UNDERSTANDING PRIVACY LEAKAGE CONCERNS IN FACEBOOK: A LONGITUDINAL CASE STUDY

Thesis submitted for the degree of Doctor of Philosophy by

Arshad Jamal

School of Information Systems Computing and Mathematics
Brunel University

April 2013

ABSTRACT

This thesis focuses on examining users' perceptions of privacy leakage in Facebook – the world's largest and most popular social network site (SNS). The global popularity of this SNS offers a hugely tempting resource for organisations engaged in online business. The personal data willingly shared between online friends' networks intuitively appears to be a natural extension of current advertising strategies such as word-of-mouth and viral marketing. Therefore organisations are increasingly adopting innovative ways to exploit the detail-rich personal data of SNS users for business marketing. However, commercial use of such personal information has provoked outrage amongst Facebook users and has radically highlighted the issue of privacy leakage. To date, little is known about how SNS users perceive such leakage of privacy. So a greater understanding of the form and nature of SNS users' concerns about privacy leakage would contribute to the current literature as well as help to formulate best practice guidelines for organisations.

Given the fluid, context-dependent and temporal nature of privacy, a longitudinal case study representing the launch of Facebook's social Ads programme was conducted to investigate the phenomenon of privacy leakage within its real-life setting. A qualitative user blogs commentary was collected between November 2007 and December 2010 during the two-stage launch of the social Ads programme. Grounded theory data analysis procedures were used to analyse users' blog postings. The resulting taxonomy shows that business integrity, user control, transparency, data protection breaches, automatic information broadcast and information leak are the core privacy leakage concerns of Facebook users. Privacy leakage concerns suggest three limits, or levels: organisational, user and legal, which provide the basis to understanding the nature and scope of the exploitation of SNS users' data for commercial purposes. The case study reported herein is novel, as existing empirical research has not identified and analysed privacy leakage concerns of Facebook users.

DEDICATION

I dedicate this thesis to my beloved parents, whose prayers were with me during my doctoral research and without whose support this thesis might not have been written.

The support and encouragement of my loving wife has been immense and persistent indeed. You have shown extraordinary patience throughout my PhD, particularly when I passed the stress of my research on to you and you stood firm, which lifted my motivation to complete this thesis.

My children Zoha, Abdullah, Abdul Rehman and Mustafa deserve special appreciation as I couldn't give them the time and company they needed due to the competing demands of my PhD. But your smiles and cheerfulness inspired me the most and I am deeply indebted to you all. In particular, Mustafa, who was born during my PhD – hence another contribution, your laughter and cheerfulness are the greatest moments of my life. I love you all.

Last but not least, I dedicate this thesis to my in-laws for their continued support and inspiration.

I thank you all,

Arshad Jamal

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr. Jane Coughlan and Prof. Mark Lycett, for their academic and emotional support, which was invaluable and without which this research might not have been possible. Dr. Jane Coughlan in particular supported me during hard times for which I am greatly indebted.

I would also like to thank my previous supervisor, Dr. Melissa Cole, for her guidance and support. I am grateful to all members of Rubicon research group within Centre of Information systems research for their valuable suggestions for improving my research.

I express my deep gratitude to you all.

PUBLICATIONS

Peer reviewed Journal and Conference articles contributed to this thesis:

Jamal, A., Coughlan, J. and Kamal, M. (forthcoming). Mining Social Network Data for Personalisation and Privacy Concerns: A Case Study of Facebook's Beacon. International Journal of Business Information Systems.

Jamal, A. and Coughlan, J. (submitted). Investigating Privacy Concerns in Relation to the Use of Business Analytics in Social Networks: A Case study of Facebook's Beacon. Journal of Enterprise Information Management.

Jamal, A. and Cole, M. (2009). "A Heuristic Evaluation of the Facebook's Advertising Tool Beacon". Information Science and Engineering (ICISE), 2009 1st International Conference , vol., no., pp.1527-1530, 26-28 Dec. 2009

Jamal, A. and Cole, M. (2009) "Rethinking privacy in social networks: A case study of beacon". In: Poulymenakou, A, Pouloudi, N, Pramatari, K (eds). 4th Mediterranean Conference on Information Systems. Athens: Athens University of Economics & Business. 2009. CDROM.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

1.1 Research Background and Motivation

The concept of *privacy* has been recognised as a vital notion since the ancient Greek and Chinese civilisations emerged (Warren and Brandeis, 1890). However, recent developments in information technologies have led to a renewed interest in this field. This is partly because online information is persistent (i.e. stored permanently); searchable; replicable (i.e. it can be copied); and 'invisibly read', in that it can be difficult to identify the viewer (boyd, 2008). In the modern information age, privacy has become a vital human right (Rotenberg, 2000), yet it is far from easy to define (Michael, 1994) because it is a complex (Solove, 2006), discipline-dependent (Xu et al., 2008), context-specific (Ajzen and Fishbein, 2005; Nissenbaum, 2004) and temporal concept (Palen and Dourish, 2003). Accordingly, various discipline-specific conceptualisations of privacy have emerged. For instance, in law it is considered as a *'right'* or *'entitlement'* (Warren and Brandeis, 1890), in philosophy and psychology as a *'state of limited access or isolation'* (Schoeman, 1984), and in information systems and social sciences as a *'control'* (Culnan, 1993; Westin, 1967).

Thus, privacy definitions, relationships and concepts are disconnected, inconsistent, underdeveloped, and not empirically validated (Xu et al., 2011). Solove (2006) sums up this situation well when he notes "*privacy is a concept in disarray. Nobody knows what it means*" (p.477). Such fragmented views of privacy are reflected in different privacy conceptualisations and some of those encompass "*(among other things) freedom of thought, control over one's body, solitude in one's home, control over personal information, freedom from surveillance, protection of one's reputation, and protection from searches and interrogations"* (Solove, 2008; p.1). However, the most relevant conceptualisation of privacy in the current digital age concerns control over personal information – also known as information privacy.

But such a one-dimensional conceptualisation of information privacy merely as a matter of control is over simplistic, as a number of information system (IS) advocates (such as Malhotra et al., 2004; Smith et al., 2011; Son and Kim, 2008; Xu et al., 2011) unanimously consent that it is a multi-dimensional concept. This indeed has opened up a plethora of information privacy conceptualisations within the discipline of IS (see Chapter two for detailed literature on information privacy). The author will use the terms information privacy and privacy interchangeably hereafter in this thesis. Furthermore, the recent rapid development in information technologies poses further challenges to online users' privacy: the continuing advancements in information technologies and ubiquitous computing environments open countless opportunities for organisations to track consumers online, and to collect, process, distribute and use immense amount of data about them, thus limiting users' ability to control their personal information (Dinev and Hart, 2006 ; Hui et al. 2007; Malhotra et al. 2004; Smith et al., 2011; Solove, 2006). Even the most prominent organisations, such as Facebook, Google, Microsoft and Yahoo, use their customers' personal information for personalised advertising and share this data with third-party partner companies, as reported by a study on evaluating organisational privacy practices of the top 50 most visited websites (Gomez et al., 2009).

Accordingly, Smith et al., (2011) warn that such sharing and use of personal information is not without privacy risks. Moreover, a Wall Street Journal study revealed a striking finding i.e. consumer tracking is one of the fastest growing businesses on the internet today (Angwin, 2010). Although tracking has existed for a long time, the study highlighted that such tracking is now so sophisticated that users' location, income, shopping interests and even medical conditions are collected in almost real time to build and subsequently sell rich consumer profiles to advertisers as well as to tracking companies, data brokers and advertising networks (Angwin, 2010). Similarly in another context, iPhone users raised concerns when Apple shared their personal data such as age, gender and location with third parties (Thurm and Kane, 2010). Such privacy breaches, perceived as privacy violations by online users, highlighted the conflict between what companies need to do (i.e. collect, disseminate, use and aggregate consumer information) and what consumers expect in terms of management of their privacy (i.e. confidentiality of their personal information). Online users (consumers) have shown heightened concerns regarding their privacy, which is unprecedented compared to any other era of history

(Solove, 2008). Such concerns of privacy violations are reflected in the mainstream media and consumer opinion polls, with one poll showing that *"72 percent [consumers] are concerned that their online behaviours were being tracked and profiled by companies"* (Consumers-Union 2008). Perhaps that is why Son and Kim (2008) suggested that "s*uccessfully addressing information privacy issues in an online environment is particularly relevant to the growth of the information age" (p.504).*

Iachello and Hong (2007), therefore, call for more studies to understand the attitudes and behaviours of online users relating to information privacy, particularly with regard to leakage of users' privacy as a consequence of organisational and business practices. Responding to this call, IS researchers carried out extensive research on information privacy in a variety of contexts. For instance: Awad and Krishnan (2006), Cranor (2003), Dinev and Hart (2006), and Hui et al., (2007) studied privacy in e-commerce; Iachello and Hong (2007), Patil and Kobsa (2004), and Grinter and Palen (2002) explored privacy in personal communication; Dinev and Hart (2004), Korzaan et al., (2009), Malhotra et al., (2004) and Son and Kim (2008) studied it in internet use; Angst and Agarwal (2009) studied privacy in the context of e-health; and Lwin et al., (2007) and Okazaki et al., (2009) explored privacy within online and mobile advertising. Though such research spans individual, organisational and societal levels (Xu et al., 2008), few studies exist, which sufficiently investigate the way organisational practices impact on consumer privacy; rather, research exploring the link between individual actions and consumer privacy has dominated (Smith et al., 2011; Xu et al., 2011). This gap in current research is surprising, particularly in the current digital age when organisations are increasingly seeking to employ innovative ways to exploit consumer data for commercial purposes, and when the popular emergence of social media tools such as social network sites (SNSs) has facilitated the disclosure and dissemination of personal information, representing a tempting source of data for business marketing. Within SNSs, users have shown increasing concerns about their privacy, such as third party use of personal data, tracking of online behaviour, unintentional (accidental) information disclosure, reputational damage due to gossip and rumours, and harassment or stalking (boyd and Ellison, 2007; Rosenblum, 2007).

1.2 Privacy Leakage in Social Network Sites (SNSs)

SNSs are free web-based services that allow people to create an online identity called a profile, connect with their family and friends and share their personal life with them (boyd and Ellison, 2007). SNSs fulfil the basic human craving for interaction and facilitate connection and communication between people while allowing them to generate their own content (Tufekci, 2008a). SNSs have become so popular that the largest SNS Facebook, has over one billion monthly active users (Facebook, 2013). Facebook is now the world's number one most visited website (Alexa, 2012). Certainly, SNSs represent huge repositories of user-generated real-time behaviour (personal actions, choices, interests, browsing) as well as personal data so rich in detail as to be able to identify a person (Krishnamurthy and Will, 2010; Rosenblum, 2007). Surely this identity rich data, such as name, email, phone numbers, gender, and age, have huge commercial potential for social network service providers as well as third parties, such as advertisers, tracking companies, advertising networks, and data brokers.

Thus, SNS users' data is an irresistible resource for SNS service providers, online marketers and information aggregators (who combine behavioural data with the personal data) (Bonneau et al., 2009a, 2009b; Martin, 2010). Also, SNS data willingly shared between friends' networks intuitively appears to be a natural source of current advertising strategies, such as word-of-mouth (WOM) and viral marketing. Kirkpatrick (2007, p.1) echoes the importance of SNS data whilst remarking "*now there's starting to be real money in the business, as every major consumer advertiser realizes that if you can engage effectively with these newly networked hordes, they become agents of your brand"*.

To exploit social network data for commercial gains, SNSs and third party organisations have started interacting by deploying innovative technologies in the *"art of the possible"* (Smith et al., 2011). Such interaction between SNSs and third parties has become so sophisticated that it enables two-way communication of data between SNSs and third parties. Specifically, Facebook users' personal data is shared with third party sites and their behavioural data from third party websites is shared back with Facebook. Since such sharing of personal information between Facebook and third parties occurs without

the explicit consent and knowledge of users, therefore, it is called privacy leakage. Privacy leakage has been reported in the studies of Krishnamurthy and Wills (2008, 2010), Bonneau et al., (2009a, 2009b) and Debatin et al., (2009). Krishnamurthy and Wills (2010) characterise privacy leakage as the ability of third parties to link personal identifiable information (PII) of SNS users with their browsing data. Indeed, the leakage poses threats to users' privacy and provokes concerns amongst SNS users.

Also, the integration between SNSs and third parties seems aggressive as it is constantly changing the private and public boundaries of personal information (Nissenbaum, 2004; Solove, 2008). This enables businesses to build digital dossiers of personal lives of individuals to gain commercial advantage. Such privacy leakage indeed makes users vulnerable, as John Battelle (2010) of *Wired Magazine* notes, warning against such data storage as being *"a database of humankind intentions that has the potential to be abused in extraordinary fashion".* In contrast, SNS service providers assume that most of the information that users share voluntarily is there to be collected, disseminated, aggregated and used. This philosophy is well reflected in the thinking exposed by Mark Zuckerberg, the CEO of Facebook: *"People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people"* (2010). Due to leakage of data, the privacy of online users is diminishing day by day and Scott McNealy, the former founder of Sun Microsystems, alarmed us when he remarked *"you have zero privacy anyway, get over it"*. Such contrasting expectations need to be further investigated, especially from the point of view of SNS users. Users' concerns could be heard in the public outcry during the launch of Facebook's social marketing tool Beacon, which broadcast (leaked) SNS users' private actions performed on third-party sites to their Facebook profiles (Schonfeld, 2007).

