tests of contingencies were conducted in order to understand and determine the differences in user privacy setting behaviours and personal information disclosure variables with gender, age, education, and privacy rating for each of the four social networks. All chi-squares were interpreted at a conservative alpha of .01 to control for multiple tests. The Chi-square analysis helps to determine whether two discrete variables have any statistical association and whether there is a statistical significance between the variables.

The following sections summarise the findings from the statistical analysis. The full SPSS-generated tables are provided in the appendix. This section is divided into four sub-sections: Facebook, Snapchat, Instagram, and Twitter. The method for presenting the results differs for each SNS due to the differences between the questions in the survey. For instance, Facebook as a social network has the highest number of personal information categories and settings. Therefore, the analysis and presentation reflects the nature of the questions participants had to answer, which differed from the questions for the other SNSs.

Some alterations were made to the collected data in order to allow statistical analysis. Age, education, and privacy level groups were reduced and grouped together because some age groups contained very few participants. The original 8 age groups were reduced to three: 16-24 (representing the younger generation), 25-35, and 35+ years old. Educational level was represented by four groups: high school, bachelor's degree, master's degree, and doctorate. For levels of privacy concern, users were asked at the beginning of the survey how they felt about their SNSs' privacy and whether it was a concern to them. The question was formatted using a rating scale. Answers to this question were statistically analysed with reference to participants' other answers to determine if there was an association between the level of privacy were ranked in three groups: always, most of the time/sometimes, and rarely/never.

4.1.3.1 Facebook

The following section analyses Facebook data from four different perspectives: gender, age, education, and privacy concerns. The results were first analysed individually for each question using chi-square tests, then the answers to the questions were accumulated in order to compare the overall sharing of information and privacy settings. The higher the score, the more privacy settings were applied by the user, and the lower the score, the fewer privacy settings were applied. A t-test was conducted for gender as there are two variables, hence two means and two standard deviations. For age, educational level and privacy concerns, a one-way ANOVA test was conducted to compare privacy groups across total privacy scores.

Gender:

The results indicate that gender had a significant impact on users' privacy settings and information disclosure. The analysis showed that males were more likely to put their full real name than females, whereas females were more likely to put part of their real name. Both groups were very unlikely to sign up with a fake name. In addition, males were more likely to accept someone they did not know personally; have a public profile picture of themselves; have the nearby friends feature enabled; and have pictures or videos available publicly than females. Table 4.9 presents a summary of the analysis of the disclosure of personal information for males and females.

Table 4. 9. Facebook: Gender Chi-square description of results

Attribute	Results of cross -tabs and Chi Square analysis
Hometown	- Males more likely than females to be public - Females more likely than males to be friends - Females more likely than males to be "don't share" - No difference between gender on customized
Current City	- Males more likely than females to be public - Females more likely than males to be friends - Females more likely than males to be "don't share" - No difference between gender on customized
Family members	- Males more likely to be public than females - Females more likely to "not share' than males - No difference on friends or customized
Relationship status	- Males more likely to be public than females - Females more likely to "not share' than males - No difference on "friends" or customised
Birthday	- Males more likely than females to be public - Females more likely than males to be friends - No difference on customised or don't share
Education	- Males more likely than females to be public - Females more likely than males to be friends - Females more likely than males to "not share" - No difference on customized
Events	- Males more likely than females to be public - No difference on friends - Females more likely to be customized - Females more likely to "not share"
Locations visited	- Males more likely than females to be public - No difference on friends - No difference on customized - Females more likely to "not share"
Friends list	- Males more likely than females to be public - No difference on friends - Females more likely to be customized - Females more likely to "not share"
Contact information	- Males more likely than females to be public - No difference on friends - Females more likely to be customized - Females more likely to "not share"

T-test results:

An independent t-test was conducted to compare the means of the total privacy settings between males and females. The settings were ranked from a high score (more private settings) to a lower score (fewer private settings). The results of the t-test indicated that females had significantly higher privacy settings (M = 23.94, SD = 6.64) than males (M =

18.56, SD = 6.98), t(218) = 5.61, p < .001. Table 4.10 displays the results of the t-test.

Table 4. 10. Facebook total privacy score t-test results

Group Statistics

	gender	N	Mean	Std. Deviation	Std. Error Mean
privacy_sett_ total	female	81	23.9383	6.64332	.73815
	male	139	18.5612	6.98354	.59234

Age

The results of the chi-square analysis indicated that age did not have an effect on whether users signed up with their real name or a fake name. However, the results indicated that as users get older, they are less likely to accept friend requests without knowing them personally. In addition, users aged 16-24 were more likely to have a profile picture of themselves than those aged 25 to 34 or 35+. Additionally, those aged 16-24 were more likely to have a public picture of themselves than those aged 24-34 or those aged 35+. The 'Nearby friend' feature was less likely to be enabled by older users than younger ones. Finally, the 16-24 age group was more likely to have photo/videos made public than the 25-4 or 35+ age groups. Table 4.11 presents the summary of findings for individual chi-square tests for each personal information category.

Table 4. 11. Facebook: Age Chi-square description of results

Attribute	Results of cross -tabs and Chi Square analysis
Hometown	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - 25 to 34 and 35+ more likely to use customised than 16-24 - No effect on "I don't share"
Current City	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - 25 to 34 and 35+ more likely to have customised than 16-24 - No effect on "I don't share"
Family members	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - 25 to 34 and 35+ more likely to have customised than 16-24 - No effect on "I don't share"
Relationship status	- 16-24 most likely to have public, followed by 25 to 34 - 35+ mostly likely to have friends - No effect on "I don't share" or customised
Birthday	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - No effect on "I don't share" or customised
Education	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - No effect on "I don't share" or customised
Events	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - No effect on "I don't share" or customised
Locations visited	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - No effect on "I don't share" or customised
Friends list	- 16-24 most likely to be public, followed by 25 to 34 - 35+ mostly likely to be friends - No effect on "I don't share" or customised
Contact information	- 16-24 and 25 to 34 most likely to be public, - 35+ mostly likely to be friends - No effect on "I don't share" - 25 to 34 more likely to be customised

One-way ANOVA was used to compare age groups across total privacy scores. Results indicated a significant effect of age on privacy total, $F(2, 219) = 7.10$, $p = .001$. Those aged 25-34 ($M = 22.02$, $SD = 7.27$) and 35+ ($M = 23.23$, $SD = 6.58$) had significantly higher privacy settings than those aged 16-24 ($M = 18.89$, $SD = 7.76$). There was no difference between those aged 24-34 and those aged 35+.

Educational level impact

Chi square test results indicated that educational level did not have an impact on whether users signed up with their real name or not. However, it did have an effect on accepting the friend requests of people users didn't personally know: the results indicated that those with only high school qualifications were most likely to accept friends they didn't know, followed by bachelor's, then master's, and finally, by those with doctoral qualifications. The same pattern applied for activating the 'nearby friends' feature; those with high school qualifications only were most likely to have it activated, with a progressive decline from bachelor's to master's and doctoral level. Furthermore, users with only high school qualifications were more likely to have pictures or videos available to the public, followed by bachelor's, master's and doctoral qualifications. Table 4.12 presents a summary of the chi-square results for education.

Table 4. 12. Facebook: Education Chi-square description of results

Attribute	Results of cross -tabs and Chi Square analysis
Hometown	- High school most likely to be public - Doctoral most likely to be friends - No effect for customised or "I don't share"
Current City	- High school most likely to be public - Doctoral most likely to be friends - No effect for customised or "I don't share"
Family members	- High school most likely to be public - Doctoral most likely to be friends - No effect for customised or "I don't share"
Relationship status	- High school most likely to be public - Doctoral most likely to be friends and customised - No effect for "I don't share"
Birthday	- High school most likely to be public - Doctoral most likely to be friends - No effect for customised or "I don't share"
Education	- High school most likely tobe public - Doctoral most likely to be friends - No effect for customised or "I don't share"
Events	- High school most likely to be public - Doctoral most likely to be friends - No effect for customised or "I don't share"
Locations visited	- High school most likely to be public - Doctoral most likely to be friends and customised - No effect for "I don't share"
Friends list	- High school most likely to be public - Doctoral most likely to be friends - No effect for customised or "I don't share"
Contact information	- High school most likely to be public - Doctoral most likely to be friends - No effect for customised or "I don't share"

One-way ANOVA was used to compare education groups across total privacy scores Results indicated a significant effect of education on privacy total, $F(3, 218) = 6.19$, $p < .001$. The results indicated that the homogeneity of variance was violated, therefore, the results of Games-Howell post-hoc tests were used. Homogeneity of variance is an assumption of one-way ANOVA. It essentially tests whether the groups have the same or similar variance (homogeneous). The result obtained had a value of $p = .024$, indicating they were not equal; hence, the assumption was violated as the criteria had not been met by the data. The post-hoc evaluation controlled for the fact that there were 6 comparisons because there were 4 education levels, so dividing the alpha value of .05 by 6 produced a

value of .008, which was the new alpha level used when looking at the post hoc tests. The results indicated that high school students had significantly lower privacy scores than bachelor's or master's degree students (no significant difference between high school and doctoral, although high school was lower).

Privacy concern impacts

Individual chi-square tests indicated that those who said that their privacy was rarely/never a concern for them were more public and shared more, as follows: they were more likely to have nearby friends enabled, and to make publicly available pictures/videos of themselves and/or their friends; lists of their family members; relationship status; events they have attended; locations they have visited; friends lists; and contact information.

One-way ANOVA was used to compare privacy concern groups across total privacy scores for information shared and settings. Results indicated a significant effect of privacy concerns on privacy total, $F(3, 218) = 9.70$, $p < .001$. Homogeneity of variance was assumed so the Tukey test was used. There were 6 comparisons because there were 4 privacy levels, hence, dividing the alpha value of .05 by 6 produced a value of .008, which was the new alpha level used when interpreting the post hoc tests. Results indicated that those who rarely/never had privacy concerns had significantly lower privacy totals than the other three groups. The other three groups did not differ from one another.

4.1.3.2 Snapchat

Account privacy setting

The statistical analysis showed that gender, age, education, and privacy attitude had a statistically significant influence on Snapchat privacy settings, specifically on who could view the user's content. The results showed that males were more likely to have public settings than females. Females were more likely to have a customised list for viewing, which is the most private and restricted privacy setting. Snapchat users aged 16-24 years were more likely to have their Instagram story as a public listing. Those aged over 35 were the most likely to have a friends-only setting. Educational status also plays a part: high school students tended to have a public setting whereas those with bachelor's, master's or doctoral degrees were more likely to have friends-only settings. Finally, users who were

less concerned about their privacy were more likely to have their settings set to public in comparison to those who were more concerned about their privacy. Table 4.13 presents a summary of the individual Chi-square analysis results.

Table 4. 13. Snapchat quantitative Chi-square analysis results for account settings

		Public	Friends	Customised	X^2
Gender	**Male**	50(39.4%)	65(51.2%)	12(9.4%)	79.697*
	Female	7(4.4%)	72(45.6%)	79(50%)	
Age	**16-24**	42(26.3%)	69(43.1%)	49(30.6%)	13.765*
	25-34	13(13.7%)	47(49.5%)	35(36.8%)	
	35+	2(6.5%)	22(71.0%)	7(22.6%)	
Education	**High school**	27(52.9%)	16(31.4%)	8(15.7%)	46.349*
	Bachelor	22(14.6%)	74(49.0%)	55(36.4%)	
	Masters	8(11.6%)	37(53.6%)	24(34.8%)	
	Doctoral	0(0%)	11(73.3%)	4(26.7%)	
Privacy	**Rarely/Never**	26(52%)	18(36%)	6(12%)	43.257*
	Sometimes	6(8%)	43(57.3%)	26(34.7%)	
	Mostly	9(14.1%)	31(48.4%)	24(37.5%)	
	Always	16(16.5%)	46(47.4%)	35(36.1%)	

*p < .001, = P- value less than 0.001 indicate statistical significance

Type of name used on sign-up

For usernames, there was no statistical significance between males and females; both genders were likely to sign up with only part of their real names and both males and females were highly unlikely to provide a fake name. The results also revealed that high school students were most likely to sign up with their full real names. Degree holders of all levels were more likely to sign up with only part of their real names. Table 4.14 displays the results.