The Committee of Privacy in the Information Age suggests that privacy concerns are better understood in a situation specific context (e.g. in social network sites) rather than in general (Waldo et al., 2007 cited in Xu et al., 2011). Likewise, Nissenbaum (2004) argues that changes in the context of information dissemination and use should lead to changes in privacy protection. Similarly, Ajzen and Fishbein (2005) addressed the situation specific nature of privacy and argued that attitudes towards privacy expressed outside a

specific context are unreliable predictors of users' behaviour. Therefore, organisations need to adjust their privacy protection and business processes according to the degree to which their business models intrude upon and damage customer privacy, rather than adopting a one-size-fits-all policy in all contexts of information dissemination and use.

The recent proliferation of SNSs has changed the information disclosure landscape quite dramatically (Smith et al., 2011). In the traditional context of e-commerce, privacy exists in 1-1 setting such that consumers' data is shared with the e-commerce service providers only. In contrast, SNSs are built to disseminate users' personal data over the network such that the data is accessible to other users as well as third party organisations (e.g. advertisers) (Jamal and Cole, 2009). Likewise, Rosenblum (2007) warns that SNSs are designed to support dissemination of personal information. Therefore, privacy problems in SNSs are more prominent and complex than traditional online environments, since what constitutes a breach of privacy is so often ambiguous, meaning different things to different people at different times. Additionally, privacy management should not be considered in isolation but rather as a function of a sequence of events in the past, present and future. This highlights the temporal nature of privacy which signifies that the current information disclosure decisions are influenced by the past and future situations (Altman, 1975; Palen and Dourish, 2003). For example, academics publish their web pages not only to publicise their expertise but also to keep information about their papers in order to limit future accessibility (Palen and Dourish, 2003). Thus, the present information disclosure is influenced by past and future situations, hence posing further challenges to managing privacy in the digital age. In recent years, the main stream media has highlighted privacy issues in SNSs (Gurses et al., 2008). This leads to the interesting question of the extent to which SNS users' perception of privacy is influenced by its perceived loss as reported in the media. So the context-specific, complex, fluid and temporal nature of privacy, together with the increasing interest of businesses in exploiting massive identity-rich social network data for commercial purposes (Jones and Soltren, 2005; Krishnamurthy and Wills, 2010), has raised concerns amongst multiple stakeholders, such as service providers, privacy activists, policy makers, researchers and individual users (Smith et al., 2011).

Privacy scholars, responding to this call, have become more and more interested in researching privacy within social networks. However, the phenomenon of privacy leakage in SNSs has been reported only recently and hence no study to date has investigated privacy leakage concerns of SNS users. Rather, most of privacy research in SNSs focuses on: users' information-sharing behaviour (e.g. Acquisti and Gross, 2006; Barnes, 2006; boyd and Marwick, 2011), privacy awareness amongst SNSs users (Debatin et al., 2009; Fogel and Nehmad, 2009), changes in privacy settings (e.g. boyd and Hargittai, 2010; Lewis and George, 2008) and exploring the link between user demographics and privacy behaviour (e.g. Lenhart and Madden, 2007). The focus of the majority of the research has been on what Debatin et al., (2009) call is the visible part (see Figure 1.1) that only consists of SNS users' profile information and their interactions, whereas few studies focus on the large part of SNS users' data which is being shared with third parties by SNS service providers (Bonneau et al., 2009a, 2009b; Debatin et al., 2009).

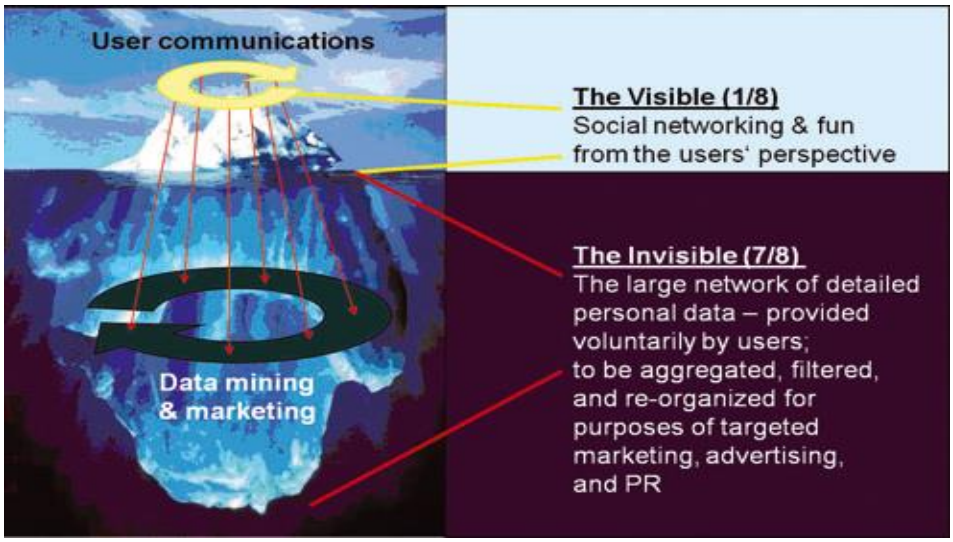


Figure 1.1: Facebook Iceberg Model

(Debatin et al., 2009: Iceberg Image © Ralph A. Clevenger/CORBIS)

Krishnamurthy and Wills (2010) in their investigation of 12 SNSs[1] demonstrate that all of them, with the exception of Google-owned Orkut, outflow users' personally identifiable information to third-party advertisers and aggregators. However, little is known about

[1] Bebo, Digg, Facebook, Friendster, Hi5, Imeem, LiveJournal, MySpace, Orkut, Twitter, Xanga and LinkedIn

how SNS users perceive such leakage of their privacy. Krishnamurthy and Wills (2010) also highlight the need to understand privacy leakage more holistically. They argue that understanding SNS users' perceptions of privacy leakage may provide a basis for understanding privacy leakage holistically. To the best of the author's knowledge, no study has examined how SNS users perceive privacy leakage in social network sites and whether such perceptions change overtime.

1.3 Research Aim and Objectives

Therefore, the aim of this research is:

> *"To identify and analyse user perceptions of privacy leakage in a social network site: Facebook. In so doing, to add to existing knowledge by developing a taxonomy of privacy leakage concerns which improves our understanding of the nature and form of privacy leakage concerns of SNS users".*

As the aforementioned research background and motivation suggest, privacy is a complex, fluid, context-specific and temporal concept and hence the privacy perceptions of users are constantly in flux and it is difficult to envisage what would be perceived as being harmful to privacy, as different people perceive privacy differently in different contexts. Online social networks further complicate concepts of privacy as organisations are increasingly exploiting SNS data for commercial gain (e.g. using personal data of SNS users for personalised advertisements without their knowledge and consent). Therefore, the empirical investigation of SNS users' perceptions of privacy leakage in real life contributes to a better understanding of the complex phenomenon of information privacy in SNSs.

Therefore, this research is set up to fulfil following research objectives:

1. To empirically investigate the privacy leakage concerns of Facebook users;
2. To analyse how and why privacy leakage is perceived as privacy violation by SNS users;

3. To investigate whether SNS users' perceptions of privacy leakage change overtime;
4. To develop a taxonomy of privacy leakage concerns through empirical means in order to understand the nature and form of privacy leakage concerns in SNSs;
5. To evaluate and validate the taxonomy of privacy leakage concerns through existing taxonomy of privacy concerns; and
6. To analyse and discuss how the developed taxonomy of privacy leakage concerns can be applied as best practice guidelines in social network sites.

1.4 Research Approach

This thesis adopts an interpretive research approach to investigate privacy leakage concerns of SNS users. The major motivation for this choice is derived from the ontological belief that reality is subjective and can be studied through the meaning people assign to reality (Klein and Myers, 1999). Further motivations to use an interpretive approach are justified below:

As mentioned previously, privacy is a complex social phenomenon and many factors such as the social, technological, psychological and commercial, influence users' perceptions of privacy leakage in SNSs. It is often not easy to study the complex and interrelated issues within the technological and social context in which they are developed and used (Wood-Harper and Wood, 2005). Moreover, IS research relies heavily on social context and not just technological context (Mingers, 2001). Therefore, the author argues that SNS users' privacy leakage perceptions can be appropriately investigated through an interpretive approach as "*[it] assume[s] that people create and associate their own subjective and intersubjective meanings as they interact with the world around them. Interpretive researchers thus attempt to understand phenomena through accessing the meanings participants assign to them"* (Orlikowski and Baroudi, 1991).

Since privacy leakage in SNSs is a recent phenomenon, little is known of how SNS users perceive privacy leakage. Therefore, this research aims to understand SNS users' perceptions of privacy leakage. As Moore and Read (2006) suggest, users' perceptions

should be studied through a subjective meaning of reality, hence the interpretive approach seems fitting within the context of this thesis. The social world cannot be ideally divided into variables and hence should be studied in its totality, as Orlikowski and Baroudi (1991) argue: *"social process is not captured in hypothetical deductions, covariances and degrees of freedom. Instead, understanding social process involves getting inside the world of those generating it".* Since privacy leakage is a social phenomenon, it should be studied through the explanations people generate for it, thus the interpretive approach seems appropriate.

Within an interpretive approach, case study research strategy is adopted in this thesis as it provides the opportunity for the researcher to study the phenomenon of privacy leakage within real-life settings and hence develop a new theory (Yin 1994; Benbasat et al., 1987) of privacy leakage concerns. Therefore, the case study seems an appropriate research strategy to adopt to study the privacy leakage concerns of SNS users. The details of research methodology are discussed in Chapter three.

1.5 Thesis Structure

This thesis consists of seven chapters (see Figure 1.2):

Chapter Two provides the review of the literature related to information privacy and also discusses privacy concerns of online users in general, as well as social network users in particular to contextualise the current research.

Chapter Three discusses research paradigms and approaches and provides justifications for their selection in this thesis. It also discusses and justifies the data collection and analysis methods.

Chapter Four reports the finding and analysis of stage one of the longitudinal case study i.e. the launch of Facebook Beacon – a personalised marketing tool aimed to maximise use of social network data for business marketing. The outcome of this chapter is the

Beacon privacy framework, which identified the privacy concerns of online social network users.

Chapter Five presents empirical findings and analysis of stage two of the longitudinal case study – the launch of Facebook Connect which enabled third-party websites to broadcast users' actions performed on these sites back to Facebook profiles. The outcome of this chapter is the Connect privacy framework, which identified the privacy leakage concerns of SNS users.

Chapter Six provides a consolidated and a developmental view of the findings of stages one and two of this research. In order to do that, the Beacon and Connect privacy frameworks were compared and then combined. Finally, the resultant taxonomy of privacy concerns is evaluated through relevant theoretical framework.

Chapter Seven provides a summary of the thesis chapters. Then conclusions are derived from the thesis. Fair information processing guidelines are presented next. Research contributions are highlighted thereafter. Finally, research limitations are identified and recommendations for future research are presented.

Chapter 1: Introduction
- Research background and motivation
- Privacy leakage in SNSs
- Research aim and objectives
- Research Approach

Chapter 2: Literature Review
- Characteristics of SNSs (history, typology, usage, business model and information disclosure)
- Information privacy concerns and privacy leakage in SNSs
- Information privacy concerns (frameworks and taxonomy

Chapter 3: Research Methodolgy
- Explanation of research paradigms and methods
- Justification for the choice of research approach
- Rationale of using inductive qualitative research method (case study)
- Data Collection (use of blogs, ethics of using blogs)
- Data Analysis Method (Grounded theory, use of NVIVO, analysis process)

Chapter 4: Longitudinal Case Study Findings- The Beacon Case
- Overview of Beacon launch and user backlash
- Details of blogs commentary (summary, time series)
- Empirical findings (Beacon privacy framework)
- Analysis of empirical findings
- Conclusion of Beacon study

Chapter 5: Longitudinal Case Study Findings- The Connect Case
- Overview of Connect launch
- Details of blogs commentary (summary, time series)
- Empirical findings (Connect privacy framework)
- Analysis of empirical findings
- Conclusion of Connect study

Chapter 6: Evaluation and Discussion of the Findings
- Delineation of research findings (consolidation of privacy frameworks)
- Discussion of privacy leakage taxonomy
- Evaluation of privacy leakage taxonomy
- Applying taxonomy to LinkedIn case

Chapter 7: Conclusions and Future Recommendations
- Recapitulation of research and findings
- Conclusions and contributions
- Research limitations
- Recommendations for future research

Figure 1.2: Thesis Structure

CHAPTER TWO: LITERATURE REVIEW

2.1 Chapter Overview

Chapter one highlighted that, despite a significant increase in the number of privacy studies, little is known about how the users of social network sites (SNSs) perceive privacy leakage from those sites. The aim of this research is, therefore, to discover what users think and how they feel about their private information being leaked from SNSs. In pursuit of this research goal, this chapter begins by characterising SNSs with a particular focus on privacy leakage of SNS users. It also identifies and discusses privacy concerns of SNS users, as reported in the literature. Then, underlying features of privacy leakage in social networks are discussed – thus highlighting a gap in the current knowledge relating to the lack of understanding of SNS users' perceptions of privacy leakage in SNSs. Finally, current theoretical frameworks to understand information privacy in the context of digital information are discussed.

2.2 Online Social Network Sites

2.2.1 Definition and Characteristics of SNSs

Online social networks have proliferated in the recent past, to the extent that more than a billion people use SNSs today. Facebook, the largest and the most popular SNS today, has one billion users alone (Facebook, 2012), thus making it the world's number one most visited website (Alexa, 2012). Other examples of popular SNSs are Bebo, Hi5, MySpace, Orkut and Twitter. SNSs are free web-based services which allow individuals to create an online identity called a profile, connect with their family and friends and share with them their photos, videos and personal stories (boyd and Ellison, 2007). SNSs fulfil the basic human desire to discourse (Danah, 2007; Tufekci, 2008b) and socialise, for which interaction people willingly disclose personal information and share personal interests. Thus, SNSs provide a tempting source of data for marketers and social network service providers. For SNSs to flourish, service providers adopt such technological platforms and information management practices as facilitate personal information disclosure and

dissemination. Moreover, as business models of SNSs largely depend on users' data SNS service providers are looking to adopt innovative ways to exploit it for commercial gains. Krishnamurthy and Wills (2010) in a recent study found that many SNSs are leaking users' data (either voluntarily or involuntarily) without the knowledge of users. As a result, SNSs have raised concerns of privacy amongst users (Acquisti and Gross, 2006; Barnes 2006; Gross and Acquisti, 2005; Krishnamurthy and Wills, 2010; Strater and Lipford, 2008) as well as among policy-makers.

In addition, Smith et al., (2011) caution that the recent proliferation of social network sites has changed the information disclosure landscape quite dramatically. Consequently, SNSs have become an important area for fieldwork into privacy research (Bonneau et al., 2009b) and a considerable amount of literature has been published on the privacy issues raised by them (e.g. Bonneau et al., 2009a; Gross and Acquisti, 2005; Krishnamurthy and Wills, 2010; Rosenblum, 2007). Facebook profiles are often more accurate because people use their real world identities to communicate with real-world friends (Dwyer, 2007a). This makes it a particularly interesting place for privacy research (Bonneau at al., 2009b). boyd and Ellison (2007) summarise features of SNSs as "*web-based services that allow individuals to: (1) Construct a public or semi-public profile within a bounded system, (2) Articulate a list of other users with whom they share a connection, and (3) View and traverse their list of connections and those made by others within the system."* (boyd and Ellison, 2007: p.2). This definition highlights the three characteristics of SNSs discussed below with particular consideration of privacy issues.

- ***User Profile and Privacy related Issues:*** First, a user can "type oneself into being" (Sunden, 2003, p. 3) by creating an online identity called a profile, which consists of users' first name, middle name, surname), photos, home address, phone number, email address, gender, birth date, educational institutions attended, work related and other information. Also, SNS users' pictures and videos can be linked with profiles. The rich detail of such profiles clearly gives them commercial value (Felt and Evans, 2008), since users can be personally identified (Sweeney, 2000). It is also why leakage of such personal information poses a threat, real or perceived, to the privacy of SNS users (Felt and Evans, 2008; Krishnamurthy and Wills, 2010; Bonneau et al., 2009a).

Users' profiles contain information which is both public and private. Public information is accessible to all SNS users, even those who do not have an account and also to web crawlers[2] (Bonneau et al., 2009a). Supposedly, private information should only be accessible to users' friends, which unfortunately is not the case as SNS data is constantly being passed on to third parties (e.g. advertisers), often without users' knowledge (Debatin et al., 2009; Krishnamurthy and Wills, 2010). Although privacy controls are provided by which users can manage their personal information, most privacy settings are public by default which enables dissemination of information within the SNS and to outside third parties, such as advertisers.

User interactions within SNSs also threaten users' privacy because organisations (e.g. SNSs and advertisers) are increasingly looking for ways to exploit SNS data. For example in September 2007, Facebook allowed access to users' public profiles listings (containing information such as user name, photo and photos of eight random friends) to those who had not logged in to their accounts, including web crawlers (Bonneau et al., 2009b). Given the sophistication of available web crawling, data mining and de-anonymising tools, the public exposure of eight random pictures of friends of an SNS user allows web crawlers to identify all his/her friends (Bonneau et al., 2009b). Such identification of a private profile user is possible by linking its profile with the profile of an insecure friend (e.g. a friend with a public profile) (Bonneau et al., 2009b). Subsequently, third-party advertisers can use public information to predict private information such as user demographic, as well as other data such as political affiliation and sexual orientation. In an interesting experimental study, two MIT students successfully identified the gay students based on their friends' data, raising unique concerns relating to users' privacy (Johnson, 2009) based on friends' data. These findings also corroborate with the Facebook's Iceberg model (Figure 1.1), which suggests that restricting a profile's visibility only restricts it within the tiny visible part of

[2] A program which browses World Wide Web pages automatically. (Information accessed on the 4th Dec,2009 at http://www.sciencedaily.com/articles/w/web_crawler.htm

the Iceberg, whilst the large invisible part is still being fed to third parties (Debatin et al., 2009).

SNS users' profiles are public by default but users have the control to change the setting in order to increase their privacy. However, research suggests that most of SNSs users do not change profile settings, and hence are more vulnerable to privacy threats. For instance, Gross and Acquisti (2005) found that only 1.2% of Facebook users studying in a University change their thumbnail profile setting and only 0.06% of SNS users changed the default profile setting. Similarly, Krishnamurthy et al. (2008) highlighted that only 1% of Twitter users change default privacy settings. Reflecting on the possible reasons for such lax user behaviour, Raynes-Goldie (2010) suggests that SNSs users show greater concerns over the privacy of their information when it is leaked to someone they know (e.g. friends and parents) compared to leakage of personal information to governments or organisations. However, this is changing as users become more aware due to media coverage of privacy issues in SNSs (boyd and Hargittai, 2010).

In a recent longitudinal study, boyd and Hargittai (2010) found that an overwhelming majority of SNS users change privacy settings as a result either of a change in organisational practices (e.g. privacy settings) or of media coverage of privacy issues. This is an interesting finding, as it challenges the general perception that SNS users (particularly the young) do not care about privacy. This shift in user behaviour may be associated with the more aggressive business practices reported in the media, making users feel more vulnerable and hence opting to change privacy settings. The way SNS users perceive their privacy can be seen to evolve overtime as organisational practices change. Now the interesting question is what organisational activities cause privacy risks and how and why these cause concerns amongst SNS users.

- ***Connections and related Privacy Issues:*** Secondly, SNSs enable online users to create connections of friends (which are usually separate from their offline friends), with whom they can contact and share information. The strength of

network is usually dependent on the number of friends (contacts). For example, any update on SNS user's page is disseminated to all contacts as well as sub-contacts (Dube and Adomaitis, 2009). Therefore, SNSs are nourished by relationships (Dube and Adomaitis, 2009). However, the number of connections influences a user's level of privacy protection because the probable number of threats to privacy increases due to data persistence (Debatin et al., 2009). For example, when SNS user A, who has a greater number of connections than user B, posts a message to his/her social network, the persistence of the transmitted message or information will be greater than that of user B's transmissions; hence user A is more vulnerable to privacy threats because of having more connections.

The number of connections of a particular SNS user is also referred to as a social graph. In the aforementioned example, the threats to a user's privacy arise from data persistence. Another type of privacy vulnerability relates to the social graph and concerns the ability of web crawlers or tracking software to easily extract details from data as opposed to lifting a complete profile (Bonneau et al., 2009a). This simply means that even if an SNS user has made all profile information strictly private, that private data can be easily extracted through his/her social graph. This indeed has implications for SNS users' privacy.

- ***Community related Privacy Issues:*** Finally, SNSs enable users to not only connect with their own friends but friends of friends and hence create a community. Dube and Adomaitis (2009) give an example of a community such as the alumni of a high school which share common thoughts on a SNS. However, this characteristic may cause secondary privacy leakage when an SNS user shares private information of her friend with another friend who may not even know her (Krishnamurthy and Wills, 2009a). What the aforementioned discussion exhibits is that the technological platform of SNSs facilitates the disclosure and dissemination of personal information, however, it can cause concerns for the privacy of SNS users. The question arises why people still use SNSs. The author contemplates that the history of SNSs can provide some insights here.

2.2.2 Brief History and Typology

Figure 2.1 shows the timeline of SNSs. SixDegrees.com, the first SNS which was launched in 1997, could not succeed and closed down in 2000, primarily because the site did not offer many features other than creation of profiles and connections with friends (Boys and Ellison, 2007). Many other SNSs which allowed the creation of profiles and articulated lists of friends, were launched between 1997 and 2000, namely: AsianAvenue, BlackPlanet, and MiGente. Interestingly, most SNSs could identify friends of users on their profiles and failed to seek prior approval (O. Wasow, personal communication, 2007: cited in boyd and Ellison, 2007), whereas most of SNSs today do not allow users publish names of their friends on their profiles. Other SNSs which emerged during 2000-2001 included Korean Cyworld and the Swedish community rich network LunerStorm. However, the first wave of SNS ended in 2001.

Began in 2001, the second wave of SNSs aimed to leverage business networks (boyd and Ellison, 2007). The sites such as LinkedIn, and Friendster were launched to support technology and business professionals (cited in boyd and Ellison, 2007). Interestingly, the founders of these sites joined together rather than compete against each other (Festa, 2003). However, only two sites such as Friendster and LinkedIn were successful to attract masses (boyd and Ellison, 2007). However, the popularity of Friendster faded because of technical factors (e.g. failures of servers to manage growth), and social factors (e.g. cultural clash between different user groups). At the same time when the popularity of Friendster was waning in the U.S., the network flourished in Indonesia, Malaysia and Philippines (Goldberg, 2007). Therefore, an Asian buyer offered $100 million to buy Friendster (Lee, 2009), however, the deal was finalised for only $24.6 million (Arrington, 2009).

SNSs hit the mainstream in 2003 when the third wave began, prompting social network analyst Clay Shirky (2003: p.1) to coin the acronym YASNS: *"Yet Another Social Networking Service"*. Most of the SNSs emerged at that time used profile features similar to Friendster (boyd and Ellison, 2007). MySpace emerged as a competing SNS when it was launched in 2003 as it was targeted to attract deprived Friendster users (boyd and

Ellison, 2007), in which perhaps they were successful, as MySpace grew fast. It also offered users the improved personalisation services that, included group management, ability to blog and use applications (e.g. comparing movies) (boyd and Ellison, 2007). A significant majority of today's popular SNSs (20) – Facebook, MySpace, Hi5, Flicker, Orkut and Twitter – were introduced after 2003.

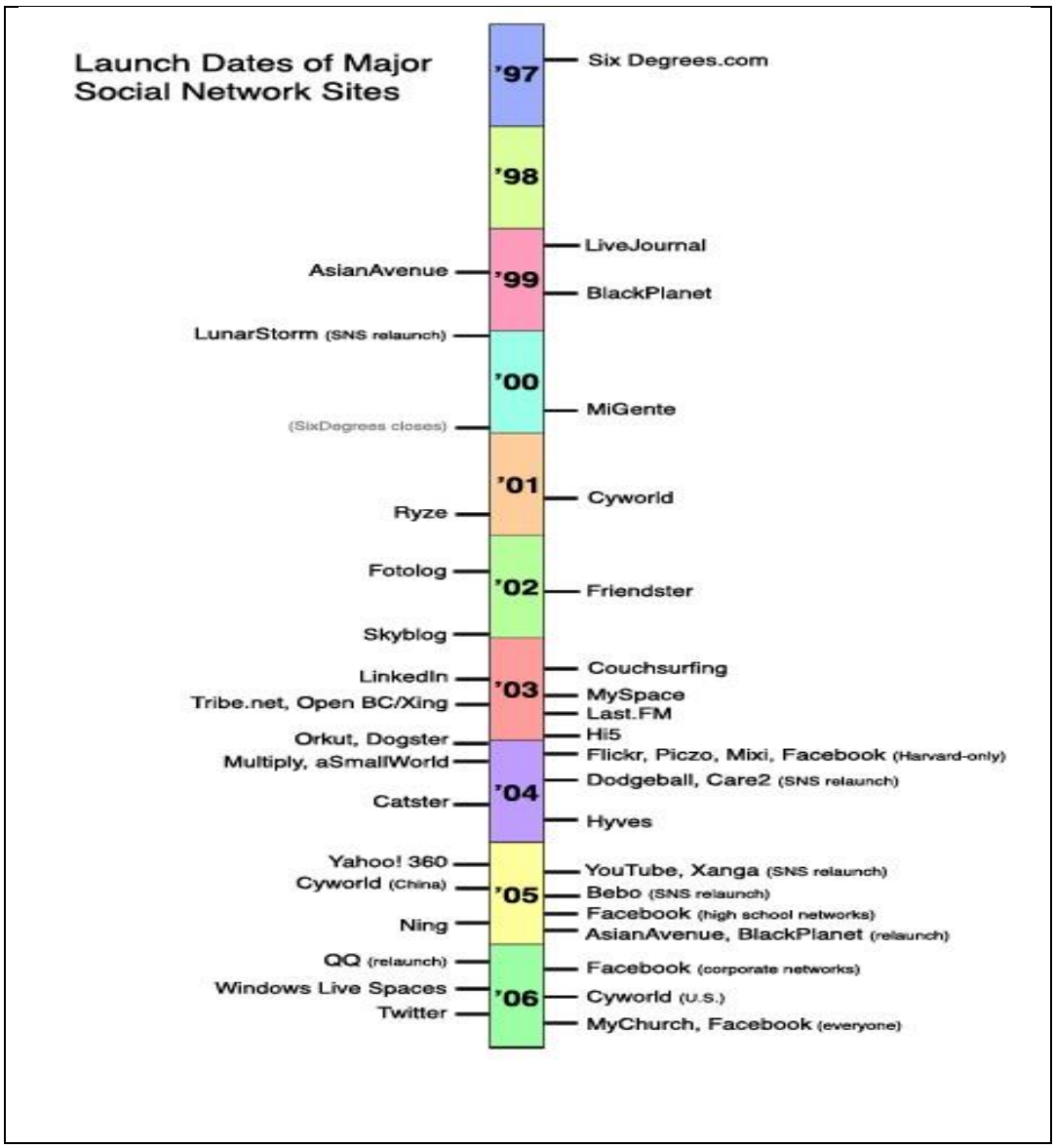


Figure 2.1: History of SNSs
(Source: boyd and Ellison, 2007)