Table 4. 14. Snapchat quantitative Chi-square analysis results for name used

		Real name	Part of real name	Fake name	X^2
Gender	**Male**	56(56%)	41(41.0%)	3(3.0%)	2.902*
	Female	36(43.9%)	44(53.7%)	2(2.4%)	
Age	**16-24**	51(54.3%)	41(43.6%)	2(2.1%)	3.792
	25-34	30(48.4%)	29(46.8%)	3(4.8%)	
	35+	11(40.7%)	16(59.3%)	0(0%)	
Education	**High school**	28(90.3%)	3(9.7%)	0(0%)	24.975*
	Bachelor	40(41.2%)	54(55.7%)	3(3.1%)	
	Masters	20(41.7%)	26(54.2%)	2(4.2%)	
	Doctoral	4(57.1%)	3(42.9%)	0(0%)	
Privacy	**Rarely/Never**	25(68.8%)	12(31%)	1(2.6%)	8.873
	Sometimes	19(42.2%)	24(53.3%)	2(4.4%)	
	Mostly	16(48.5%)	15(45.5%)	2(6%)	
	Always	32(47.8%)	35(52.2%)	0	

*p < .001,

Accepting friend requests from other users the user does not know personally

The results also showed that there was a statistically significant relationship between gender, age, education, and privacy attitudes and respondents' answers to whether they accepted friend requests from strangers or people they did not know personally. The results showed that males, survey participants aged 16-24 years, high school students and survey participants who were rarely or never concerned about their privacy were more likely to accept friend requests from people they did not know personally. The results are summarised in Table 4.15

Table 4. 15. Snapchat quantitative Chi-square analysis results for accepting friend requests from unknown users

		Yes	No	X2
Gender	Male	74(54.8%)	61(45.2%)	33.673*
	Female	35(22%)	124(78.0%)	
Age	16-24	74(46%)	87(54%)	12.364*
	25-34	26(26.3%)	73(73%)	
	35+	9(25.7%)	26(74.3%)	
Education	High school	35(68.6%)	16(31.4%)	27.609*
	Bachelor	51(32.7%)	105(67.3%)	
	Masters	19(26.4%)	53(73.6%)	
	Doctoral	4(25%)	12(75%)	
Privacy	Rarely/Never	32(64%)	18(36%)	26.021*
	Sometimes	16(20.8%)	61(79.2%)	
	Mostly	29(42%)	40(58%)	
	Always	32(32%)	67(67.7%)	

*p < .001,

Posting personal pictures of the user

In terms of content posted, both males and females were likely to post pictures and videos of themselves on Snapchat. However, those aged 16-24 years and high school students were the most likely to post videos or pictures of themselves in Snapchat. Results are summarised in Table 4.16.

Table 4. 16. Snapchat quantitative Chi-square analysis results for posting personal pictures of the user

		Yes	No	X^2
Gender	Male	110(80.3%)	27(19.7%)	.0008
	Female	127(79.9%)	32(20.1%)	
Age	16-24	142(88.2%)	65(44.8%)	16.472*
	25-34	72(72%)	28(28%)	
	35+	23(63.9%)	13(36.1%)	
Education	High school	50(98%)	1(2%)	14.692*
	Bachelor	121(76.6%)	37(23.4%)	
	Masters	56(77.8%)	16(22.2%)	
	Doctoral	10(62.5%)	6(37.5%)	
Privacy	Rarely/Never	44(88%)	16(12%)	4.494
	Sometimes	64(81%)	15(19%)	
	Mostly	56(81.2%)	13(18.8%)	
	Always	73(73.7%)	26(26.3%)	

*p < .001,

Posting pictures of family

The results indicated that gender and age had a significant effect on posting pictures/videos of family or friends. The results showed that females and those aged 16- 24 and 25-34 were most likely to post pictures that included their family or friends. The results are summarised in Table 4.17

Table 4. 17. Snapchat quantitative Chi-square analysis results for posting pictures/videos that include family members/friends

		Yes	No	X^2
Gender	Male	74(56.9%)	56(43.1%)	14.187*
	Female	122(77.7%)	35(22.3%)	
Age	16-24	119(75.8%)	38(24.2%)	12.414*
	25-34	61(62.9%)	36(37.1%)	
	35+	16(47.1%)	18(59.9%)	
Education	High school	41(83.7%)	8(16.3%)	8.533
	Bachelor	104(68%)	49(32%)	
	Masters	42(59.2%)	29(40.8%)	
	Doctoral	9(60%)	6(40%)	
Privacy	Rarely/Never	42(84.0%)	8(16%)	8.025
	Sometimes	46(60.5%)	30(39.5%)	
	Mostly	44(67.7%)	21(32.3%)	
	Always	64(66%)	33(34%)	

*$p < .001$,

Including real location when posting pictures/videos

Whether the four variables had an effect on the inclusion of the geographical location of the user, a feature that is offered by Snapchat, was also investigated. The results indicated that only age had a statistically significant effect. Those aged 16-24 and 25-34 years had the highest likelihood of including their location. The results are presented in Table 4.18.

Table 4. 18. Snapchat quantitative Chi-square analysis results for including location information in pictures/videos

		Yes	No	X^2
Gender	Male	97(72.9%)	36(27.1%)	1.152
	Female	123(78.3%)	34(21.7%)	
Age	16-24	130(81.8%)	29(18.2%)	10.118
	25-34	71(73.2%)	26(26.8%)	
	35+	20(57.1%)	15(42.9%)	
Education	High school	45(90%)	5(10%)	23.18*
	Bachelor	111(75.5%)	42(27.5%)	
	Masters	54(75%)	18(25%)	
	Doctoral	11(68.8%)	5(31.3%)	
Privacy	Rarely/Never	41(83.7%)	8(16.3%)	3.570
	Sometimes	62(79.5%)	16(20.5%)	
	Mostly	49(72.1%)	19(27.9%)	
	Always	69(71.9%)	27(28.1%)	

*p < .001,

4.1.3.3 Instagram

Account privacy settings

The analysis indicated a significant effect of gender, education, and privacy attitude on privacy settings, with males more likely than females to use public settings. High school students were the most likely educational group to use public settings; and those who rated privacy as never/rarely a concern were the most likely to use public settings. There was, however, no effect of age on privacy settings. Results are displayed in Table 4.19.

Table 4. 19. Effect of gender, age, education and privacy concerns on the use of public settings in Instagram

		Public	Private	X^2
Gender	Male	89(75.4%)	29(24.6%)	53.89*
	Female	47(30.5%)	107(69.5%)	
Age	16-24	80(55.2%)	65(44.8%)	7.35
	25-34	35(38.5%)	56(61.5%)	
	35+	22(57.9%)	16(42.1%)	
Education	High school	37(77.1%)	11(22.9%)	23.18*
	Bachelor	60(41.7%)	84(58.3%)	
	Masters	37(54.4%)	31(45.6%)	
	Doctoral	3(21.4%)	11(78.6%)	
Privacy	Rarely/Never	33(75.0%)	11(25.50)	18.26*
	Sometimes	28(38.4%)	45(61.6%)	
	Mostly	26(40.0%)	39(60.0%)	
	Always	50(54.3%)	42(45.7%)	

*p < .001,

Accepting friend requests from other users the user does not know personally

The results indicated a significant effect of gender, education, age, and privacy attitudes on who survey participants accepted as followers (Table 4.21). Males were more likely to always accept, females were more likely to never accept and there was no difference in those males and females who sometimes accepted. Those aged 16-24 were more likely to always accept, those aged 25+ were more likely to never accept and there was no difference in those who sometimes accepted. For education, high school students were more likely to always accept. Finally, those who stated they were never/rarely concerned about privacy were more likely to always accept, those who rated sometimes/mostly/always were more likely to sometimes accept and there was no difference in those indicating 'never accept'. The results are summarised in Table 4.20

Table 4. 20. Effect of gender, age, education and privacy concerns on the acceptance of friend requests from unknown applicants on Instagram

		Always	Sometimes	Never	X^2
Gender	Male	42(38.9%)	44(40.7%)	22(20.4%)	42.91*
	Female	9(6.9%)	55(42.3%)	66(50.8%)	
Age	16-24	37(29.4%)	56(44.4%)	33(26.2%)	22.55*
	25-34	14(17.9%)	29(37.2%)	35(44.9%)	
	35+	0	14(40.0%)	21(60.0%)	
Education	High school	28(62.2%)	9(20.0%)	8(17.8%)	57.84*
	Bachelor	16(13.0%)	58(47.2%)	49(39.8%)	
	Masters	7(12.1%)	27(46.6%)	24(41.4%)	
	Doctoral	0	5(38.5%)	8(61.5%)	
Privacy	Rarely/Never	26(59.1%)	8(18.2%)	10(22.7)	53.98*
	Sometimes	3(5.5%)	30(54.5%)	22(40.0%)	
	Mostly	7(11.9%)	32(54.2%)	20(33.9%)	
	Always	15(18.5%)	29(35.8%)	37(45.7%)	

*$p < .001$

Type of name used on sign-up

With regard to the type of name used when signing up, age, gender, education and privacy attitudes had no effect, as shown in Table 4.21.

Table 4. 21. Effect of gender, age, education and privacy concerns on the type of name used on Instagram

		Real name	Part of the real name	Fake name	X^2
Gender	Male	62(70.5%)	22(25.0%)	4(4.5%)	8.99
	Female	37(48.1%)	36 (46.8%)	4(5.2%)	
Age	16-24	58(68.2%)	24(28.2%)	3(3.5%)	7.594
	25-34	31(57.4%)	20(37.0%)	3(5.6%)	
	35+	11(39.3%)	15(53.6%)	2(7.1.0%)	
Education	High school	27(87.1%)	4(12.9%)	0	12.553
	Bachelor	48(54.2%)	35(39.8%)	5(5.7%)	
	Masters	22(51.2%)	18 (41.9%)	3(7%)	
	Doctoral	3(60%)	2(40.0%)	0	
Privacy	Rarely/Never	29(87.9%)	3(9.1%)	1(3%)	15.31
	Sometimes	18(46.2%)	19(48.7%)	2(5.1%)	
	Mostly	19(57.6%)	12(36.42%)	2(6.1%)	
	Always	34(54.8%)	25(40.3%)	3(4.8%)	

Posting personal pictures of the user

Table 4.22 displays the analysis results indicating that gender and education have an effect on posting personal pictures of the user. Males and high school students are more likely to post personal pictures of themselves.

Table 4. 22. Effect of gender, age, education and privacy concerns on posting personal pictures of the user on Instagram

		Yes	No	X^2
Gender	Male	89(75.4%)	22(25.0%)	31.766*
	Female	62(41.1%)	36 (46.8%)	
Age	16-24	88(62.0%)	54(38.0%)	4.768
	25-34	44(49.4%)	45(50.6%)	
	35+	19(47.5%)	21(52.5%)	
Education	High school	37(78.7%)	10(21.3%)	15.213*
	Bachelor	79(54.9%)	65(45.1%)	
	Masters	30(45.5%)	36(54.5%)	
	Doctoral	5(35.7%)	5(64.3%)	
Privacy	Rarely/Never	34(79.1%)	9(20.9%)	11.422
	Sometimes	35(50%)	35(50%)	
	Mostly	33(50.8%)	32(49.2%)	
	Always	49(52.7%)	44(47.3%)	

Posting pictures of family

As shown in Table 4.23, only education had an effect on posting pictures or videos of family on Instagram; high school students tended to post family pictures more than the other educational groups.