But Facebook won unprecedented success and popularity and is now not only the largest SNS, but also the world's most visited website (Alexa, 2012). Facebook has achieved a landmark of more than one billion monthly active users (Facebook, 2013). Launched as a Harvard-only network in 2004 (Cassidy, 2006), Facebook gave access to high school

students and professionals in September 2005 and finally opened to everyone in 2006 (boyd and Ellison, 2007). Since then the network has gained mass popularity. Due to its massive popularity and the richness of its users' data, it has become a hugely tempting source of data for organisations for the purpose of business marketing. However, this has raised the risks not only for SNS users to protect privacy (Bonneau et al., 2009), but also for researchers, business owners, privacy advocates and policy makers (Smith et al., 2011). Particularly, privacy leakage of Facebook users due to its features and privacy management practices, has emerged as a challenging issue for Facebook users and policy makers. In concluding, the author states that like other websites, SNSs should be organised in the interests of individuals rather than those of businesses (boyd and Ellison, 2007; Dube and Adomaitis, 2009; Mislove et al., 2007). But the question is: *are they?* The number of privacy breaches cited recently appears to illustrate a different story. The latter can be attributed to an outcry amongst SNS users whenever their interests are overridden by commercial priorities.

Thelwall (2009) states that the current SNSs have been categorised according to the purpose they serve based on a typology (see Figure 2.2). The typology suggests SNSs serve three purposes such as socialising, social navigation and networking. In socialising SNSs, communication between members is the key. Examples of such networks are Facebook, Hi5, Bebo, MySpace, and Cyworld (Thelwall, 2009). In networking SNSs, such as LinkedIn, members communicate for professional purposes rather than for social reasons. For example, members of LinkedIn communicate to establish professional contacts. The third and final type of SNSs are used for social navigation purposes e.g. to find out information about videos on YouTube (other sites are digg and CiteUlike). However, YouTube has recently been used as a socialising rather than a navigational SNS as people share their videos, create channels, promote themselves, their work and their choices. Though the typology broadly outlines the purposes for which people use SNSs, the important question remains why users still use SNSs. The next section discusses this detail.

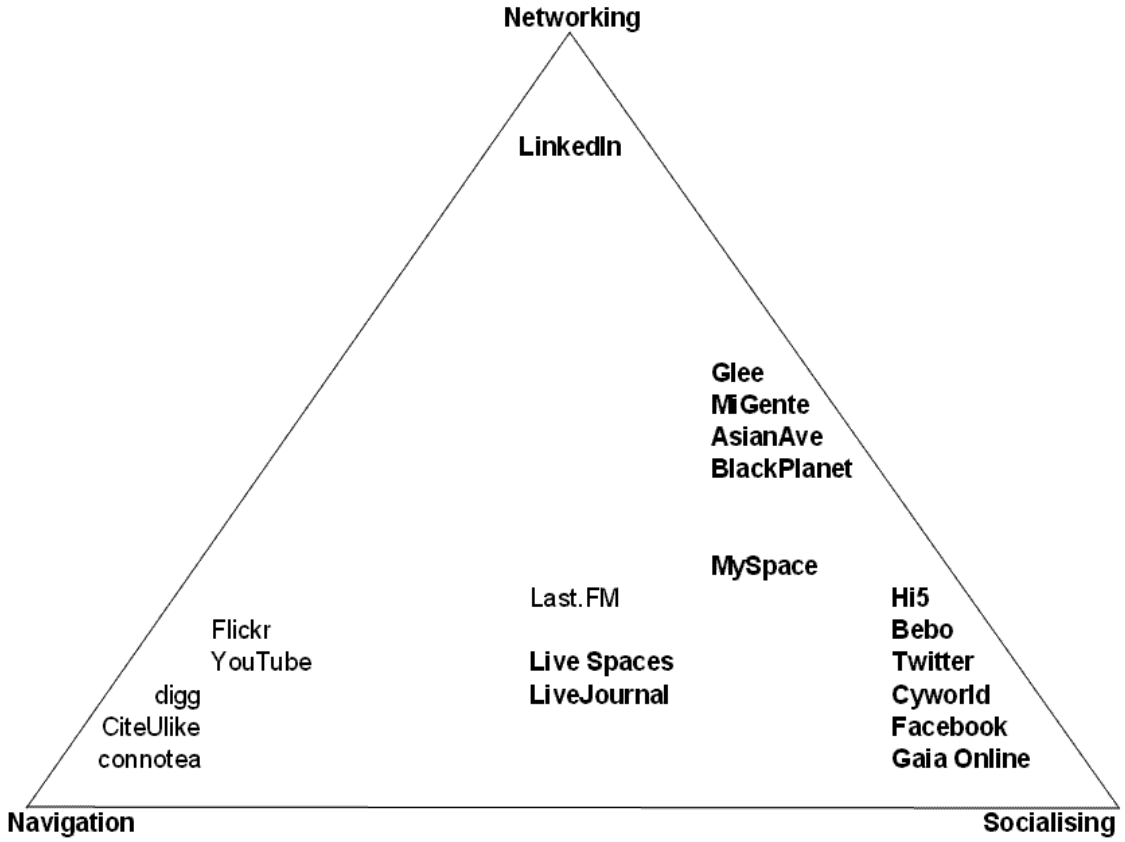


Figure 2.2: A Typological Presentation of Popular SNSs
(Source: Thelwall, 2009)

2.2.3 Uses of SNSs

Given the massive popularity of SNSs, Thelwall (2009:p.15) poses two legitimate questions: why SNSs have become so popular? and are SNSs a temporary phenomenon or are they here to sustain? These questions can be answered in the way SNSs fulfil a human desire to 'gossip' and exchange information about human relationships which has existed and will remain forever (Donath, 2007; Tufekci, 2008a). Thus, SNSs are likely to be sustainable unless something that more powerfully satisfies the same need replaces them (Thelwall, 2009). SNSs use a technological platform which facilitates people to gossip (Donath, 2007). Tufekci (2008a) argues that SNSs help in social grooming (e.g. gossip and small talk) as people who use SNSs are more likely to value socialising compared to those who do not use SNSs. Therefore, SNSs help build communities with their relationship creation features.

Likewise, Joinson (2008) suggests that users use Facebook for many reasons such as: to socialise, to share their identities, to share photographs, to share content, to view profiles of people, and to find and contact new people. Similar findings were found by Donath and boyd (2004) and Ellison et al. (2006) that people use SNSs to publish their profiles, photos and videos. These motivations are the cornerstone and prerequisites of the creation and nourishment of the online friendships that consequently determine the success of SNSs. In the absence of disclosure of personal information such friendships would wither and subsequently die. Therefore, the technological platform of SNSs and the applications are designed to support increasing level of disclosures of personal information. However, as Debatin et al., (2009) suggest, such technological platforms (e.g. Facebook) often raise privacy concerns amongst SNS users.

Another dominating motivation to use SNSs as suggested by Rosenblum (2007) is that they provide a simple and usable technological platform. The technological platform of SNSs and their applications are implemented in a simple and usable way to attract more people and more communications which is the life-blood of SNSs. Moreover, a slightly different but interesting perspective of the use of SNSs has been highlighted in a recent study of the use of SNSs by students in India. Agarwal and Mital (2009) highlight that students use SNSs for three purposes: networking (e.g. using LinkedIn for better job prospects, better understanding of the business environment, and career building); and socialising(e.g. using Bebo to make plans with friends/contacts and using Facebook to make new friends, sharing opinions, staying in touch with friends /family/contacts/strangers, etc). The last two of these have already been reported by Donath and boyd (2004), Donath (2007), Ellison et al., (2006) and Joinson (2008). The finding that students use SNSs to improve their prospects is interesting, particularly in a different culture. This finding may be due to the fact that Indian students (who were study participants) wanted to explore career prospects and study options in Europe or US and think SNSs will help them find useful information, since overwhelming majority of SNS users are based in Europe or US.

Organisations are also increasingly using social network sites. They are integrating SNS applications into their corporate environment to improve marketing and overall performance of business (Li and Bernoff, 2008). SNSs help businesses to find the right employees, to market their products, to find a manufacturer or supplier (Swearingen, 2008). Companies are now using public profiles of potential employees to scrutinise their job applications (Rosenblum, 2007). For example, an officer of a company states *"You really do get a lot of information you can't ask for in the job interview, but you go on the Web and it's all right there"* (Belluck, 2006). However, the use of personal profile information by employers raises many questions about the privacy of SNSs users, as the content to which employers might object (cheeky pictures and remarks about their past) can be extremely damaging. For instance, Rosenblum (2007) refers to the case of a promising job applicant whose application was blocked because the hiring officer found through web chat that the applicant's interests lay in "smoking blunts" and "obsessive sex", even when the hiring officer understood that the remarks might have been the result of Net posturing. For users, this is unacceptable because they do not expect their personal data to be used by third party organisations. While most employers justify their practices because of the public nature of SNSs (Rosenblum, 2007), most SNSs users lack real awareness of the privacy risks associated with sharing and dissemination of personal information. However, there has been a recent shift towards greater user awareness regarding the harm that can be done through privacy leakage (boyd and Hargittai, 2010).

2.2.4 Business Model of Social Network Sites

SNSs depend on three business models for revenues which include: advertisements (e.g., Facebook), micropayments via the social function of gift exchange (e.g. Cyworld) and premium membership fee for enhanced features (e.g. flickr, mixi) (Thelwall, 2009). SNS data is rich in both quantity and quality as more than one billion people not only disclose their personal information and demographic information, but also share their preferences (what they like and what they don't like). Morrissey (2009) acknowledges the importance of SNS data for advertisers when he remarks that the *"universe of social network sites*

presents a tempting pool of data for advertisers to use in order to improve targeting techniques" (p.1). Similarly, Thelwall (2009) recognises the potential of using SNS users' data for targeted advertising and suggests that advertising will remain the dominant source of revenue for SNSs. Facebook's privacy policy also acknowledges that advertising is the primary source of revenue support for its business (Facebook, 2010).

Realising the potential of users' data, SNSs are building alliances with third-party companies called advertising networks (Krishnamurthy and Wills 2008; Preibusch et al., 2007). Advertisement networks (e.g. DoubleClick) include those companies which provide content and advertisements to first-party websites (e.g. Facebook) (Krishnamurthy and Wills; 2010). The data of SNS users also provide search engines with the opportunity to improve their search techniques (Zimmer, 2008). Furthermore, SNSs are building advertising networks with bigger companies including Google and Microsoft (Johnston, 2007). For example, Microsoft bought 1.6% stakes in Facebook by spending $240 million which gave Microsoft exclusive rights (only in the U.S.) to publish banner and sponsored links on Facebook. In 2006, Google also signed a similar advertising deal with MySpace (Johnston, 2007). Therefore, advertisers are using new ways to exploit data of SNS users for better customer relationship management (Bernoff and Li, 2008). However, the interaction of SNSs with third parties (both companies and advertisers) has raised privacy concerns amongst SNSs users. For instance, the social marketing tool Beacon was withdrawn by Facebook because of user backlash on privacy grounds (Jamal and Cole, 2009).

2.2.5 Information Disclosure Characteristics in SNSs

Disclosing personal information in SNSs has inherently contradictory results: on the one hand, such disclosure is essential for the use of SNSs facilities; on the other hand, it renders users vulnerable to several real-world and cyberspace threats (such as identity theft, stalking, personalised spamming) if their personal information is leaked to unintended parties such as users, SNSs and third party sites (Acquisti and Gross, 2006; Lam et al., 2008). Also discussed earlier in Chapter one, digital information is persistent, searchable, copiable and invisible. Together, these properties create a serious threat to

user's privacy given the availability of sophisticated data mining and tracking tools which can link behavioural data with profile information and thus make it possible to uniquely identify a person (Bonneau et al.,2009a; Krishnamurthy and Wills,2010). Social network site users' data can be disclosed voluntarily or involuntarily. Voluntary disclosure of information occurs when a user shares photos and other information with friends and family, whereas involuntary disclosure is characterised by the leakage of information by other users (Lam et al., 2008) or third parties (e.g. SNS service providers, advertisers and/or aggregators) (Debatin et al., 2009; Krishnamurthy and Wills, 2010).