Table 4. 23. Effect of gender, age, education and privacy concerns on posting pictures of family on Instagram

		Yes	No	X^2
Gender	Male	68(59.2%)	47(40.8%)	0.00
	Female	90(59.2%)	62 (40.8%)	
Age	16-24	91(63.6%)	52(36.4%)	4.173
	25-34	49(56.3%)	38(43.7%)	
	35+	18(46.2%)	21(53.8%)	
Education	High school	37(78.7%)	10(21.3%)	13.109*
	Bachelor	84(58.7%)	59(41.3%)	
	Masters	29(44.6%)	36(55.4%)	
	Doctoral	8(57.1%)	6(42.9%)	
Privacy	Rarely/Never	33(76.7%)	10(23.3%)	8.797*
	Sometimes	34(48.6%)	36(51.5%)	
	Mostly	38(59.4%)	26(40.6%)	
	Always	53(57.6%)	39(42.4%)	

Including real location when posting pictures/videos

Table 4.24 shows that age was the only variable that had an effect on posting pictures with locations attached, as Instagram users in the survey aged over 35 years were the least likely to include location information.

Table 4. 24. Effect of gender, age, education and privacy concerns on the inclusion of location in Instagram photos

		Yes	No	X^2
Gender	Male	85(72.0%)	33(28.0%)	.460
	Female	103(68.2%)	48 (31.8%)	
Age	16-24	106(74.1%)	37(25.9%)	10.731*
	25-34	63(71.6%)	25(28.4%)	
	35+	19(47.5%)	21(52.5%)	
Education	High school	37(78.7%)	10(21.3%)	3.786
	Bachelor	99(68.3%)	46(31.7%)	
	Masters	41(63.1%)	24(36.9%)	
	Doctoral	11(78.6%)	3(21.4%)	
Privacy	Rarely/Never	38(86.4%)	6(13.6%)	7.379*
	Sometimes	47(68.1%)	22(31.9%)	
	Mostly	43(66.2%)	22(33.8%)	
	Always	60(64.5%)	33(35.5%)	

Including house location when posting pictures/videos

The results indicate that all four variables; gender, age, education, and privacy attitude had an impact on whether Instagram users posted pictures or videos including their house location in the map. Males, high school students, and survey participants who were less concerned about their online privacy were more likely to include their house location when posing a picture. Participants who were over 35 years of age were the least likely to include this information. The results of the analysis are listed in Table 4.25.

Table 4. 25. Effect of gender, age, education and privacy concerns on inclusion of house location data in Instagram photos

		Yes	No	X^2
Gender	Male	65(55.6%)	52(44.4%)	21.870*
	Female	41(27.3%)	109(72.7%)	
Age	16-24	67(47.9%)	73(52.1%)	10.89*
	25-34	32(36.0%)	57(64.0%)	
	35+	8(20%)	32(80%)	
Education	High school	30(65.2%)	16(34.8%)	16.893*
	Bachelor	48(33.3%)	96(66.7%)	
	Masters	26(40%)	39(60%)	
	Doctoral	3(21.4%)	11(78.6%)	
Privacy	Rarely/Never	29(65.9%)	15(34.1%)	16.894*
	Sometimes	20(28.6%)	50(71.4%)	
	Mostly	22(34.9%)	41(65.1%)	
	Always	16(39.1%)	56(60.9%)	

Profile pictures that contain a personal picture of the user

The results of the analysis showed that only gender had an effect on having a profile picture that contained a personal picture of the user: males were more likely than females to do so, as shown in Table 4.26.

Table 4. 26. Effects of gender, age, education and privacy concerns on use of personal pictures in Instagram profiles

		Yes	No	X^2
Gender	Male	67(79.8%)	17 (20.2%)	28.842 *
	Female	28 (37.8%)	46 (62.2%)	
Age	16-24	55(66.3%)	28(33.7%)	3.014
	25-34	27(55.1%)	22(44.9%)	
	35+	14(50%)	14(50%)	
Education	High school	29(93.5%)	2(6.5%)	57.84*
	Bachelor	46(54.1%)	39(45.9%)	
	Masters	18(46.2%)	21(53.8%)	
	Doctoral	3(60%)	2(40%)	
Privacy	Rarely/Never	30(90.9%)	3(9.1%)	20.616*
	Sometimes	17(47.2%)	19(52.8%)	
	Mostly	13(46.6%)	19(59.4%)	
	Always	36(61%)	23(39%)	

4.1.3.4 Twitter

Privacy settings

The results that were gathered and analysed from the Twitter section of the survey are displayed in Table 4.27. The results indicate that there was no significant effect of gender, age, education, or privacy concerns on account privacy settings.

Table 4. 27. Effect of gender, age, education and privacy concerns on the use of public/private settings in Twitter

		Public	Private	X^2
Gender	Female	106(80.9%)	25(19.1%)	0.913
	Male	116(85.3%)	20(14.7%)	
Age	16-24	112(87.5%)	16(12.5%)	7.35
	25-34	72(77.4%)	21(22.6%)	
	35+	39(81.3%)	9(18.8%)	
Education	High school	41(93.2%)	3(6.8%)	4.753
	Bachelor	109(80.1%)	27(19.9%)	
	Masters	61(83.6%)	12(16.4%)	
	Doctoral	12(75%)	4(25%)	
Privacy	Rarely/Never	41(87.2%)	6(12.8%)	4.432
	Sometimes	67(87%)	10(13%)	
	Mostly	44(74.6%)	15(25.4%)	
	Always	71(82.6%)	15(17.4%)	

*$p < .001$,

Name used

The results revealed that gender was the only variable that had an effect on the sign-up name used, as shown in Table 4.28. Males were more likely to sign up with their real name, whereas females were more likely to sign up with only part of their real name.

Table 4. 28. Effect of gender, age, education and privacy concerns on the use of real or fake sign-up names on Twitter

		Real name	Part of the real name	Fake name	X^2
Gender	Female	74(55.6%)	48(36.1%)	11(8.3%)	17.907*
	Male	105(77.8%)	19(14.1%)	11(8.1%)	
Age	16-24	88(69.3%)	33(26%)	6(4.7%)	4.774
	25-34	62(66.0%)	23(24.5%)	9(9.6%)	
	35+	30(61.2%)	12(24.5%)	7(14.3%)	
Education	High school	31(72.1%)	9(20.9%)	3(7%)	2.351
	Bachelor	92(66.7%)	37(26.8%)	9(6.5%)	
	Masters	47(64.4%)	18(24.7%)	8(11%)	
	Doctoral	10(62.5%)	4(25%)	2(12.5%)	
Privacy	Rarely/Never	37(80.4%)	7(15.2%)	2(4.3%)	10.487
	Sometimes	44(57.1%)	27(35.1%)	6(7.8%)	
	Mostly	36(61%)	17(28.8%)	6(10.2%)	
	Always	63(71.6%)	17(19.3%)	8(9.1%)	

*$p < .001$,

Profile picture that contains a personal picture of the user

With regard to having a profile picture that contains a personal picture of the user in Twitter, the results indicated that there was a significant effect of gender only. The results of the analysis listed in table 4.29 show that males were more likely to have a real personal picture of themselves than females.

Table 4. 29. Effect of gender, age, education and privacy concerns on use of real profile pictures in Twitter

		Yes	No	X^2
Gender	Female	24(18.3%)	107(81.7%)	84.352*
	Male	100(74.6%)	34(25.4%)	
Age	16-24	64(50.4%)	63(49.6%)	1.521
	25-34	39(42.9%)	52(57.1%)	
	35+	21(42.9%)	28(57.1%)	
Education	High school	28(65.1%)	15(34.9%)	8.609
	Bachelor	54(39.7%)	82(60.3%)	
	Masters	34(47.2%)	38(52.8%)	
	Doctoral	8(50%)	8(50%)	
Privacy	Rarely/Never	30(65.2%)	16(34.8%)	8.172
	Sometimes	30(40%)	45(60%)	
	Mostly	26(44.1%)	33(55.9%)	
	Always	38(43.7%)	49(56.3%)	

*p < .001,

Including real location on profile

Table 4.30 shows that gender and age had a statistically significant effect on inclusion of the user's location in their Twitter profile, as males were more likely than females to include it. Older survey participants, over the age of 35 years, were less likely to include information about their location in their Twitter profile.

Table 4. 30. Effect of gender, age, education and privacy concerns on inclusion of location in Twitter profile

		Yes	No	X^2
Gender	Female	40(30.8%)	90(69.2%)	32.182*
	Male	88(65.7%)	46(34.3%)	
Age	16-24	70(55.1%)	57(44.9%)	10.013
	25-34	44(48.9%)	46(51.1%)	
	35+	14(28.6%)	35(71.4%)	
Education	High school	28(651%)	15(34.9%)	6.047
	Bachelor	60(44.1%)	76(55.9%)	
	Masters	33(46.5%)	38(53.5%)	
	Doctoral	7(43.8%)	9(56.3%)	
Privacy	Rarely/Never	31(67.4%)	15(32.6%)	8.808
	Sometimes	31(41.9%)	43(58.1%)	
	Mostly	25(42.4%)	34(57.6%)	
	Always	41(47.1%)	46(52.9%)	

*p < .001,

Posting public pictures and videos

Table 4.31 shows that gender, age, education, and privacy concerns all had an effect on whether Twitter users posted public pictures of themselves on Twitter. Males were more likely to have pictures or videos posted to their account. Similarly, users aged 16-24 years were more likely to have pictures or videos posted to their account. The educational status of the user also had an effect; high school students were more likely to have pictures or videos posted publicly than the other educational groups. Lastly, survey participants who stated that they rarely/never had concerns about their online privacy were more likely to have pictures or videos posted publicly on their accounts.

Table 4. 31. Effect of gender, age, education and privacy concerns on the inclusion of publicly accessible personal pictures on Twitter

		Yes	No	X^2
Gender	Female	19(14.6%)	111(85.4%)	46.181*
	Male	73(54.5%)	61(45.5%)	
Age	16-24	58(47.7%)	69(54.3%)	15.408*
	25-34	26(28.9%)	64(71.1%)	
	35+	8(16.3%)	41(83.7%)	
Education	High school	27(62.8%)	16(37.2%)	23.18*
	Bachelor	44(32.4%)	92(67.6%)	
	Masters	18(25.4%)	53(74.6%)	
	Doctoral	3(18.8%)	13(81.3%)	
Privacy	Rarely/Never	28(60.9%)	18(39.1)	18.832*
	Sometimes	19(25.7%)	44(74.6%)	
	Mostly	15(25.4%)	44(74.6%)	
	Always	30(34.5%)	57(65.5%)	

*$p < .001$,

4.2 SOCIAL EXPERIMENT RESULTS

The purpose of this experiment was to collect the responses of SNS users to a stranger friendship request to access their private accounts. Accepting a friend request from a private SNS profile immediately grants the request-sender the right to access, view, and communicate with the SNS user.

4.2.1 Experiment scenario and results:

A fake account was created for each SNS in this research: Facebook, Twitter, Instagram, and Snapchat. The accounts created had a fake name and no other information was provided in the account. The only thing that was recorded was whether the user accepted the friendship/follow request or not. No other information was collected for ethical reasons. Table 4.32 displays the results of this experiment. For each SNS, 400 different users were sent a friendship/following request. As discussed in section 3.5, the users were unknown to the researcher and they were added randomly using the random search function in the SNSs or by finding the users on public pages. After adding those SNS users' random accounts, their responses to the request were collected for a period of one month on a daily basis. The reason for collecting the responses daily and not waiting for a month is that some users could subsequently un-friend the fake account. Therefore, each accepted friendship was recorded instantly to provide accurate results.

As seen from the results in table 4.32, the majority of Facebook, Twitter and Instagram users did accept the friendship/follow request, whereas Snapchat users showed the lowest percentage of acceptance.