Voluntary Information Disclosure: The majority of SNS users disclose real personal information voluntarily (Gross and Acquisti, 2005). In a study of 4000 students in a US University, most students were found disclosing personal data on their Facebook profiles voluntarily. Specifically, 89% SNS users used their real name on profiles and 90% uploaded a profile image (Figure 2.3 highlights other types of information). Such a high information disclosure of personal information by students on Facebook profiles may be because of the information disseminating design of Facebook and its perceived connection to a physical and a bounded community (Gross and Acquisti (2005). The connection of Facebook with physical entities is also confirmed by Dwyer (2007a). Similar findings are reported in the studies of Acquisti and Gross (2006) , Barnes (2006) , Govani and Pashley (2005) and Strater and Lipford (2008).

Another interesting dimension of information disclosure relates to the effect of gender on information disclosure. No major difference in information disclosure by men and women was found, except that 47% males disclosed phone numbers compared to 28.9% of women who disclosed their phone numbers on profiles (Gross and Acquisti, 2005). In contrast, study of teenagers reveal that boys and girls have different information disclosure behaviour (Lenhart and Madden; 2007). For example, girls seem to share their own photos as well as their friends' photos, whereas boys seem to share their address, last names and phone numbers on their profiles.

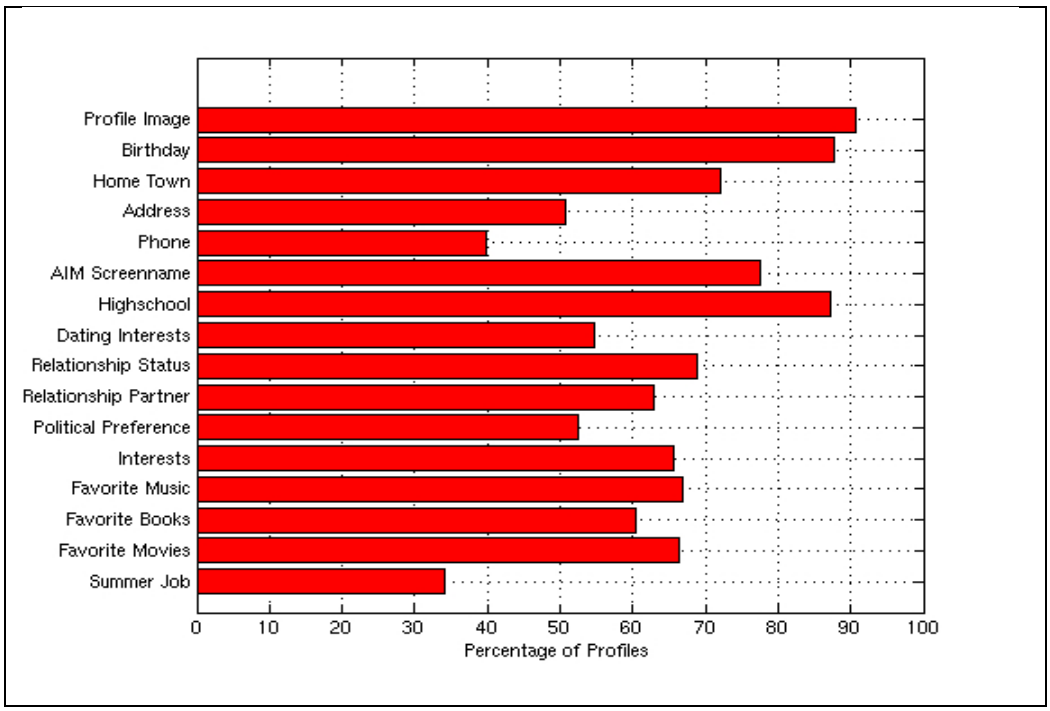


Figure 2-3: Disclosure of Personal Information on Facebook Profiles
(Source: Gross and Acquisti; 2005)

This discrepancy of research findings suggests another characteristic i.e. age since both the study participants were in different age groups. Participants in the Gross and Acquisti (2005) study were in age group 18-24, whilst those in the later one were aged 12-17. Similarly, older teens (15-17) seem more likely to share personal information, that is why there was no significant difference between disclosures of overall information between the two groups (teens and adults) (Lenhart and Madden, 2007). (Table 2.1 highlights this comparison).

Comparing information disclosure between young adults (18-24) and teenagers (12-17)		
Information item	Adults (Source: Gross and Acquisti;2005)	Teens (Source: Lenhart and Madden;2007)
Real Name	89%	82%
Profile Image	90%	79%
Date of Birth	87%	Not available
Home Town	72%	61%
Home Address	51%	Not available
Phone Number	40%	Not available
School Name	87%	49%
AIM Screen Name	77%	40%

Table 2.1: Comparing Information Disclosure between Adults and Teenagers

Likewise, Campbell et al., (2001) and Fogel and Nehmad (2009) studied the effect of gender on information disclosure. Accordingly, males were more aware of privacy

concerns compared to females but they were involved in more risky behaviour (Campbell et al., 2001). Explaining this behaviour Campbell et al., 2001) argue that because men have better awareness of privacy issues and perhaps more awareness of privacy protection methods. Similar finding was discussed by Acquisti and Gross (2006) that people who demonstrate better understanding of privacy share large amount of information on their profiles. However, other studies reported contradictory results that women have more awareness of privacy issues than men and are less likely to divulge personal information (Lewis et al., 2008). Similarly, Acquisti and Gross (2006) found that women are less likely to disclose their sexual orientation, personal information and phone numbers.

Another interesting finding is that culture influences information disclosure as U.S. users were found to disclose more data than other cultures, and hence require more privacy controls (Lenhart, 2009). This has implications for SNS service providers as they have to offer users stronger privacy controls as well as put in place more powerful security measures to protect the confidentiality of personal data.

Involuntary Information Disclosure: The involuntary disclosure of information, also called privacy leakage, happens when other users (mostly online friends) disclose SNSs users' personal information online (Lam et al., 2008) or when SNS providers leak (either intentionally or unintentionally) SNS users' personal information to third-party advertisers or when third-party advertisers pass on SNSs users' personal information to other third parties mostly without their consent and without them realising (Krishnamurthy and Wills, 2010). Lam et al., (2008), in an investigation of 592,548 profile of Wretch (a popular SNS in Taiwan), show that 78% of users were subject to involuntary name leakage by their friends through annotations. Another instance of privacy leakage concerns the passing of names, friends' list, and other profile information by SNSs to third-party advertisers and online trackers (Krishnamurthy and Wills, 2010). Krishnamurthy and Wills (2010) in their investigation of 12 SNSs[3] demonstrate that all SNSs with the exception of Google-owned Orkut leak users' personally identifiable information to third-party advertisers and/or aggregators.

[3] Bebo, Digg, Facebook, Friendster, Hi5, Imeem, LiveJournal, MySpace, Orkut, Twitter, Xanga and LinkedIn

2.2.6 Information Privacy Concerns in SNSs

Various studies have highlighted privacy concerns within the context of online social networks. A few privacy concerns are discussed here.

1. ***Identity Theft:*** SNS users' disclose personally identifiable information (PII) on their profile in order to use SNS service. Therefore, leakage of rich in detail personally identifiable information of SNS users pose a threat called identity theft (Krishnamurthy and Wills, 2010). With the advancement in technologies, it is now possible to identify a person from even non-identifiable information by combining data elements. As the boundary between PII and non-PII becomes ever more indistinct, the problem of identity theft becomes even more complicated. Therefore, all stakeholders such as SNS users, network service providers, regulators and SNS designers should work together to combat this threat.

2. ***Physical and Online Stalking:*** The proliferation of SNSs has raised concerns about stalking, both physical and virtual (Gross and Acquisti, 2005; Krishnamurthy and Wills, 2010). Stalkers can determine the actual location of a user for larger portion of the day, cyber-stalking is a particular threat to users of SNSs as they disclose rich in detail personal data which stalkers can use for wrong doings. However, Lenhart and Maddel (2007) suggest that male teenagers are more likely to avoid this threat by falsifying information.

3. ***Constructing a Digital Record:*** SNS users share excessive data (e.g. their names, location, phone numbers, political interests, etc.) on SNSs (Gross and Acquisti, 2005) that helps SNSs and marketers of third parties construct rich in detail profile. Because digital data is permanent in nature as it cannot be deleted by SNS users once stored on companies servers, thus threats to users' privacy become more serious as the indefinite retention of information increases the possibility that the profile information may get into the wrong hands. Moreover, the identifiable profile information of SNS users pose further harms to users when their identity is stolen.

4. ***Secondary Privacy Diffusion:*** An entirely new privacy concern, not discussed in any privacy studies, called "secondary privacy diffusion" (SPD) is mentioned in the study of Krishnamurthy and Wills (2009). Secondary privacy diffusion happens when one person disclose personal information of another either intentionally or unintentionally. An example of such privacy diffusion is where one person give email addresses of another person to a company because the company may be using email address without the consent of actual user. The researchers admit that it is not easy to avoid the damage from SDP before it occurs. However, they provided a possible solution involving post-leakage notification: the user would be notified when it had occurred and the source website would stop propagating information further and avoid further damage. Unfortunately, their proposed solution would involve tracking other websites, SNSs to check if leakage had happened, which would incur huge costs. Also, the method itself require to track websites and SNS, hence can be problematic. While Krishnamurthy and Wills (2009a) only refer to secondary privacy diffusion damage caused by information leaked by other users, the author argues that instances of information leakage by SNSs operators and third parties should also be included. To date, there is no literature reported on this new privacy issue.

2.5 Privacy Leakage in Social Network Sites

Contrary to user expectation that personal data is only accessible to friends and the SNSs, other entities such as third-party advertisers and data aggregators, as well as social network users who are not their friends and third-party applications all have access to private bits of information (Krishnamurthy and Wills, 2008; 2010). Moreover, third-party companies (including SNSs) also track SNS users' behaviour and actions on external websites and not only use that data for business marketing but publish it to their profiles, often without their knowledge, thus causing personal embarrassments for SNS users (Krishnamurthy and Wills, 2008; 2010). Also, SNSs leak personally identifiable information (PII) which third-party sites are able to link with user actions performed within and outside SNSs (Krishnamurthy and Murthy, 2010). Krishnamurthy and Wills (2010) refer to this ability to linking PII to behavioural data as leakage.

Within the context of this thesis, privacy leakage refers to sharing of personal data of SNS users with third parties without their explicit consent and knowledge. Such privacy leakage occurs as a result of the integration between SNSs and third parties. Recently, Facebook has launched social marketing tools which share Facebook users' personal information with third party sites (e.g. Blockbuster) and behavioural data of users is pulled into the social network from third party websites. As this sharing happens without explicit consent and knowledge of SNS users, this is referred to as privacy leakage. Often, the leaked information is able to identify a person and hence has provoked privacy concerns amongst SNS users. The leakage of personally identifiable information (PII) raises grave concerns about users' privacy because it may lead to privacy harms such as stalking and identity theft (boyd and Ellison, 2006; Rosenblum, 2007). An example of such privacy leakage was Facebook's social advertising tool, Beacon, which tracked SNS users' action (e.g. renting a movie) on third-party websites (e.g. Blockbuster) and shared that information back to their Facebook friends even without their knowledge. The alarming part of Beacon (for SNS users) was that it tracked Facebook users' behaviour even when they had opted out of it and logged off Facebook (Perez, 2009).

Personally identifiable information includes first name, last name, address (street, city, zip), email address, telephone numbers, and photos (both personal and group) (Krishnamurthy and Wills, 2010). Even non identifiable data (e.g. age, gender, birthday, year of birth, educational institution, and employer name when joined to other parts of PII e.g. name or email address, can also personally identify an SNS user (Krishnamurthy and Wills, 2010). A complete list of PII compiled from McCallister et al. (2009) is shown in Table 2-2.

Personally Identifiable Information and examples	
PII Item	**Examples**
Name	Full name, maiden name, mother's maiden name, or alias.
Personal Identification Number	SSN, passport number, driver's license number, taxpayer identification number, patient identification number, and financial account or credit card number.
Address Information	Street address or email address.
Asset Information	Internet Protocol or Media Access Control address or other host-specific persistent static identifier that consistently links to a particular person or small, well-defined group of people.
Telephone Numbers	Mobile, business, and personal numbers.
Personal Characteristics	photographic image especially of face or other distinguishing characteristic, x-rays, fingerprints, or other biometric image or template data such as retina scans, voice signature, facial geometry.
Information Identifying Personally Owned Property	Vehicle registration or identification number and title numbers and related information.
Information about an Individual that is Linked or Linkable to one of the Above	Date of birth, place of birth, race, religion, weight, activities, or employment, medical, education, or financial information.

Table 2.2: Examples of Personally Identifiable Information
(Source: McCallister et al., 2009:p.2)

An interesting finding relating to USA users shows that date of birth , zip code and gender could be used together to identify 87% Americans (Sweeney, 2000). It is because this so called anonymous when combined will reduce the population to a large extent that it uniquely identifies a majority (87%) of Americans (Schoen, 2009). Similarly, in a recent study, users were identified from their web browsing behaviour such the websites visited (Narayanan and Shmatikov, 2008), this consequently refutes the claim by organisations (at least theoretically) that they only collect user behaviour which cannot identify a person.

These findings suggest that advances in de-anonymisation or re-identification techniques calls into question the old promise that user data which have been collected for decades by online trackers would be kept anonymous (Ohm, 2009). PII is treated as highly sensitive information and protected in the USA law by, for instance, the Federal Telecommunications privacy laws (U.S. Code, 2008), the U.S. based Health Insurance Portability and Accountability Act (HIPPA, 1996) and the EU Data Protection Directive (1995). This contradictory situation calls for a comprehensive review of existing data protection laws, avoidance strategies by online trackers and adoption of safe browsing habits by users (e.g. deletion of cookies, providing false information, avoiding

information disclosures). Surprisingly, most SNS users are not aware of privacy leakage and more than one billion SNS users' personal information is being used for commercial purposes without their knowledge or consent.