Table 4. 32. Acceptance rate for fake profiles on Snapchat, Facebook, Twitter and Instagram

	Users added	Users accepted the add	Frequency of acceptance
Snapchat	400	120	30%
Facebook	400	245	61.25%
Twitter	400	233	58%
Instagram	400	224	56%
Period: 1 month 1st March- 1st April 2016			

4.3 CONCLUSION

Chapter 4 has presented the findings of the applied research methodology reported in Chapter 3. The chapter was divided into two main sections: survey findings and social experiment findings. The chapter started with the survey findings, as the survey was the main source of data in this research. Firstly, the main findings of the survey questions were reported in either a graphical or tabular manner. Following that, the results of the statistical analysis were reported. In this section, findings from each of the four social networks were presented separately. The results showed that gender, age, and education had a great effect on most users' self-disclosure habits and privacy setting applications. The final part of this research presented the results of users' reaction to stranger's follow/friendship requests. The results showed that the majority of users accepted the request in Facebook, Twitter, and Instagram. More detailed explanation of the results is presented in the following chapter, which will also link the research findings to the research questions and related sub-questions.

Chapter 5: Discussion of Results

5.0 INTRODUCTION

In Chapter 4, the findings and the analysis of the survey and social experiment data were reported and presented according to the methodology presented in Chapter 3. Chapter 5 presents a comprehensive discussion of the results that were collected and analysed in Chapter 4 and evaluates their significance to the field of SNS privacy. The findings from Chapter 4 enable the research questions and related sub-questions to be answered in this chapter.

This chapter consist of four main sections. Section 5.1 answers the research questions and sub-sections introduced in Chapter 3. Section 5.2 discusses the significance of the results, guided by the literature review in Chapter 2. Based on the findings, section 5.3 presents best practices and recommendations for protecting user privacy and security when using SNSs. The final section, 5.4, presents the conclusions reached in this study.

5.1 ANSWERING THE RESEARCH QUESTIONS

This section provides a basis for answering the main research questions and related research sub-questions that were established in Chapter 3. This research answers three main questions that define the scope of this research. The first and primary research question has three related sub-questions and is answered systematically in section 5.1.1. The secondary research question is answered in section 5.1.2 and the final research question is addressed in section 5.1.3.

5.1.1 Primary research question and sub questions

Table 5. 1: Sub-Question 1 and Answer

Sub-Question (SQ1): **What is the influence of gender on SNS users' information disclosure and privacy settings for Facebook, Snapchat, Instagram and Twitter?**
Answer: **Facebook:** The results indicate that gender has a significant influence on all privacy settings and information disclosure. Males were more likely than females to share their full name, maintain a public profile picture of themselves, accept strangers into their private profile without knowing them personally, enable location-tracking features, post personal pictures and videos to the public and share personal information such as hometown, current city, family members, relationship status, birthday date, education, events, friends lists, contact information (email, phone number, websites) and locations visited (check-ins). The results of the statistical analysis showed that females had significantly higher privacy settings. **Snapchat:** The results indicate that gender has a significant influence on some users' privacy settings and information disclosure. The results showed that males were more likely to have public settings than females. Females were more likely to have a customised list for viewing, which is the most private and restricted privacy setting. In addition, males were more likely to accept friend requests from people they did not know personally. **Instagram:** The analysis indicated a significant effect of gender, with males more likely than females to use public settings, accept follow requests from users they didn't know personally, post personal pictures/videos of themselves, include their house location when posing a picture, and have a public profile picture using a personal picture. **Twitter**

The results indicate that gender had a significant influence on some users' privacy settings and information disclosure. Males were more likely to sign up with their real name, have a public profile picture that contained a personal picture, include their real location on their profile, and tweet personal videos/pictures from their account.

Table 5. 2: Sub-Question 2 and Answer

Sub-Question (SQ2): **What is the influence of age on SNS users' information disclosure and privacy settings for Facebook, Snapchat, Instagram and Twitter?**
Answer: **Facebook*:*** The results of the analysis showed that age influenced users' privacy settings and information disclosure habits. Younger users (16-24 years old) were more likely to have a public profile picture of themselves, accept strangers into their private profile without knowing them personally, enable location-tracking features, post personal pictures and videos to the public and share personal information such as hometown, current city, family members, relationship status, birthday date, education, events, friends lists, contact information (email, phone number, websites) and locations visited (check-ins) than those aged 25-34 and 35+. The results of the statistical analysis showed that those aged 25-34 and 35+ years old had significantly higher privacy settings than those aged 16-24. There was no difference between those aged 24-34 and those aged 35+. **Snapchat:** Snapchat users aged 16-24 years were more likely to have their Snapchat story as a public listing. Those aged over 35 were the most likely to have a friends-only setting. In addition, Snapchat users aged 16-24 years were more likely to accept friend requests from people they did not know personally, and to post videos or pictures of themselves in Snapchat. Those aged 16-24 and 25-34 were most likely to post pictures that included their family or friends and had the highest likelihood of including their location. **Instagram:**

The analysis indicated a significant effect of age, with users aged 16-24 years more likely to accept follow requests from users they didn't know personally, post personal pictures/videos of themselves, include their house location when posing a picture, and have a public profile picture that contained a personal picture of the user. Those aged 16-24 years tended to post pictures with their house location on the map. Instagram users aged over 35 years were the least likely to include location information when posting pictures or/and videos.

Twitter:

The results showed age had a statistically significant effect on inclusion of the user's location in their Twitter profile. Older survey participants, over the age of 35 years, were less likely to include information about their location in their Twitter profile. Age also had an effect on whether Twitter users posted public pictures of themselves on Twitter: users aged 16-24 years were more likely to have pictures or videos posted to their account.

Table 5. 3: Sub-Question 3 and Answer

Sub-Question (SQ3): **What is the influence of education on SNSs users' information disclosure and privacy settings for Facebook, Snapchat, Instagram and Twitter?**
Answer: **Facebook:** The results of the analysis showed that educational level influenced users' privacy settings and information disclosure habits. High school students tended to have a public profile picture of themselves, accept strangers into their private profile without knowing them personally, enable location-tracking features, post personal pictures and videos to the public and share personal information such as hometown, current city, family members, relationship status, birthday date, education, events, friends lists, contact information (email, phone number, websites) and locations visited (check-ins). This tendency progressively reduced with the possession of bachelor's, master's and doctoral qualifications. In addition, the results indicated that high school students have significantly lower privacy settings than bachelor's or master's students.

Snapchat:

High school students tended to have a public setting whereas those with bachelor's, master's or doctoral degrees were more likely to have friends-only settings. The results also revealed that high school students were most likely to sign up with their full real names. Degree holders of all levels were more likely to sign up with only part of their real names. High school students were more likely to accept friend requests from people they did not know personally, post videos or pictures of themselves, and include their location in their photos/videos.

Instagram:

Education had a statistically significant effect: high school students were the most likely educational group to use public settings, always accept followers they did not know personally, post personal pictures of themselves, post family pictures, include their house location when posing a picture, and have a public profile picture that contained a personal picture of the user.

Twitter:

Education had a slight effect on Twitter users' information disclosure and privacy settings. The results showed that high school students were more likely to have pictures or videos posted to their account than degree holders.

5.1.1.1 First main research question

What are the personal attributes that can have an influence on information disclosure and privacy settings of SNS users?

This research focused on the effect of three demographic characteristics: age, gender, and education. The purpose of this was to identify any trends or associations between these attributes and the degree of self-disclosure on SNSs. The answers to the three sub-questions established that these three attributes did have an effect on the amount of information users disclosed in SNSs and on their privacy settings.

Firstly, gender played an important role, as indicated by the statistically significant differences between the answers for males and females. Males showed more revealing and unconcerned behaviour than females, specifically in social networks that require disclosing

a vast amount of information about the user, such as Facebook. Their information was more available to the public and very identifiable. However, in social networks that are not exactly designed to reveal specific identifying information such as Twitter, males and females showed similar behaviour, with males being more likely to reveal information.

Age also played a significant part with regard to information disclosure and privacy settings. Younger people, aged 16-24 years, revealed more about themselves and applied less restrictive privacy settings, thus potentially exposing their personal identity online to the public, not just to their friends. The older users were, the more restrictive they apparently became about the information they posted and the privacy settings they applied.

Lastly, education had a significant effect on information disclosure and privacy settings. High school students were less concerned about their privacy online as they revealed personal identifying information such as full real name, personal photos and videos, and considerable information about their location. However, this kind of behaviour decreased with increasing levels of educational attainment.

Overall, males, younger users aged 16-24 years and/or high school students or holders of high school qualifications only disclosed more personal and identifying information and applied fewer privacy settings than females, users over the age of 25 years and/or holders of higher educational qualifications.

5.1.2 Second main research question

How do users' levels of privacy concern affect the amount of information they disclose in social networking sites?

This research studied how users perceive privacy compared to their actual actions on SNSs that might have an effect on their privacy. The research firstly identified the level of SNSs users' online privacy concerns, which were then compared with their answers to the information disclosure and privacy setting questions. The research had four levels of privacy concerns: always concerned, mostly concerned, sometimes concerned, and rarely/never concerned. The results of the analysis showed that the less users cared about their privacy, the more they shared identifying information and applied privacy settings that could threaten their security, such as enabling live-tracking features.

5.1.3 Third main research question

How aware are users of the extent to which their information is protected by SNS providers according to the privacy policies that the users have agreed to?

An important part of this research focused on analysing SNS privacy policies in Chapter 2. The outcome of that analysis indicated that SNSs could legally perform certain actions without prior approval from users, such as tracking user web browsing histories. Facebook, Snapchat, Instagram, and Twitter all state in their privacy policies that they monitor the web pages accessed by the user. Therefore, this research investigated users' awareness and knowledge of how their personal information is collected, stored, processed, and shared with other parties. The survey participants were presented with statements that were derived directly from each SNS's privacy policies and were asked whether they believed that those SNSs had the legal right to perform those actions on the user's data, based on the terms and condition that the users had agreed to prior to using the service. The results, which are listed in detail in Chapter 4, indicated that a statistically significant majority of users were not aware, as they stated that they did not believe the statements listed were accurate. The results indicate that the majority of survey takers were not aware of how their information was handled. The survey results also show that 84.1% of the respondents did not read the privacy policies/terms of condition before signing up

5.2 DISCUSSION OF THE RESULTS

The results presented in Chapter 4 indicated that SNS users share a significant amount of personal information. This section discusses the significance of these findings.

Findings from the survey revealed that the younger the users, the more frequently they used SNSs and the more information they revealed. In addition, the younger the users, the less they cared about privacy online; the results showed that 54% of 16-19 year-olds rarely cared about their privacy. A high proportion of Facebook users revealed personal information, which puts them at high risk of online crimes such as fraud, identity theft and stalking. The survey results revealed that users disclosed a significant amount of personal information; however, the amount differed between social networks. Facebook users shared the highest amount of personal information compared to the other SNSs in this research and

Twitter users shared the least. This is due to differences between SNSs in terms of content and purpose. For instance, Facebook has more categories for personal information whereas Twitter is mainly for posting news rather than specific categorised personal information; in other words, it is less personal than other SNSs. Tables 5.4, 5.5, 5.6, and 5.7 provide a summary of the type of information revealed and the degree of disclosure for each SNS studied.