Although the aforementioned studies have highlighted the issue of privacy leakage, little is yet known about how users perceive this leakage of personal information, which usually occurs as a consequence of commercial organisation practices, such as passing information to advertisers. Additionally, the most worrying thing for SNS users is that such leakage of private information is often hidden or secret, giving them little control to mitigate the privacy risks associated with such leakage (Bonneau et al., 2009b; Krishnamurthy and Wills, 2008).

2.6 Privacy Evolution

Privacy has long been recognised as a moral phenomenon in the history of humankind. The value of privacy in the modern global information societies remains undisputed (Xu et al., 2011) as *"[it] has become one of the most important human rights of the modern age"* (Rotenberg, 2000). Whilst privacy is a recognised as a moral issue, even so it remains difficult to define (Michael, 1994). Primarily because it is a complex (Solove, 2006), discipline-dependent (Xu et al., 2008), context-specific (Ajzen and Fishbein, 2005; Nissenbaum, 2004) and temporal concept (Palen and Dourish, 2003). Accordingly, various discipline-specific conceptualisations of privacy emerge. For instance, in law privacy is considered as a *'right' or 'entitlement'* (Warren and Brandeis, 1890), and in information systems and social sciences as a *'control'* (Culnan, 1993; Westin, 1967). Furthermore, the fragmented definitions, relationships and concepts of privacy are inconsistent, underdeveloped, and not empirically validated (Xu et al., 2011).

For instance, the application of the notion of 'privacy as a right' in consumer behaviour results in a privacy paradox because consumers have been disclosing their personal information despite serious privacy concerns (Smith et al., 2011). This causes a shift in 'privacy as a right' to 'privacy as a commodity' (Bennet, 1995) with an assigned economic value to be determined in a cost-benefit relationship (Smith et al., 2011). Similarly, the widely referred notion of privacy as control over personal information

(Altman, 1975; Westin, 1967) emerged within information systems and social sciences. However, some researchers argue that control does not equate with privacy, but rather it is one of the key elements conceptualising privacy (Margulis 2003). Similarly, Malhotra et al., (2004) argue that control over information within the internet environment can implemented by consent, and opt-in or opt-out mechanism. But the question arises: can control in any other environment and context be exercised via these three factors (i.e. approval, modification, opt-in or opt-out)? The answer lies in another privacy concept: that it is context-specific and ideally should be studied within a particular context (Ajzen and Fishbein, 2005; Nissenbaum, 2004). This further complicates the situation and Solove (2006) rightly sums up this situation as confusing: *"privacy is a concept in disarray. Nobody knows what it means"* (p.477).

Although various notions of privacy exist, in the context of this thesis the author will use the term information privacy as it is the most relevant to the aim of this research. Information privacy is referred to as "the claim of individuals, groups and institutions to determine for themselves when, how and to what extent information about them is communicated to others" (Westin, 1967:p.7). Similarly, Stone et al. defines it as *"the ability of the individual to personally control information about one's self"*. Boyle and Greenberg (2005) refer to it as a process of controlling personal information. However, the widespread adoption of technologies has made it difficult for technology users to control their personal information (Dinev and Hart, 2006; Hui et al., 2007; Malhotra et al., 2004; Solove, 2006). Smith et al., (2011), whilst adapting the work of Westin (2003) argue that information privacy is associated with the evolution of technology and divide it into four eras (see Table 2-3).

Evolution of the Information Privacy Concept Following the Evolution of IT	
Privacy Baseline 1945-1960	Limited information technology use, high people trust in government and businesses, and general comfort with the information collection.
First Era of Contemporary Privacy Development 1961-1979	Rise of information privacy as an explicit social, political, and legal issue. Early recognition of potential dark sides of the new technologies (Brenton, 1964), formulation of the Fair Information Practices (FIP) Framework and establishing government regulatory mechanisms established such as the U.S. based Privacy Act of 1974.
Second Era of Privacy Development 1980-1989	Rise of computer and network systems, database capabilities, federal legislation designed to channel the new technologies into FIP, including the U.K. based Data Protection Act of 1984. European nations move to national data protection laws for both the private and public sectors
Third Era of Privacy Development 1990-present	Rise of the Internet, Web 2.0 and the terrorist attack of 9/11/2001 dramatically changed the landscape of information exchange. Reported privacy concerns rose to new highs.

Table 2.3: Evolution of Privacy Concept and Legislation
(Source: Smith et al., 2011:p.991)

Interestingly, there is a direct relationship between evolution of privacy and information technologies, suggesting that advances in technologies have brought new challenges to the management of privacy and hence heightened privacy concerns. As in, the third era of privacy, people happen to report higher privacy problems compared to in the second era. This is not surprising, as organisational needs provide an impetus to use innovative technologies (often to exploit personal information for commercial purposes) and hence cause concerns amongst online users.

2.7 Information Privacy Concerns (IPC)

Information systems researchers are increasingly adopting the concept of privacy 'concerns' to measure privacy related issues (Xu et al., 2011). This approach is consistent with the suggestion of Solove (2008) and Phelps et al., (2000) who argue that instead of choosing an abstract concept of privacy, it is better to investigate privacy issues by identifying and studying the root causes of privacy concerns. The majority of IS studies using a privacy 'concerns' construct, concentrate on identifying a link between privacy concerns and behavioural variables, such as the willingness to disclose personal information (Chellappa and Sin, 2005) and the intention to transact (Dinev and Hart, 2006). Surprisingly, the most of IS studies were positivists and hence fail to provide any depth of insight into the subjective views of people. Also, the majority of studies relied on general privacy concerns rather than on a context-dependent situational construct. However, as privacy is a contextual phenomenon and user behaviour can be better investigated within a particular context (Margulis, 2003; Solove, 2008; Xu et al., 2011), therefore, the research reported in this thesis adopted a context-specific concept of privacy concerns by using definitions proposed by Smith et al., (1996); Son and Kim (2008) and Campbell (1997). Smith et al., (1996) and Son and Kim (2008) define IPC as the extent to which online users show concerns about the collection and use of personal information by online companies. Whereas, Campbell (1997) refers to IPC as the subjective view of users regarding the fairness of information practices of businesses. Within this thesis, both conceptualisations seem fitting and hence are combined to represent privacy concerns as: the subjective views of social network site users about the fairness of business practices relating to the use of personal information.

2.7.1 Privacy Concerns Frameworks

The focus of past research on information privacy has been to determine those factors that motivate the divulgence of personal information and those that inhibit it (Dinev and Hart, 2006; Son and Kim, 2008). Most of frameworks adopted privacy concerns to deconstruct privacy. Realising that information privacy is not a single dimensional concept based on control over personal information, most of the information privacy frameworks conceptualise privacy as a multifaceted phenomenon. For example, Smith et al., (1996) developed a multidimensional scale of concern for information privacy (CFIP). It was designed to record individuals' concerns about organisations' practices. CFIP consisted of four dimensions and 15 items. The four dimensions were: *collection*, *unauthorised secondary use*, *improper access*, and *errors*. However, CFIP was used mostly offline or in the context of direct marketing (Malhotra et al., 2004). Stewart and Segars (2002) empirically tested CFIP with 355 participants in an online context. However, Malhotra et al., (2004) argued that as the context of direct marketing was supported by one directional communication hence CFIP could not be effectively used in the context of the internet which is characterised by two-way communication.

Malhotra et al., (2004) devised a framework to measure internet users' information privacy concerns (IUIPC). They conceptualised privacy as a multifaceted phenomenon and presented privacy framework called IUIPC. IUIPC comprised of three elements: collection, control and user awareness. Malhotra et al. (2004) applied this theoretical framework in conventional e-commerce environment which are featured on mandatory information disclosure (Nov and Wattal, 2009), hence other contexts, such as online social networks where information disclosure is voluntary, cannot benefit. Also, in SNSs information flow is much more complicated than in a conventional two-way communication. For example, information within SNSs flows: from user to an SNS; from one user to another; from user to third-party trackers (mostly without user knowledge) and from service providers to advertisers.

However, most of the aforementioned privacy frameworks exist in conventional online context where information disclosure is mandatory and usually happen with the consent

of users. However, these frameworks are not effective in the context SNSs where data is secretly collected and shared with third-party companies without users' knowledge. Surprisingly, given the richness of SNS user data and the associated privacy concerns, no privacy framework exists to date which could be used to conceptualise the privacy leakage concerns of SNS users.

2.7.2 Privacy Taxonomy

Another stream of research concentrates on the need to devise a taxonomy of privacy concerns. As Etzioni (1999) suggests, the first step towards addressing any privacy issue is to identify the activities which can cause a privacy violation. Likewise, Solove (2006) suggests the need to have a comprehensive taxonomy of privacy which identifies all the activities posing threats to users' privacy. The researcher agrees with Solove's (2006) thinking that in order to address privacy problems effectively we need first need to find out the root cause of the problem. Accordingly, Solove (2006, 2008) devised privacy taxonomy (see Table 2.4) that focuses on activities that cause privacy problems divided into four main groups: information collection, information processing, information dissemination and invasion. It provides a useful mechanism to understanding the activities that cause privacy problems, as well as how they do so.

As Solove (2006) suggests, individuals, institutions and governments all engage in activities that conflict, thus causing social friction. Consequently, the relief from such social friction ensures users' privacy. For example, in SNSs, third-party organisations track users without their consent, thus posing threats to their privacy because SNS users perceive such tracking as invasive and a source of social friction (Solove, 2008; Solove, 2006). Though Solove's (2006) taxonomy improves our understanding of the nature of privacy concerns by identifying the activities which are perceived as privacy violations, yet the taxonomy conceptualises common privacy problems for people (such as surveillance, aggregation, secondary use, disclosure, intrusion, etc.) rather than focusing on privacy leakage concerns of SNS users, thus highlighting the need to develop a comprehensive taxonomy of privacy leakage concerns of social network users.

Privacy Taxonomy	
Principle group activity	**Activities creating privacy problems**
Information collection	**Surveillance** - Involuntary gathering of information , mostly without the awareness of subject **Interrogation**- Involuntary gathering of information with conscious awareness
Information processing	**Aggregation** - The gathering together of information about a person. **Identification**- It enables us to attempt to verify identity of a person. **Insecurity** - Problem caused by the way our information is handled and protected. **Secondary use**-The use of data for purposes unrelated to the purposes for which the data was initially collected without the data subject's consent. **Exclusion** - The failure to provide individuals with notice and input about their records.
Information dissemination	**Breach of confidentiality** -Is not simply that information has been disclosed, but that the victim has been betrayed because of the violation of trust in a relationship. **Disclosure** - Also protects relationships of trust, but disclosure must result in the release of embarrassing secrets or discrediting data before courts will consider it to be harmful **Exposure** - Involves the exposing to others of certain physical and emotional attributes about a person. **Increased accessibility** -The existence of information in a government database can increase the potential accessibility of that information. **Blackmail** - **The threat to disclose personal information which creates power imbalance in a relationship.** **Appropriation** - Use of one's identity or personality for the purposes and goals of another. **Distortion** - Consists of the dissemination of false or misleading information about individuals.
Invasion	**Intrusion** - Concerns invasive acts that disturb one's tranquillity or solitude. **Decisional interference** Involves the government's incursion into the data subject's decisions regarding her private affairs.

Table 2.4: Privacy Taxonomy
(Solove, 2008; Solove, 2006)

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Chapter Overview

This chapter details and justifies the research approach used to investigate user perceptions to privacy leakage in the context of online social network sites. It starts by discussing the challenge and importance of finding a fitting approach in information systems (IS) research. The next section contrasts different research paradigms (philosophical perspectives) employed in IS research. Thereafter, the rationale for adopting qualitative research is discussed. This rationale is followed by a discussion on selecting a case study based research strategy and the selection of cases. Then, the rationale for using blogs data along with related ethical issues is discussed. Finally, the data analysis method and process used in the current research is examined.

3.2 Choosing a Fitting Research Approach in IS

As Walsham (1995a) notes, selecting an appropriate research approach is a major task in the research design process. In IS research, this task becomes even more challenging given the multidisciplinary nature of IS (Land, 1992), to which many disciplines including engineering, mathematics, natural and behavioural sciences contribute. Moreover, different research methods suit different situations (Galliers, 1992) as they focus on different aspects of reality (Mingers, 2001). This further complicates the decision-making process of researchers as to which research approach (es) to choose in a certain context. For example, the case study research method is a viable approach to answer 'why' and 'how' questions, which enable researchers to understand the nature and complexity of the processes taking place often in real life setting (Benbasat et al., 1987; Yin, 2008). Also, Galliers (1992) argues each research methodology has its own strengths and weaknesses.