Table 5. 4. Personal information disclosure in Facebook

Personal information revealed	Percentage of disclosure
Hometown	95.24%
Current city	94.32%
Education	93.79%
Birthday	92.61%
Friends lists	84.72%
Relationship status	88.15%
Events	91.23%
Family members	85.15%
Location visited	84.35%
Contact Information (emails, address, phone number)	79.13%
Public profile picture that contains a picture of the user	74.36%
Pictures or videos of yourself available to the public (non-friends)	49.57%
Real name	88.89%

Table 5. 5. Personal information disclosure in Snapchat

Personal information revealed	Percentage of disclosure
Real name	50.27%(Real first and last name) 46.99% (Part of the real name)
Post personal pictures/videos of themselves	79.80%
Post personal pictures/videos of family/friends	68.06%
Location information when posting pictures/videos	75.93%

Table 5.6. Personal information disclosure in Instagram

Personal information revealed	Percentage of disclosure
Real name	59.88%(Real first and last name) 35.33% (Part of the real name)
Post personal pictures/videos of themselves	55.72%
Post personal pictures/videos of family/friends	58.74%
Location information when posting pictures/videos	69.37(General location) 39.78%(House location)
Display public contact information on profile	59.63%
Public profile picture that contains a picture of the user	60%

Table 5. 6. Personal information disclosure in Twitter

Personal information revealed	Percentage of disclosure
Real name	66.67%(Real first and last name) 25.19% (Part of the real name)
Post personal pictures/videos of themselves	34.80%
Post personal pictures/videos of family/friends	33.21%
Location information on profile	48.12%
Public profile picture that contains a picture of the user	46.44%

According to Holm (2014), sharing personal details such as age, gender, full name, current city and other personal information can tremendously assist in establishing an identity. This is a weakness that identity thieves exploit to launch major identity theft attacks on SNS users, using the information that the user has already provided. This research has proven that accessing even private profiles is not difficult. The results from the experiment that was performed in the four networks showed that the majority of Facebook, Twitter, and Instagram users accepted total strangers' friend requests. This is alarming because it indicates that it is easy to access users' private information. This is a serious issue because identity theft crimes have the potential to reach anyone. Research conducted at Carnegie Melon University indicates that teenagers 15-18 years old are those most likely to be

victimized by identity criminals (Power, 2011). The results from this survey also revealed that the younger the user, the more they reveal. The enabler of these crimes is the availability of personal identification information, which increasingly has a measurable monetary value and an emotional cost as well (Al-Daraiseh, Al-Joudi, Al-Gahtani, & Al-Qahtani & 2014). Personally Identifiable Information (PII) is defined as information that can distinguish or trace an individual identity either alone or when linked with other information available about a specific individual (Krishnamurthy & Wills, 2009). Not only do users post PII, but their real locations are often very trackable. More than 45% of Facebook users have the nearby feature enabled. This feature enables Facebook to track your exact GPS location and inform other users who are nearby of your location. The issue with this feature is that not only friends, but other strangers the user has accepted might be informed of the user's real physical location. This can be an issue for younger users and children who could be at risk from child predators.

Overall, privacy is a subjective topic and depends on the user's perception of privacy and how much they value their privacy. In addition, privacy perceptions can vary from one culture to another. Chapter 3 reviewed a study that compared privacy concerns in two different countries: Hong Kong and France. The results showed that nationality had a significant effect, indicating that HK users tended to share more identifying information than French users. Further results, combined with users' privacy concerns, suggested that French SNSs users' higher privacy worries probably resulted in a lower level of disclosure of personal information. The results of the present research also indicated that the higher the privacy concern, the lower the self-disclosure. The second piece of research that was reviewed in Chapter 4 studied the degree of self-disclosure at different ages and between genders. The survey results revealed some serious privacy threats for SNSs users, particularly for minors and young people. The group most willing to disclose their personal information in that particular study were firstly men (81%) aged 18-25, followed by women (77%) in the same age group, which is similar to what the present research found.

5.3 SAFE PRACTICES AND SUGGESTIONS TO PROTECT USER PRIVACY

It is important to emphasise that privacy protection is mostly a personal responsibility. There is no doubt that the Internet has become a daily part of many lives. The results of this

research indicate that many of the younger generation share significant amounts of information about themselves over SNSs, not only with friends, but also with the public. In addition, the research also shows that many users are not aware of what they are agreeing to when they sign up with SNSs, or to what extent SNSs providers can legally use the information that the user knowingly and unknowingly provides. When visiting SNSs, whether to update a status or post a picture, users inevitably leave digital footprints that can invade privacy and make users vulnerable to identity theft. The good news is that there are some methods that can be applied to improve the privacy of user's information. These methods are technical and human-related aspects, which are described below:

- **Privacy protection from a technical aspect:**

One of the greatest security and privacy threats to users' data happens when the data is in transit, as it passes to and from a user's device. Connecting to the Internet happens via a massive network of undersea cables between countries, which means that our personal information can be intercepted in transit by those countries and by other service providers. That means that users' personal information is not only exposed to their friends (or the public based on their settings), but to many other parties. Internet service providers (ISP), SNSs owners, third party applications, and possibly also countries' governments have access to our information. One can argue that this cannot be an invasion of privacy because the user has willingly and knowingly shared their information via an SNS. The counter-argument here is that users post this information primarily for the purpose of connecting with friends and family, as has been shown in the findings of the survey. However, their information can be collected and used by other parties for other purposes, as discussed in Chapter 2. In addition, other parties collect more than just what the users post; they collect information about users' locations, their devices, what web pages they have visited, and many other digital footprints that users may not be aware of. Third party companies use and sell users' data for advertisement purposes without the knowledge of the user. In addition, ISPs can easily intercept connections between the user and the SNS, hence collecting all the information the user provides, when the user only intended that a small group of people would view it. Protecting users' privacy in SNSs depends on many factors. It depends on the privacy settings the users apply, and it also depends on the audience the user allows to view their information. Figure 5.1 illustrates the dangers of the traditional connection method, taking Snapchat as an example. When using Snapchat, many parties can intercept

and get hold of users' data, particularly if users are accessing SNSs via public Wi-Fi in airports, hotels, coffee shops or other public places. In addition, even secure Wi-Fi networks at home or work have the possibility of passing on the user's data before it reaches its final destination across a broadband modem and through in-between points on the Internet. Cellular networks, unless the user is individually targeted or under government censorship, are considered to be quite secure.

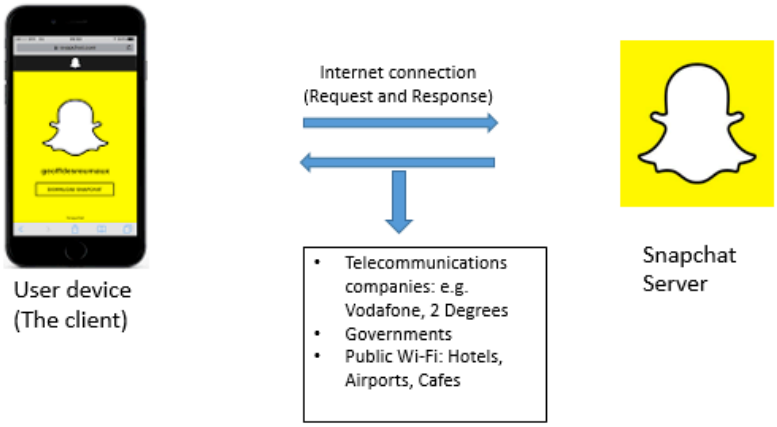

Figure 5. 1. Connection between the user device and Snapchat

There are some methods users can apply to reduce the amount of information that gets collected from them. One of the most effective methods for people who want to protect their privacy from unwanted interception or even protect their identity and location from SNSs providers who can collect browsing information, connections and location, is by connecting to the Internet via a virtual private network (VPN) (Fleishman, 2015).

A VPN enables the user to establish a secure connection via the Internet. It can be used for many purposes; for instance, employees wanting to connect to their work database outside of work. One of the best features of VPNs is that they provide a secure encrypted connection between the user's device, whether it is a smartphone or a computer, and the server (Fleishman, 2015). However, choosing the right VPN is crucial to ensure maximum benefit. There are some free VPN providers; however, these services are usually slow and may have bandwidth limits. In addition, there are hundreds of paid VPN services available online; however, these companies are profit-driven, meaning they will most likely store users' information and their browsing histories then sell it to advertisement companies or

even governments, which defeats the purpose of having a VPN. If users want to subscribe to these services, it is recommended they carefully read the VPN service's privacy policy (Fleishman, 2015). If the policy states that the VPN keeps a log of user data they should be avoided because this defeats the purpose of using a VPN. In addition, if the VPN service truly values its users' privacy, it will not record and collect user's browsing data. The third and most recommended option is creating a personal VPN. The benefit of creating a private and personal VPN is that the user can be assured that their private information remains private as the server is only accessed by the owner, who is the user. In addition, the server is secure and encrypted, which provides a high level of security. The user can have the server in their own house, but if the Internet is slow that can cause problems. A solution for this is to have a cloud-based sever. Digital-Ocean is considered one of the best websites for creating a personal and private cloud for beginners at an affordable price (Lee, 2014). The user can choose the physical location of the cloud server and can follow simple instructions to connect the VPN to the server and browse SNSs more privately and securely.

However, advertisers and third parties can still track users' browsing habits through the use of cookies and/or with other technologies. In order to avoid that, it is recommended that the user turns on private browsing and uses a VPN at the same time for maximum privacy (Diallo, 2014).

- **Privacy protection from a human aspect:**

Another aspect of protecting user privacy and security is increasing users' awareness of what information they provide in SNSs, particularly since this research has proven that many users accept friend requests from strangers. According to a UK financial services group, over 33% of SNSs users post information about when they're going to be away from their home (Consumer Reports, 2010). Posting such information can significantly threaten users' safety and security. Analysing privacy settings and privacy policies in Chapter 2 also indicated that most SNSs default settings are set to public and generally don't provide the security and privacy that users would want. Therefore, it is highly recommended that users maintain their privacy by updating their privacy settings and changing them from the SNSs default settings.

SNSs have become part of the lives of many people, specifically younger generations. They can have a great effect on their lives if used wrongfully. It is highly recommended to organise campaigns that target younger generations at school, to teach children and teenagers how to use SNSs properly.

5.4 CONCLUSION

Chapter 5 has provided a discussion of the findings presented in Chapter 4. Firstly, the three main research questions and related sub-questions were answered. The results showed a significant amount of personal identifying information was revealed by the users. The chapter provided more detail by identifying the factors that may have influenced such behaviour. It was apparent that age, gender, and education have a significant influence on users' behaviour. In addition, the chapter also discussed users' perceptions of privacy and whether these had any effect on their actions online. The chapter also discussed the results in terms of privacy policies and users' levels of awareness. After discussing the results, the chapter linked the results to the literature reviewed in Chapter 2 and briefly reflected on the consequences of sharing such information on users' safety. Lastly, the chapter suggested some practices that may provide users with more private SNS browsing experiences.

Chapter 6: Conclusion and Future Research

6.0 INTRODUCTION

Chapter 1 briefly introduced the research topic, discussed the motivation behind conducting the research, and outlined the thesis structure. SNSs have become part of daily life for billions of users. In SNSs, users build explicit networks that reflect their social relationships and usually share a wealth of personal information. The possible privacy risks of such behaviours are sometimes underestimated or ignored. The issue is exacerbated by a lack of knowledge and awareness in SNS users, as well as poorly designed tools for privacy protection on the part of the SNSs. Additionally, the centralized nature of SNSs makes the user dependent and puts the SNS provider in a position of power. In addition, privacy breaches can occur due to the lack of awareness of the complications of supplying identifying and trackable information. This research investigated the type of information users disclose, the level of protective privacy settings applied, and their knowledge of the rights of SNS providers in terms of handling their data, with the goal of providing recommendations to enhance users' privacy.

Chapter 2 reviewed the available literature on the topic of user privacy in SNSs, which provided a better understanding of the problems and consequences. It also analysed Facebook, Twitter, Instagram, and Snapchat privacy settings and security issues that can violate users' privacy, and privacy policies.

In Chapter 3, three relevant research methodologies were reviewed from three academic studies in order to establish the research methodology for this study. The three main research questions and three sub-questions were established. The research methodology consisted of an online survey as the main data collection method. The research questions were carefully designed in order to provide reliable data that would answer the research questions. The secondary data collection method was to test SNS users' reactions to friendship requests from strangers. Chapter 3 also presented the plan for data collection and analysis.

Chapter 4 reported the findings of the survey and the social experiment conducted by applying the methodology established in chapter 3. Statistical analysis was also conducted. Four attributes were used to determine any trends or associations in the data: gender, age, education, and level of privacy concern.

Chapter 5 firstly answered the sub-questions and then the three main research questions. The results were discussed and linked back to the literature review in Chapter 2. In addition, recommendations for enhancing users' privacy were outlined.

This chapter presents the conclusion of this thesis. Section 6.1 presents a summary of the research findings, which were previously presented in Chapter 4 and discussed in Chapter 5. Section 6.2 provides a summary of the limitations predicted and encountered during the research. Finally, potential future research areas and recommendations are outlined in section 6.3.

6.1 SUMMARY OF RESEARCH

This research has focused on the amount of self-disclosure of SNS users by identifying the type of information they reveal about themselves and their tendency to disclose personal information to the public. Firstly, the research studied the effect of three demographic factors on users' self-disclosure: gender, age, and educational status. The research also studied whether users' privacy concerns and perceptions had any influence on their behaviour on SNSs in terms of applying privacy settings and their revelation of personal information.

In order to collect the data and answer the research questions, two methods were used: an online survey and a social experiment. The focus of both the survey and the social experiment was on four different social networks: Facebook, Twitter, Snapchat and Instagram. Each of these SNSs has its own unique purpose and functionality, and displays different kinds of personal information, as explained in section 2.2. It was decided that data

would be collected for each one, rather than generalizing the survey questions and the social experiment to focus on all social networks as a single entity.

The results of the survey showed that there was a difference in information revelation between the four social networks. This is mainly due to the design of each social network. Social networks that encourage users to supply information about themselves, such as Facebook, had the highest level of information disclosure, whereas Twitter had the lowest amount of information disclosure compared to the other networks. This is because Twitter's main communication method is tweets, which are a maximum of 140 characters.

For Facebook, 88.89% of users revealed their full legal name; 53.85% said that they accepted strangers into their private profiles (although the experiment showed that 61.25% accepted strangers' friend requests); 74.36% had a public profile picture that contained a picture of themselves; 45.5% enabled location tracking services that notify their friends of their location; 59% had personal pictures and videos available to the public; and 61.11% posted pictures of family and friends. In addition, the vast majority of Facebook users displayed information about their hometown, current city, what school and university the studied in, their date of birth, their relationship status, lists of their friends, events they went to, family members' information, locations visited, and contact information.

Snapchat was the most used SNS out of the four studied. Users posted personal content on it: 79.80% posted personal pictures and videos of themselves; 68.06% posted pictures and videos of their family and friends; 75.95% included the geographical location of their pictures and videos; 70.93% posted pictures and videos of family/friend's children under the age of 16; and 50.27% revealed their full legal name, while 46.99% revealed part of their real name; and 36.95% said that they accepted strangers into their private profiles.

Instagram users also disclosed information that could comprise their privacy. Although 50% of the survey participants had private Instagram accounts, 62% of private account owners accepted strangers' follow requests. Of Instagram users surveyed, 59.88% used their full legal name; 55.72% posted personal content that included pictures of them; and 58.74% posted family pictures and videos. As for location disclosure, Instagram provides a function that allows the user to include a GPS location (a map) of where their picture was

taken. They use image EXIF data to determine it and the results showed that more than 69% of users included the real location of their images and videos when posting. In addition, it was alarming to find that 40% posted their house location publicly, while 59.63% included contact information in their public profile and 60% had a picture of themselves as their public profile picture.

Lastly, although 82.90% of Twitter users made their profile public and 66% had their real name in the profile, the type of the content they posted was far less personal. The majority did not use their real picture in their profile, include location information, or post personal pictures and videos of themselves, family or friends.

The research revealed that gender, age, and education had a significant influence on information disclosure and users' privacy settings. In general, males, young people between the ages of 16 and 24, and high school students displayed reckless and very identifying behaviour in SNSs that might actually comprise their privacy and in the worst cases, their safety, as they become more vulnerable to attacks from identity thieves and other malicious attackers.

Chapter 2 analysed the privacy policies of four social networks: Facebook, Twitter, Instagram and Twitter. The analysis revealed that social networks can collect information from users whether it has been directly provided or not. Examples of information provided by the user directly are username, profile picture, phone number, and user content that includes all posted information such as photos, videos, comments and other material. Examples of indirect sources of information are web browsing activities, logged information about the user's usage of the service, such as which IP was used to access the service, domain names, pages viewed, and similar information. In addition, SNSs such as Snapchat can access and collect users' contact lists on their phones, access their cameras and track their location, unless they disable those options. After analysing privacy policies, users' awareness of how their data is handled was measured. The results indicated that the majority of survey takers were not aware of how their information was handled. The survey results also showed that 84.1% of the respondents did not read the privacy policies/terms and conditions before signing up. The survey respondents were presented with four main statements selected from Facebook, Snapchat, Instagram, and Twitter. All selected

statements were factual and were selected from each SNS's terms and conditions or privacy policies. The vast majority of survey respondents stated that none of these statements were true and that SNSs did not use or access their data in these ways. This clearly indicated a lack of awareness among survey takers about how their information is handled by SNSs. Overall, the survey participants were mostly concerned about the privacy of their information and showed a lack of trust in SNS service providers' handling of their information. However, reviewing users' privacy settings showed that users were overly confident about their privacy protection perceptions because their actual privacy settings showed a lack of protection.

Overall, this research has performed an intensive analysis on users' privacy in SNSs. It has assessed the intensity of personal information sharing behaviours and performed an analysis of the factors that may have an influence on these behaviours. There is no doubt that many social networks offer a wide range of advantages to users. However, those benefits are not free of risks. Chapter 2 presented many severe consequences of privacy breaches that can comprise the safety of users. The fact that the majority of users underestimate those risks or are not aware of such risks is very concerning. Joining SNSs has becoming a phenomenon of modern existence, and both the young and the old are part of it. The lack of awareness that this research has revealed suggests the need for ongoing research in this field to provide better privacy protection mechanisms, both technically and socially.

6.2 LIMITATIONS OF THE RESEARCH

This research investigated the effects of gender, age, and education on information disclosure and privacy settings. However, there are other factors that can influence SNS users' information-sharing behaviour online that were not studied due to time and scope limitations. For instance, this research was conducted on users all over the globe; no specific cultural group or country was targeted. However, cultural beliefs or traditions may have an effect on users' behaviour on SNSs. For instance, in some conservative cultures, such as Middle Eastern cultures, there is a difference between male and female behaviour. For example, women might not display their pictures publicly whereas males may be more

open. This is considered to be a limitation of this research because race and culture can have a significance influence on SNS users' online behaviour.

Chapter 2 reviewed privacy issues and security holes that affect users' privacy on Facebook, Twitter, Snapchat and Instagram. Due to the limited availability of academic publications in this area, some of the literature that discussed privacy holes was from commercial and industrial sources. Nevertheless, every effort was made to ensure that the sources were as reputable as possible.

The experiment that was conducted to test SNS users' reactions to strangers' friend requests had some limitations because of ethical concerns. Firstly, the ethics committee did not approve of collecting demographic information about the users, such as their age or gender; it only approved collecting their responses to the friendship/follow requests. This was a limitation; factor analysis could not be conducted because no personal attributes were known. In addition, responses were collected over a one-month period. Although it is very rare that users do not check their accounts within a month, it is still possible. However, due to time limitations, if there was no response within a month of sending the request, it was recorded as an SNS user who did not accept strangers' requests.

Lastly, this research surveyed only four social networks due to scope and time limitations. Investigating more social networks and a larger SNS user audience could reveal deeper insights into user's privacy and personal information disclosure behaviours. Additionally, questions on whether users had been targeted into social engineering, or had been victims of identity theft may have been valuable in understanding the connection between information disclosure and privacy issues and their consequences.

6.3 FUTURE RESEARCH

This section discusses potential further research in the area of SNS privacy and security, which could provide great value and benefit to society. In this research, three personal demographics were studied as factors that could influence SNS usage. As discussed in the limitations section, culture can also be a significant factor. Future research could address

the differences between cultures and whether these have any effect on personal information disclosure.

Another way to develop this research in the future would be to study how disclosure of personal information in SNSs can enable identity theft. Chapter 2, section 2.3.3.4 discussed identity theft in relation to SNSs and the factors that can contribute to its occurrence. Future SNS privacy research could be conducted specifically in this field by analysing identity theft cases that have occurred involving SNS users. For instance, what is the common online behaviour of the victims that leads online criminals to steal those users' identities? What information do identity thieves want most from SNS users' profiles? The present research provided information about the type of information provided by SNS users and users' privacy settings. Future research could expand on this and examine the effect of users' sharing behaviours and their current privacy protection methods on the likelihood of their identities being stolen. By analysing previous cases of identity theft and comparing them to users' disclosure of information, protective methods could be developed and introduced to users to help eliminate identity theft cases that can cause great loss to victims.

The second significant area for future research in SNS privacy is to research location tracking in SNSs. Due to the advances in smartphone technology, location-tracking technology has become more sophisticated and can capture very precise contextual data such as users' movements, orientation, and location. The benefit of this development is location-based social networks. While location disclosure is crucial to enable many interesting location-based SNS features, it has many significant privacy implications. Since current GPS-equipped mobiles can detect accurate position information down to the metre level, sharing information about user location becomes even more problematic. For instance, when analysing a user's location history, it is possible to identify how many times a user went to the hospital in a year; hence, intruders may able to derive her health situation to some degree. In addition, if a user checks a site more frequently, it makes it easier to infer her home location as being near the location of the most common updates posted at night. Sharing information like this with friends is quite different from sharing it with the public. Future research can empirically study how users share location updates with the public. Obtaining a location-based SNS's API enables the collection of such data for analysis.

The final recommendation for future research is security in SNSs. The focus of this survey was mainly on what information users disclose and what privacy settings users apply that can compromise their privacy through their own lack of awareness. Future studies could investigate the technical security issues that comprise users' privacy.

REFERENCES

Al-Daraiseh, A. A., Al-Joudi, A. S., Al-Gahtani, H. B. & Al-Qahtani, M. S. (2014). Social networks' benefits, privacy, and identity theft: KSA case study. *International Journal of Advanced Computer Science and Applications*, *5*(12), 129-143. doi:10.14569/IJACSA.2014.051218

Aldhafferi, N., Watson, C., & Sajeev, A. (2013). Personal Information privacy settings of online social networks and their suitability for mobile internet devices. *International Journal of Security, Privacy and Trust Management*, *2*(2), 1-17. doi:10.5121/ijsptm.2013.2201

Alexa. (2016). The top 500 sites on the web. Retrieved from http://www.alexa.com/topsites

Altshuler, Y., Elovici, Y., Cremers, A.B., Aharony, N., & Pentland, A. (2013). *Security and privacy in social networks*. New York, NY: Springer.

Angwin, J., & Stecklow, S. (2010, October 12). 'Scrapers' dig deep for data on the web. *Wall Street Journal*. Retrieved from http://www.wsj.com/articles/SB10001424052748703358504575544381288117888

Ayoub, M. (2015). Snapchat privacy violation facts [Video file, in Arabic]. Retrieved from https://www.youtube.com/watch?v=rM7qOtB32Xo

Baden, R., Bender, A., Spring, N., Bhattacharjee, B., & Starin, D. (2009). Persona: an online social network with user-defined privacy. *Proceedings of the ACM SIGCOMM 2009 conference on Data communication - SIGCOMM '09*, 135-146. doi:10.1145/1592568.1592585

Barnes, S.B. (2006) A privacy paradox: Social networking in the United States. *First Monday*, *11*(9). doi:10.5210/fm.v11i9.1394. Retrieved from http://firstmonday.org/article/view/1394/1312

Beye, M., Jeckmans, A., Erkin, Z., Hartel, P., Lagendijk, R., & Tang, Q. (2010). Literature overview-privacy in online social networks. Centre for Telematics and Information Technology, University of Twente.

Bilge, L., Strufe, T., Balzarotti, D., & Kirda, E. (2009). All your contacts are belong to us: automated identity theft attacks on social networks. *Proceedings of the 18th international conference on World wide web* (pp. 551-560). New York, NY, USA: ACM.