Given the multidisciplinary nature of IS research, utilising diverse research approaches seems attractive. However, there are differing viewpoints amongst IS researchers whether such diversity is beneficial for the discipline. For instance, Robey (1996) argues that such

diversity of research methods and paradigms in IS research strengthens the discipline since variety offers flexibility, and that promotes creativity. Likewise Mingers (2001) suggests that combining various research approaches and methods improves the value and benefits of the research. Additionally, Robey (1996) argues that different research paradigms and methods are suitable for answering different types of research questions, investigating different types of research phenomena, and supporting the different philosophical stances of researchers. In contrast, others express concerns about utilising diverse research approaches and methods in IS research – Benbasat and Weber (1996), for instance, solicit for the uniformity of a research paradigm in information systems research , fearing that, the absence of such uniformity would make the discipline vulnerable to be taken-over by a more established discipline. But in recent years, IS researchers have increasingly adopted research methods from other disciplines – the growing use by IS researchers of the grounded theory method (GTM) that originated in sociology is one such example (Hughes and Jones, 2003; Urquhart et al., 2010).

Orlikowski and Baroudi (1991) caution that in order for the research to benefit from pluralistic research methods, these must be applied appropriately and effectively, because pluralist methodology based on different ontological beliefs usually lacks a solid basis (Cavaye, 1996). Therefore, understanding and awareness of different research paradigms and methods is critical if researchers are to adopt a fitting research approach. Understanding of different research paradigms and methods enables researchers to remain open to using all research approaches and methods and thus helps reduce any bias towards using a particular approach (Orlikowski and Baroudi, 1991). Reflecting on the aforementioned discussion, the author argues that the type of research questions and the context of research are the key considerations informing the choice of appropriate research approach (es) and method(s). Similarly, Cavaye (1996) suggests that researchers should choose methodology based on what they want to achieve without committing themselves to a particular research approach.

3.3 Research Approaches in IS and Justification of Choice

3.3.1 Research Approaches in IS

The terms paradigm, methodology, method and technique have been used differently by different researchers. Mingers (2001) and Guba and Lincoln (1994) provide a good distinction of the terms: *paradigm* consists of a general set of philosophical assumptions such as *ontology* (what is assumed to exist); *epistemology* (what is valid knowledge - the nature of valid knowledge), and *methodology* (a set of activities whose aim is to produce valid and reliable research results). Mingers (2001) also considers *ethics or axiology* (what is valued or considered right) an important element of any research approach. Also, the distinction between a method and methodology is unclear. As Mingers (2001) reports, some researchers have simply used these terms interchangeably, such as Tashakkori and Teddlie (1998) and Livari et al., (1998). Also, these terms are used differently in North America to how they are used in Europe (Livari et al., 1999). However, within the context of this thesis, Mingers's (2001) standpoint that methodology is general and less prescriptive than a method is adopted. Mingers (2001) refers to methodology as a set of guidelines or activities which assist in producing reliable and valid results within a certain research paradigm, whereas according to Strauss and Corbin (1998), a research method consists of a set of procedures and techniques used to gather and analyse data.

Therefore, a research methodology may often contain various research methods. IS research is classified into three major research paradigms, namely positivist, interpretive, and critical (Chua, 1986; Orlikowski and Baroudi, 1991). A brief description of these three research paradigms is as follows (see Table 3.1 for summary).

- Positive research is concerned with formal propositions, quantifiable measurements of variables (dependent and independent), testing of hypothesis, and drawing inferences to arrive at conclusions about the phenomenon understudy from the sample representing a research population (Orlikowski and Baroudi, 1991; p.5). Thus, positivists assume that reality is objectively given and can be measured independently of the researcher or his/her instrument (Myers, 1997).

- Interpretive research, on the other hand, is concerned with the assumption that knowledge of reality can be gained through social constructions such as consciousness, shared meanings, language, documents, tools and other artefacts. As Walsham (1995a) argues, interpretive research is "*aimed at producing an understanding of the context of the information system, and the process whereby the information system influences and is influenced by the context"* (p. 4-5). Therefore, unlike positivist research, interpretive research does not predefine dependent and independent variables; rather, it allows researchers to focus on the complex way in which humans make sense of situations as they emerge (Kaplan and Maxwell, 1994).

  Finally, critical research is concerned with the assumption that *"social reality is historically constituted and that it is produced and reproduced by people"* (Avison and Pries-Heje, 2005: p. 244). Therefore the ability to change peoples' social and economic conditions is constrained by different kinds of social, cultural and political domination (Myers, 1997).

Primary Beliefs	**Research Paradigms**		
	Positivist	**Interpretive**	**Critical**
Ontology	Reality is objective and exists independent of humans.	Reality is subjective or inter-subjective and exists only through human actions.	Reality is historically constituted.
Epistemology	Knowledge and values are distinct. Valid knowledge exists as it is and can be measured independent of the researcher and its instrument.	Knowledge and values are interwoven. Knowledge is socially constructed through language and shared meaning	Knowledge and values are interwoven. Knowledge is founded in social and historical practices
Axiology	Deterministic explanation and prediction of reality.	Descriptive and situated understanding of a phenomenon.	Aims to initiate change. Totality of a phenomenon.
Methodology	Quantitative (e.g. surveys and controlled experiments sources of data) and use of statistical data analysis techniques.	Field studies, qualitative, hermeneutics and phenomenology.	Longitudinal historical studies and ethnographic studies, historical analysis

Table 3.1: Summary of Research Paradigms in IS
(Based on Orlikowski and Baroudi, 1991)

3.3.2 Justification of the Choice of Research Approach

Considering various perspectives of research paradigms in information systems and the nature of this research study, the interpretive approach seems appropriate to study users' perceptions of privacy leakage in social network sites. The following reasons informed the researcher's choice:

- First and foremost is the aim of this study is to understand users' perceptions of privacy leakage in the context of SNSs. The essence of this understanding is based on the subjective meanings users assign to privacy concerns. The interpretive approach which is concerned with gaining knowledge of reality through language, consciousness, shared meanings, documents, tools, and other artefacts (Deetz, 1996; Klein and Myers, 1999; Myers, 2009; Orlikowski and Baroudi, 1991) seems appropriate to study the issue of privacy leakage in SNSs, because it enabled the researcher to understand privacy leakage through the words used by SNS users' when they shared and described their experience regarding leakage of their privacy in SNSs.

- Chapter two highlighted the gap in knowledge related to privacy leakage in social networks that Krishnamurthy and Wills (2008, 2010) recommended studying holistically in order to understand the bigger picture. This research, therefore, focuses on not only identifying privacy concerns related to the leakage of SNS users' information which occurs as a consequence of organisational practices, but also on how and why these cause privacy concerns. Thus, interpretation, which is the basic premise of the interpretive approach, is seen as an appropriate approach to study the phenomenon of privacy leakage in totality and holistically, rather than looking at its parts.

- Third, privacy is a complex social phenomenon which is highly contextual (Altman,1975; Garde-Perik et al., 2008) and the attitudes expressed and behaviours exhibited outside the context are not good indicators of privacy behaviours (Ajzen and Fishbein, 2005; Garde-Perik et al., 2008). As Orlikowski and Baroudi (1991) argue *"..... social process can be usefully studied with an interpretive perspective, which is explicitly designed to capture complex, dynamic,*

social phenomena that are both context and time dependent"(p.20). Thus, the interpretive approach seems fitting for the study of the complex, context-dependent, and time specific nature of information privacy.

- Fourth, because SNSs have emerged recently and privacy leakage is a recent phenomenon, little is known about how SNS users' perceive leakage of their privacy which occurs as a consequence of organisational practices. Hence, this study could not take advantage of existing constructs and variables. Therefore, the interpretive approach was seen as relevant as it does not require predefining variables, as does the positivist approach, but rather focuses on the complex nature of how humans make sense of emerging phenomena (Kaplan and Maxwell, 1994).

- Finally, Dwyer (2007a, 2007b) argue that *"privacy within SNS is [usually] . . . . .undefined"* because it is a complex social phenomenon and different people perceive privacy differently in different contexts. Therefore there is a need to follow a research approach which enables the researcher to interpret subjective understanding of reality from users' perspectives. Accordingly, the interpretive approach is adopted.

3.4 The Rationale of Using Inductive Qualitative Research Method

A research method is the strategy by which researchers move from philosophical assumptions to research design and data collection (Myers, 1997). Specifically, it provides guidelines (e.g. about data collection) which assist researchers to advance towards achieving a research goal. Therefore, the choice of the selection of a particular research method is mainly dependent on the aim of a research study. For instance, this research aims to understand user's perceptions of privacy leakage in SNSs and in the absence of established frameworks which could be used to examine these perceptions, a qualitative approach seems fitting. The rationale for using inductive qualitative research is presented below.

- Qualitative research offers a variety of techniques to interpret data in order to describe, decode, translate and discover meanings of a real world social phenomenon, rather than focusing on how many times an event or phenomenon

has occurred (Van Maanen, 1983). Therefore, a qualitative approach seems fitting for the study of real-world social phenomenon of privacy leakage in social networks.

- The major motivation for adopting an inductive qualitative approach is the epistemological assumptions this research has made that reality is subjective and socially constructed through language and shared meanings. Indeed, quantitative research methods stand in contrast (as aforementioned) to these epistemological assumptions and hence could not be adopted in the context of this thesis.

- Furthermore, the purpose and the focus of qualitative research methods is typically to understand a phenomenon (often complex) from the participants' point of view (Leedy and Ormrod, 2005). This fits well within the aim of this research, which is to understand user perceptions of privacy leakage in online social networks. Also, the researcher supports the standpoint of Kaplan and Maxwell (1994) who argue that *"the goal of understanding a phenomenon from the point of view of the participants and its particular social and institutional context is largely lost when textual data are quantified"*. Given the contextual nature of privacy leakage (as aforementioned), a qualitative inductive approach is adopted in this thesis.

- As online social networks emerged only recently and the phenomenon of privacy leakage has only been recently reported in research studies (e.g. Krishnamurthy and Wills, 2008, 2010), little is known about how SNS users perceive privacy leakage in SNSs which occur as a consequence of organisational practices. Rather, most privacy research in SNSs has focused on the impact of excessive disclosure of users (Acquisti and Gross, 2006; Barnes, 2006; boyd and Marwick, 2011). Thus, the qualitative research method seems appropriate. Also, this study could not take advantage of existing constructs and variables, which are the cornerstones of using quantitative research methods.

- Finally, Myers (1997) recommends using the qualitative approach in IS research to study social, managerial and organisational issues as it aims to understand such issues from the subjective views of the people concerned. Accordingly, many

qualitative research studies have been published in top IS journals, providing IS researchers (particularly junior researchers such as PhD students) with the opportunity to benefit from utilising established qualitative research guidelines. This research, therefore, adopts an inductive qualitative method to examine users' perceptions of privacy leakage in SNSs which is a social as well as organisational issue.

In short, this research, whilst keeping in view the research problem, adopted qualitative methods so as to benefit from the distinctive characteristics of qualitative methods as highlighted by Strauss and Corbin (1998, p.41): that qualitative methods provide suitable guidelines for understanding the meaning or nature of individual experience and enable researchers to gather complex details about the research phenomenon.

3.5 Case Study Research Method

A research method provides guidelines for researchers on how to design a research and collect data. Four commonly used qualitative methods in IS research include: *action research*, *case study*, *ethnography* and *grounded theory* (Myers, 1997). Yin (2002: p.13) defines case study as an empirical inquiry that examines a real-life phenomenon within its context, particularly when the contextual boundaries are not clear, while Merriam (1988) defines it as *"an examination of a specific phenomenon, such as a program, an event, a process, an institution, or a social group"* (1988, p.9). Creswell (1998) and Eisenhardt (1989) argue that case study research design enables researchers to study a phenomenon more holistically because of the real-world setting provides researchers with opportunities to understand almost all processes relating to a new phenomenon (Benbasat et al., 1987). It is one of the most appropriate research methods in IS research (Myers, 1997) because the IS studies often take place in real-life settings (e.g. the implementation of an automated system in an organisation). Case study design can be applied qualitatively and/or quantitatively (Stake, 1994; Yin, 2002), however, its use in IS research as a qualitative method is gaining much recognition (Orlikowski and Baroudi, 1991). Walsham (1995) further advocates the value of qualitative interpretive case studies in IS research.

Case studies can be single or multiple. Single case studies are selected as they are unusually revelatory, extreme exemplars, or opportunities for unusual research access (Yin, 2002). For example, Weick's (1993) study of the loss of sense making in the wilderness fire-fighting disaster at Mann Gulch in 1949 represents an extreme case. The unusual access gained through friends to study the New York Port Authority by Dutton and Dukerich (1991) is another example of a single case study. Hence, single-case studies attempt to explore significant phenomenon in unusual or extraordinary circumstances (Eisenhardt and Graebner, 2007), whilst multiple-case studies allow researchers to clarify that an emergent finding is not idiosyncratic to a single case but rather replicated consistently by several cases (Eisenhardt, 1991). Therefore, multiple cases provide much stronger foundation for theory building (Yin, 2008). Likewise, Eisenhardt and Graebner (2007, p.27) assert multiple cases build robust, generalisable, and testable theories since the propositions are deeply grounded in varied empirical evidence. The author argues that the choice of a single case or multiple case studies depends on the aim(s) and objectives of the research as well as the opportunities to collect data (research design). The motivations for using case study design, as well as the selection of case (s) are discussed below.

3.5.1 Motivation to use Case Study Research Strategy in this Thesis

The current research is inspired by the case study research method due to the following reasons:

- First and foremost is the nature of the phenomenon under study i.e. privacy leakage in social network sites. As highlighted in chapter one that nature is complex, context-specific, and temporal. Hence, the researcher argues that it should be studied holistically so that context-dependent and temporal conceptualisations of privacy can be explored via empirical data. As Creswell (1998) and Eisenhardt (1989) argue, case study research design enables researchers to study a phenomenon more holistically because real-world settings provide researchers with opportunities to better understand processes concerning a relatively new phenomenon (Benbasat et al., 1987). Therefore, the current research is motivated by case study strategy.