Boyd, D.M., & Ellison, N.B. (2007). Social network sites: definition, history and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210-230. doi:10.1109/EMR.2010.5559139

Bryman, A. & Bell, E. (2011). *Business research methods* (3rd ed.). New York, NY: *Oxford University Press.*

Chapman, C. (2009, October 7). The history and evolution of social media. Retrieved from http://www.webdesignerdepot.com/2009/10/the-history-and-evolution-of-social-media/

Cheng, Y., Park, J., & Sandhu, R. (2013). Preserving user privacy from third-party applications in online social networks. *Proceedings of the 22nd International Conference on World Wide Web - WWW '13 Companion*. doi:10.1145/2487788.2488032

Chi, M. (2011) Security policy and social media use. Retrieved from https://www.sans.org/reading-room/whitepapers/policyissues/reducing-risks-social-media-organization-33749

Crandall, D. J., Backstrom, L., Cosley, D., Suri, S., Huttenlocher, D., & Kleinberg, J. M. (2010). Inferring social ties from geographic coincidences. *Proceedings of the National Academy of Sciences*, *107*(52), 22436–22441.

Cross, M. (2014). *Social media security: Leveraging social networking while mitigating risk*. Rockland, MA: Syngress (Elsevier Science).

Cutillo, L.A. (2014). *Security and privacy in online social networks* (Doctoral dissertation, Télécom ParisTech, Paris, France). Retrieved from https://tel.archives-ouvertes.fr/file/index/docid/932360/filename/these_Cutillo_V2.pdf

Cutillo, L.A., Molva, R., & Onen, M. (2011). Analysis of privacy in online social networks from the graph theory perspective. *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011* (pp. 1-5). Piscataway, NJ: IEEE. doi:10.1109/glocom.2011.6133517

Coyle, C.L., & Vaughn, H. (2008). Social networking: Communication revolution or evolution? *Bell Labs Technical Journal*, *13*(2), 13-17. doi:10.1002/bltj.20298

De Montjoye, Y. A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports*, *3*, Article No. 1376. doi:10.1038/srep01376.

Denning, D.E. (1999). *Information warfare and security*. New York, NY: ACM Press.

Diallo, A. (2014). Want privacy on the internet? Then you need a VPN. Forbes.com. Retrieved 1 September 2016, from http://www.forbes.com/sites/amadoudiallo/2014/03/07/want-privacy-on-the-internet-then-you-need-a-vpn/#717a13e99b73

Donath, J. (2007). Signals in social supernets. *Journal of Computer-Mediated Communication*, *13*(1), 231-251. doi:10.1111/j.1083-6101.2007.00394.x

Facebook. (2016). Data policy. Retrieved from https://www.facebook.com/full_data_use_policy

Felix, S. (2012, September 10). This is how Facebook is tracking your internet activity. *Business Insider*. Retrieved from http://www.businessinsider.com.au/this-is-how-facebook-is-tracking-your-internet-activity-2012-9?r=US&IR=T#it-started-off-as-just-a-normal-day-1

Fisher, D. (2015). Married woman left $41,000 out of pocket after online scam. *NZ Herald*. Retrieved from http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=11476193

Fleishman, G. (2015). Protect your data's final mile using a VPN. *Macworld - Digital Edition*, *32*(3), 99-104.

Gao, H., Hu, J., Huang, T., Wang, J., & Chen, Y. (2011). Security issues in online social networks. *IEEE Internet Computing*, *15*(4), 56-63. doi:10.1109/mic.2011.50

Ge, J., Peng, J., & Chen, Z. (2014). Your privacy information are leaking when you surfing on the social networks: A survey of the degree of online self-disclosure (DOSD). *2014 IEEE 13th International Conference on Cognitive Informatics and Cognitive Computing*, 329 - 336. doi:10.1109/icci-cc.2014.6921479

He, B.-Z., Chen, C.-M., Su, Y.-P., & Sun, H.-M. (2014). A defence scheme against Identity Theft Attack based on multiple social networks. *Expert Syst. Appl., 41*(5), 2345-2352. doi:10.1016/j.eswa.2013.09.032

Hendricks, D. (2013, May 8). The complete history of social media: then and now. Retrieved from http://smallbiztrends.com/2013/05/the-complete-history-of-social-media-infographic.html

Holm, E. (2014). Social networking and identity theft in the digital society. *The International Journal on Advances in Life Sciences*, *6*(3&4), 157-166.

Humphreys, L., Gill, P., & Krishnamurthy, B. (2013). Twitter: a content analysis of personal information. *Information, Communication & Society*, *17*(7), 843-857. doi:10.1080/1369118x.2013.848917

Identity Theft Resource Center. (2016). ITRC Fact Sheet 138: Social networking and identity theft. Retrieved from www.idtheftcenter.org/Fact-Sheets/fs-138.html

Instagram. (2013). Privacy policy. Retrieved from https://www.instagram.com/about/legal/privacy/

IPTC. (2016). Many social media sites still remove image rights information from photos. Retrieved from https://iptc.org/news/many-social-media-sites-still-remove-image-rights-information-from-photos/

Jones, C. (2013). Twitter says 250,000 accounts have been hacked in security breach. *The Guardian.* Retrieved from https://www.theguardian.com/technology/2013/feb/02/twitter-hacked-accounts-reset-security

Krishnamurthy, B., & Wills, C. E. (2009). On the leakage of personally identifiable information via online social networks. In *Proceedings of the 2nd ACM workshop on Online social networks* (pp. 7-12). ACM.

Kumar, R. P., Srikanth, C., & Sailaja, K. (2016). Location identification of the individual based on image metadata. *Procedia Computer Science, 85*, 451-454. International

Conference on Computational Modelling and Security (CMS 2016). doi: 10.1016/j.procs.2016.05.191.

Leavitt, T. (2014). Snapchat unveiled: An examination of Snapchat on Android devices. Retrieved from http://www.decipherforensics.com/snapchat/

Lee, J. (2014). DigitalOcean: The best VPS host for newbies. Retrieved from http://www.makeuseof.com/tag/digitalocean-the-best-vps-host-for-newbies/

Lewis, K. (2015). How social media networks facilitate identity theft and fraud. Retrieved from https://www.eonetwork.org/octane-magazine/special-features/social-media-networks-facilitate-identity-theft-fraud

Li, K., Lin, Z., & Wang, X. (2015). An empirical analysis of users' privacy disclosure behaviors on social network sites. *Information & Management*, *52*(7), 882-891. doi:10.1016/j.im.2015.07.006

Luo, W., Liu, J., Liu, J., & Fan, C. (2009). *An analysis of security in social networks.* Paper presented at the Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, 12-14 December, Chengdu. doi:10.1109/DASC.2009.100

Madrigal, A. C. (2012, March 1). Reading the privacy policies you encounter in a year would take 76 workdays. *The Atlantic.* Retrieved from http://www.theatlantic.com/technology/archive/2012/03/reading-the-privacy-policies-you-encounter-in-a-year-would-take-76-work-days/253851/

McKeon, M. (2010). The evolution of privacy on Facebook. Retrieved from mattmckeon.com/facebook-privacy/

Mclean, S., & Samavi, M. (2015). Data for the taking: Using website terms and conditions to combat web scraping. Retrieved from http://www.sociallyawareblog.com/2015/03/12/data-for-the-taking-using-website-terms-and-conditions-to-combat-web-scraping/#page=1

McKinsey Global Institute. (2012). *The social economy: Unlocking value and productivity through social technologies.* New York: McKinsey & Co.

Mislove, A., Marcon, M., Gummadi, P.K., Druschel, P., & Bhattacharjee, B. (2007). Measurement and analysis of online social networks. *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement* (pp. 29-42). New York, NY: ACM. doi:10.1145/1298306.1298311

Narayanan, A., & Shmatikov, V. (2009). *De-anonymizing social networks*. Presented at the 30th IEEE Symposium on Security and Privacy, Berkeley, CA. (pp. 173-187). Piscataway, NJ: IEEE. doi:10.1109/SP.2009.22

Online Security: Consumer Reports. (2010). Retrieved from http://www.consumerreports.org/cro/magazine-archive/2010/june/electronics-computers/social-insecurity/overview/index.htm

Pascual, A. (2014, February 5). 2014 Identity fraud report: card data breaches and inadequate consumer password habits fuel disturbing fraud trends. Retrieved from https://www.javelinstrategy.com/coverage-area/2014-identity-fraud-report-card-data-breaches-and-inadequate-consumer-password-habits#simple-table-of-contents-9

Patil, S., & Kobsa, A. (2009) Privacy considerations in awareness systems: designing with privacy in mind. In. P. Markopoulos, W. Mackay & B. Ruyter (Eds.) *Awareness systems: advances in theory, methodology and design.* Heidelberg, Germany: Springer.

Power, R. (2011). Child identity theft: new evidence indicates identity thieves are targeting children for unused social security numbers. Retrieved from https://www.cylab.cmu.edu/files/pdfs/reports/2011/child-identity-theft.pdf

Prall, L. (2010, September 20). SixDegrees - social networking in its infancy. Retrieved from http://ezinearticles.com/?SixDegrees---Social-Networking-In-Its-Infancy&id=5064109

Riederer, C. J., Zimmeck, S., Phanord, C., Chaintreau, A., & Bellovin, S. M. (2015). *I don't have a photograph, but you can have my footprints: revealing the demographics of location data.* In Proceedings of the Third ACM Conference on Online Social Networks (pp.185–195).

Riederer, C., Kim, Y., Chaintreau, A., Korula, N., & Lattanzi, S. (2016). *Linking users across domains with location data: theory and validation.* Paper presented at the Proceedings of the 25th International Conference on World Wide Web, Monterey, Canada.

Shin, D. (2010). The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with Computers*, *22*(5), 428-438. doi:10.1016/j.intcom.2010.05.001

Shu, C. (2013, December 31). Confirmed: Snapchat hack not a hoax, 4.6M usernames and numbers published. Retrieved from https://techcrunch.com/2013/12/31/hackers-claim-to-publish-list-of-4-6m-snapchat-usernames-and-numbers/

Singh, K., Bhola, S., & Lee, W. (2009, August). xBook: Redesigning privacy control in social networking platforms. In *USENIX Security Symposium* (pp. 249-266).

Snapchat. (2016). Snap Inc. Terms of Service. Retrieved from https://www.snapchat.com/terms

Statista. (2016). Number of social media users worldwide from 2010 to 2020 (in billions) | Statistics. Retrieved from https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Talib, S., Ismail, N., Olowolayemo, A., Naser, S., Haron, S., & Yusof, A. (2016). Social networks privacy policy awareness among undergraduate students: The case of Twitter. In *Information and Communication Technology for The Muslim World (ICT4M)* (pp. 1 - 5). Kuching: IEEE. doi: 10.1109/ICT4M.2014.7020674

Thomas, R.M. (2003). *Blending qualitative & quantitative research methods in theses and dissertations*. Thousand Oaks, CA: Corwin Press.

Torres, A.M. (2012). Social networking and online privacy: Facebook users' perceptions. *Irish Journal of Management*, *31*(2), 63-97. Retrieved from http://hdl.handle.net/10379/4059

Tosdr. (2012). Terms of services; Didn't read. Retrieved from https://tosdr.org/#twitter

Tubaro, P., Casilli, A. A., & Sarabi, Y. (2013). *Against the hypothesis of the end of privacy: An agent-based modelling approach to social media.* SpringerBriefs in Digital Spaces, doi: 10.1007/978-3-319-02456-1_1.

Tsoi, H.K., & Chen, L. (2011). From privacy concern to uses of social network sites: a cultural comparison via user survey. *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing*, (pp.457 – 464). doi:10.1109/passat/socialcom.2011.71

Tweney, D. (2016, May 26). Engagement to die for: Snapchat has 100M daily users, 65% of whom upload photos. Retrieved from http://venturebeat.com/2015/05/26/snapchat-has-100m-daily-users-65-of-whom-upload-photos/

Twitter. (2016). Twitter privacy policy. Retrieved from https://twitter.com/privacy?lang=en

Warner, T.L. (2014). Protect your online privacy by removing Exif data from your photos. Retrieved from http://www.quepublishing.com/articles/article.aspx?p=2216446

Warren, S., & Brandeis, L. (1890). The right to privacy. *Harvard Law Review 4*(5): 192-220.

White, D.S. (2013, February 9). Social media growth 2006 to 2012. Retrieved from http://dstevenwhite.com/2013/02/09/social-media-growth-2006-to-2012/

Yanofsky, D. (2015, January 12). Your private Instagrams weren't as private as you thought they were. Retrieved from http://qz.com/323307/instagram-privacy/

Young, D. (2013). Now you see it, now you don't, or do you: Snapchat's deceptive promotion of vanishing messages violates Federal Trade Commission Regulations. *Marshall Journal of Information Technology and Privacy Law*, *30*(4), 827-850.

Zheleva, E.M., Terzi, E., & Getoor, L. (2012). *Privacy in social networks*. San Rafael, CA: Morgan & Claypool.

Zuckerberg, M. (2006, September 8). An open letter from Mark Zuckerberg:. Retrieved from https://www.facebook.com/notes/facebook/an-open-letter-from-mark-zuckerberg/2208562130

APPENDICES

Appendix 1: Ethical approval document

Ethics Application: **15/429 Privacy and security concerns in social networking sites: Suggested safe practices.**

Thank you for submitting your application for ethical review. I am pleased to confirm that the Auckland University of Technology Ethics Committee (AUTEC) has approved your ethics application for three years until 7 December 2018.

As part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online through http://www.aut.ac.nz/researchethics. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 7 December 2018;
- A brief report on the status of the project using form EA3, which is available online through http://www.aut.ac.nz/researchethics. This report is to be submitted either when the approval expires on 7 December 2018 or on completion of the project;

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to obtain this. If your research is undertaken within a jurisdiction outside New Zealand, you will need to make the arrangements necessary to meet the legal and ethical requirements that apply there. To enable us to provide you with efficient service, we ask that you use the application number and study title in all correspondence with us. If you have any enquiries about this application, or anything else, please do contact us at ethics@aut.ac.nz.

All the very best with your research,

Kate O'Connor
Executive Secretary
Auckland University of Technology Ethics Committee

Appendix 2: Survey information sheet

Hello, my name is Mashael Aljohani. I am a master student, studying Masters of Information Security and Digital Forensics. This survey is part of my Masters' degree thesis at Auckland University of Technology and it studies security and privacy concerns when using Social Networking Sites. I would like to invite you to participate in my research study by completing this short survey. The survey should only take 10-15 minutes to complete. All survey responses will be anonymous and we will keep your answers confidential. The Auckland University of Technology Ethics Committee (AUTEC) has approved this study.

What will happen in this research?

You will be asked to answer several questions about how you use Social Networking Sites. The survey is only concerned about four Social Networking sites: Facebook, Instagram, Twitter, and Snapchat. If you use one or more of these networks, please complete the survey. You can navigate through each page by clicking the button Next.

What are the discomforts and risks?

There are no discomforts or risks. Participation is voluntary. You can stop being in the study at any time. You will not be penalized.

How will my privacy be protected?

Your survey responses will be anonymous. You will not be asked for personal or identifying information at any time.

What are the costs of participating in this research?

We expect the survey to take about 15 minutes of your time.

What opportunity do I have to consider this invitation?

The survey will be open until 15 of February 2016.

How do I agree to participate in this research?

Completion of this questionnaire will be taken as an indication of your consent to participate.

Will I receive feedback on the results of this research?

A summary of the findings can be found on the following link in July 2016 when the thesis is completed:

https://www.facebook.com/MasterThesisSummaryOfFindings/?skip_nax_wizard=true

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor:

Dr.Alastair Nisbet: alastair.nisbet@aut.ac.nz +64 9 921-9999 ext 5879

Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTEC, Kate O'Connor, ethics@aut.ac.nz , 921 9999 ext 6038.

Whom do I contact for further information about this research?

Researcher Contact Details:

Mashael Aljohani: xvf2283@autuni.ac.nz

Project Supervisor Contact Details:

Dr.Alastair Nisbet: alastair.nisbet@aut.ac.nz +64 9 921-9999 ext 5879

Appendix 3: Survey Questions

Page 1: Choice of SNSs

* 1. Which of the following Social Networking sites do you currently have an active account with and use? (Check all that apply)

- [ ] Facebook
- [ ] Twitter
- [ ] Instagram
- [ ] Snapchat
- [ ] None

Page 2: Demographics

* 2. What is your gender?

- ○ Female
- ○ Male
- ○ Prefer not to say

* 3. What is your age?

- ○ 16 to 19
- ○ 20 to 24
- ○ 25 to 34
- ○ 35 to 44
- 45 to 54
- ○ 55 to 65
- ○ ○ 64 to 74
- ○ 75 or older

* 4. What is your current level of education?

Page 3: General Questions about SNSs use and privacy perception

* 5. What is your main reason for joining Social Networking sites?

- ○ To keep in touch with family and friends
- ○ To make new friends
- ○ To check news and be more updated
- ○ To express your thoughts and express your opinions
- ○ Other (please specify)

* 6. On average, how many times do you check your Social Networks sites accounts?

- ○ Daily
- ○ 1- 4 times a week
- ○ Once a week
- ○ Every two weeks
- ○ Once a month
- ○ Less than once a month

* 7. Do you read the terms and conditions (Privacy policy) before creating a Social Networking profile?

- ○ Yes
- ○ No

8. If you answered the pervious question with "No", what are the main reasons behind you choosing not to read the terms and conditions of Social Networking sites before signing up?

- ☐ It takes too much time
- ☐ I don't understand legal terms
- ☐ I don't care about the terms and condition
- ☐ Other (please specify)

* 9. Is the privacy of your information on Social Networking sites a major concern for you?

Always	Most of the time	Sometimes	Rarely	Never
○	○	○	○	○

* 10. Do you trust Social Networks service providers with storing your Information?

○ Yes

○ No

11. I am confident that my current account privacy settings on Social Networking sites will protect my privacy and security

	Strongly agree	Somewhat agree	Neither agree/disagree	Somewhat Disagree	Strongly disagree	I don't use this website
Facebook	○	○	○	○	○	○
Twitter						
Snapchat	○	○	○	○	○	○
Instagram						

Page 4: Facebook Questions

If you actively use Facebook, please answer the following questions. If you do not use Facebook, please *SKIP* to the next page

12. How many people do you have on your friends list on Facebook?

13. How many of those friends on Facebook have you met in person?

14. What name do you use in your Facebook profile?

○ Real name (first and last)

○ Part of my real name

○ Fake name

15. Please answer the following

	Yes	No
Have you ever accepted a friend request without knowing for sure that you knew them personally?	◯	◯
Does your Facebook profile picture contain a picture of yourself?	◯	◯
Is "Nearby Friends" feature enabled in your Facebook?	◯	◯
Do you have some pictures or videos of yourself available to the public(non-friends)?		◯
◯	◯	◯

Have you ever posted pictures/videos of family or friends children under the age of 16

16. Please review your privacy settings and choose your display preference for each of the following types of information on your Facebook account

	Public	Friends	Customised group of friends	I don't share this Information with others
Profile picture	◯	◯	◯	◯
Hometown	◯	◯	◯	◯
Current city	◯	◯	◯	◯
Family members	◯	◯	◯	◯
Relationship status	◯	◯	◯	◯
Birthday	◯	◯	◯	◯
Education	◯	◯	◯	◯
Events	◯	◯	◯	◯
Locations visited (check in's)	◯	◯	◯	◯
Friends List	◯	◯	◯	◯
Contact Information (emails, address, phone number ...)	◯	◯	◯	◯

17. Out of the following, please check what do you believe Facebook has the right to do based on their terms and conditions

- [ ] Collect and use all the information they receive about you to suggest advertisement for you Analyse
- [ ] your posts including the posts that you have typed but end up deciding not to post Track your web
- [ ] surfing anytime you're logged in into the site
- [ ] Use your public information such as profile picture in Ads without asking you first and without any compensation to you
- [ ] Collect information about your device locations, including specific geographic locations, such as through GPS, Bluetooth, or WiFi signals
- [ ] None of the above

Page 4: Snapchat Questions

If you actively use Snapchat, please answer the following questions. If you do not use Snapchat, please *SKIP* to the next page

18. What is your default setting for who can

	Everyone	Friends	Custom
View your Story	☐	☐	☐

19. What name do you use in your Snapchat account?

- ☐ Real name (first and last)
- ☐ Part of my real name
- ☐ Fake name

20. Have you ever accepted a friend request without knowing who they are for sure?

- ☐ Yes
- ◯ No

21. Please answer the following

	Yes	No
I post pictures/videos of myself	◯	◯
I post pictures/videos of family members	☐	☐
I sometimes include the location of my pictures/video when posting it	◯	◯
I post pictures/videos of family or friends children under the age of 16	☐	☐

22. Out of the following, please check what do you believe Snapchat has the legal right to do based on their terms and conditions that you have agreed on before signing up

- ☐ Collect Information about the pages you have visited before using Snapchat
- ☐ Access and collect your device camera and photos (Once you enable the option of saving your snaps to your photo Album) Collect
- ☐ information from your device's phonebook including all your contacts (Once you enable the option of adding friends from

you address book)

☐ Collect information about your precise location while using the application (Once you enable location services for Snapchat) None of the above

Page 4: Instagram Questions

If you actively use Instagram, please answer the following questions. If you do not use Instagram, please *SKIP* to the next page

23. Your Instagram account is

- ○ Public
- ○ Private

24. For private account users, Do you accept followers request from people you do not know in real life?

- ○ Always
- ○ Sometimes
- ○ Never

25. What name do you use in your Instagram profile?

- ○ Real name (first and last)
- ○ Part of my real name
- ○ Fake name

26. Please answer the following

	Yes	No
I post pictures/videos of myself	○	○
I post pictures/videos of family members	○	○
I sometimes include the real location of my pictures/video when posting it	○	○

I once or more have posted a photo with my house location in the map	◯	◯
I post pictures/videos of family or friends children under the age of 16	◯	◯
Does your profile picture contain a picture of yourself?	◯	◯

27. Out of the following, please check what do you believe Instagram has the legal right to do based on their terms and conditions that you have agreed on before signing up

- ☐ Use all your public information including pictures and posts for purpose such as advertising or with other third party providers
- ☐ Keep hold of all your Information even when you delete a post or delete your account
- ☐ Track your web surfing anytime you're logged in into the site
- ☐ Make your posts viewable to others even if you remove information you have posted to the Service through cached and archived pages of the Service
- ☐ None of the above

Page 4: Twitter Questions

28. My twitter account is

- ◯ Public
- ◯ Private

29. What name do you use in you Twitter profile?

- ◯ Real name (first and last)
- ◯ Part of my real name
- ◯ Fake name

30. Please answer the following

	Yes	No
Do you use your real picture in your Twitter profile?	◯	◯

Do you indicate your location in your Twitter profile?

Do you post some pictures or videos of yourself?

I post pictures/videos of family or friends children under the age of 16

31. Out of the following, please check what do you believe Twitter has the right to do based on their terms and conditions

- [ ] To use and post your public tweets anywhere, and don't have to pay you for them
- [ ] Collect "log data" from you including IP address, browser type, operating system, the referring web page, pages visited, location, your mobile carrier, device information (including device and application IDs), search terms, and cookie information.
- [ ] Use all your public information including pictures and posts to customise advertisement for you
- [ ] Use all your public information including pictures and posts with Twitter partners and other third parties, including search engines, developers, and publishers that integrate Twitter content into their services, and insights without your permission.
- [ ] None of the above

Note: This survey is designed to be an online survey, hence; the design presented here does not reflect how the survey has been presented. For full access of the survey. Please refer to this link: https://www.surveymonkey.com/r/JWB2SYW