- Second, case study research strategy is more suitable to answer 'why' and 'how' questions than 'how much' and 'how many' (Benbasat et al., 1987; Yin, 2002). And the goal of the current research is to not only to identify organisational activities perceived as privacy violations by SNS users but also to discover how and why these cause privacy problems. Therefore, a case study research design seems appropriate.

- Also, online social networks emerged recently and little is known about what organisational activities cause privacy leakage and how and why they are perceived as privacy violations by SNS users. Therefore, this research adopted case study strategy to examine the issue of privacy leakage in a real-life setting in SNSs (Benbasat et al., 1987). Moreover, case study research strategy has the potential to build a new theory by better understanding the underlying processes (Benbasat et al., 1987). Indeed, case study research strategy seems to promise an improved understanding of the fundamental processes causing privacy leakage in SNSs. The case study provides rich sources of data to build new theories inductively, as new concepts or theories usually emerge from the analysis of data (Eisenhardt and Graebner, 2007).

- Finally, case study strategy in IS research has been widely applied, particularly as an approach to new theory building (Yin, 2008; Benbasat et al., 1987). Consequently, current research may be informed by the practical guidance obtained from studies employing case study strategy published in leading IS journals such as Beynon-Davies (1994), Bussen and Myers(1997), Cavaye and Cragg (1995), Lee (1994), Orlikowski (1993), and Walsham, and Waema (1994). Furthermore, current research may be benefit from theoretical guidance offered in studies such as Benbasat et al., (1987), Cavaye (1996), Eisenhardt (1989), Eisenhardt and Graebner, (2007); Walsham (1995b).

3.5.2 Case Study as a Theory Building Research Design

Case study is a useful strategy in theory building which involves using one or more cases to create theoretical constructs, propositions and /or midrange theory from case-based, empirical evidence (Eisenhardt, 1989). As Eisenhardt and Graebner (2007) note, the

central premise of case study strategy is to develop theory inductively from cases, therefore theory is emergent because it is situated in and developed through patterns of relationships discovered within and across cases. Case studies are rich, empirical investigations of particular instances of a phenomenon often based on different data sources (Yin, 2002). Cases can be historical, but are most likely to be contemporary descriptions of recent events (Eisenhardt and Graebner, 2007).

As Yin (2002) argues, multiple cases like laboratory experiments are discrete experiments that can replicate, contrast and extend an emergent theory. However, Eisenhardt and Graebner (2007) distinguish case studies from laboratory experiments which isolate phenomenon from their context, whereas case studies are essentially rich descriptions of real-world phenomenon studied within their context. Also, case studies build theories through recursive cycles amongst the case data, emergent theory, and then, extant literature, with the result that a well-designed case study usually generates a theory which is '*objective'* and is closely tied with the data that keeps researchers *'honest'* (Eisenhardt and Graebner, 2007:p.25). Accordingly, inductive theory building from cases is *"likely to produce theory that is accurate, interesting and testable"* (Eisenhardt and Graebner, 2007: p.26).

3.5.3 Selection of the Case and Theoretical Sampling

In case study research strategy, another important consideration a researcher has to make concerns case selection (Eisenhardt and Graebner, 2007). Different researchers view a case differently. For example, while Merriam views a case as a delimiting object, Miles and Huberman (1994, p.25) contend that a case is a *"Phenomenon ..... occurring in a bounded context"*, a view also supported by Bromley (1986, p.21) – a case study *"must be limited in scope ..... there must be conceptual boundaries and empirical limits to it".* Merriam also contends that a case (phenomenon under study) must be bounded. But the important question is how to determine if a case is bounded. Adelman et al., (1983, p.3) note that bounded phenomena should have obvious boundaries such as an individual or a single organisation.

Based on Adelman et al., (1983) common sense interpretation of a case, this research selects Facebook, the largest social network site, as a case organisation. Facebook now has one billion users (Facebook, 2012); meaning one in seven people in the world is a Facebook member. A distinguishing feature of Facebook is that the profiles of its users are often more accurate (compared with other SNSs) because people use their real-world identities to communicate with real world-friends (Dwyer, 2007a). This makes Facebook an interesting place for privacy research (Bonneau at al., 2009b). Together, these reasons motivated the researcher to choose Facebook's implementation of social advertising initiative (social Ads, in short) to holistically investigate users' perceptions of privacy leakage in social network sites. The selection of a single case, i.e. implementation of social Ads by a single organisation, i.e. Facebook, is consistent with the main philosophy of case study research as echoed by Siggelkow (2007) that even a single case provides deep insights into the phenomenon under study. Also, the author argues that the selection of Facebook's social Ads enables the researcher to study the context-dependent and temporal nature of privacy in SNSs because longitudinal data was collected between November 2007 and December 2010.

The social Ads programme was implemented by Facebook in two stages: the launch of Beacon in November 2007 and of Facebook Connect in December 2008. Consequently, longitudinal data was collected relating to these two launches between November 2007 and December 2010 (three years). The longitudinal data enable the researcher to fulfil one of the research objectives – how privacy perceptions of online social network users evolve overtime due to organisational practices. Longitudinal data was collected in two stages: during stage one data was collected relating to the launch of Beacon and during stage two, data was collected relating to the launch of Facebook Connect. Specifically, during stage one, qualitative data was collected between November 2007 (i.e. when Beacon was launched) and September 2009 (when Beacon was shut down by court order). See chapter four for the details of Beacon launch, data collection and analysis. Similarly, during stage two, qualitative data was collected between December 2008 (i.e. when Facebook Connect was launched) and December 2010 (when theoretical saturation was achieved and data failed to offer any new insights). See chapter five for the details of Connect launch, data collection and analysis.

Because the purpose of this research is not to test a theory but rather to build (possibly) a new framework, theoretical and not random or stratified sampling is appropriate (Eisenhardt and Graebner, 2007). According to Eisenhardt and Graebner (2007) theoretical sampling means that: *"cases are selected because they are particularly suitable for illuminating and extending relationships and logic among constructs. - - - - - Cases are sampled for theoretical reasons such as revelation of an unusual phenomenon, replication of findings from other cases, contrary replication, elimination of alternative explanations, and elaboration of the emergent theory" (p.27).* The author argues that understanding the complex, fluid and temporal nature of privacy, theoretical sampling provides an opportunity for the researcher to choose revelatory cases so that the issue of privacy leakage can be examined in the context of SNSs in a longitudinal case study Accordingly, the current research selected Facebook as a case organisation and implemented a two-stage longitudinal case study design to study users' perceptions of privacy leakage in SNSs by specifically focusing on the case of Facebook's implementation of social Ads. Although the social Ads programme was implemented in two stages, it has been treated as a single case study design because both launches (of Facebook Beacon and Facebook Connect) related to a single organisation (Facebook) and a single system (Social Ads).

The research design process as highlighted in Figure 3.1 is iterative and there is a continuous interplay between data collection and analysis which is consistent with the grounded theory approach proposed by Glaser and Strauss (1967). This is well supported by Van Maanen (1998) who argues that the nature of qualitative research is complex and hence should have a flexible and emergent character. Similar views are echoed by Gephart (2004, p.455) – qualitative research is designed while it is done. Therefore, qualitative research often mandates *"highly contextualised individual judgements"* (Van Maanen, 1998:p.xi). Finally and more importantly, qualitative research provides researchers the flexibility to be open to unanticipated events while allowing them to depict realities more holistically (Gephart, 2004). This research also benefits from such flexibility and utilises unanticipated events relating to the launch of Facebook's social Ads programme (the launches of Beacon and Facebook Connect) to depict social reality holistically i.e. how social network users perceive privacy leakage which often occurs as a consequence of organisational practices.

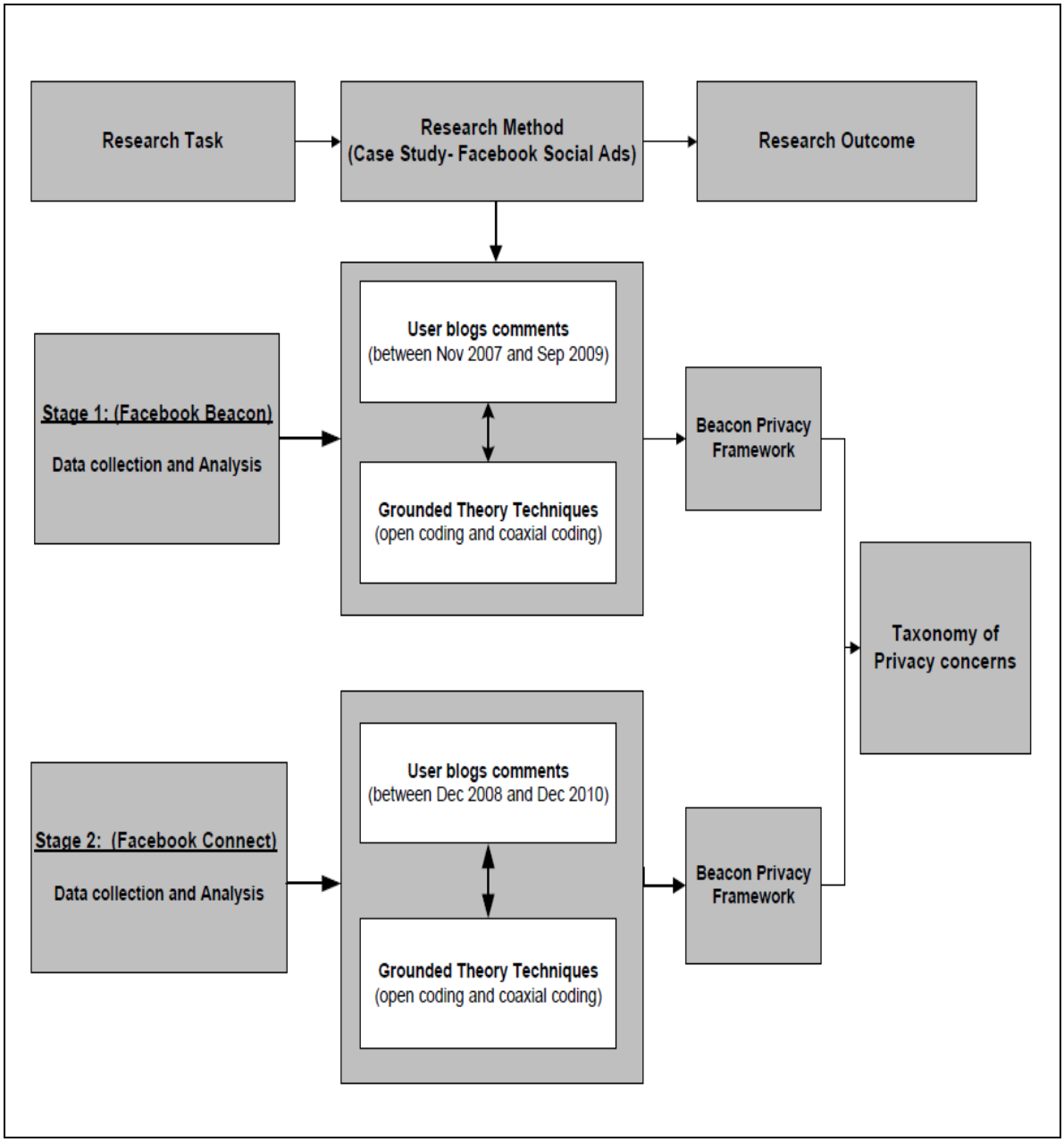


Figure 3-1: Research Design

3.5.4 The Case Organisation-Facebook

As aforementioned, Facebook is selected as a case organisation because it is: the largest social network site with over one billion active users (Facebook, 2012a); the most visited website in the world (Alexa, 2012); and because research reported that Facebook users' data was leaked (Krishnamurthy and Wills, 2008; 2010; Bonneau et al., 2009b; Debatin et al., 2009). Facebook, founded in 2004 by a Harvard student to enable students to connect and socialise within Harvard University, soon became available to other universities and colleges in the US (Cassidy, 2006). Expanding even further, Facebook included high

schools in 2005 and companies in 2006 (Zhao et al., 2008) and allowed open public access soon afterwards.

Compared with other SNSs such as MySpace, Facebook's distinctive feature is that most of its users are likely to be real people because it originated in educational institutions (which required an educational email account to join). Facebook provides a technological platform for people to create profiles (free of charge) which usually consist of their photos, personal information (e.g. name, gender, date of birth, email address, telephone number, mobile phone), personal interests (e.g. likes, music, movies, games), and political and religious views. The information such as name, email, DOB, telephone numbers, address, photos is personally identifiable information (Krishnamurthy and Wills, 2010). Additionally, Facebook users carry out many social activities of which a few relate to connecting with friends and family, create groups to support an event or cause, and create pages. According to an estimate, 13 million transactions (interactions) per second happen on Facebook (Catanese et al., 2011).

Facebook allows outside people called developers to create social applications, websites and devices and integrate into Facebook's core functions (Facebook, 2012b). Just to mention a few, these third-party social applications include features like games, music, movies, Top friends, maps, stock exchange, and YouTube videos. According to the latest Facebook statistics, 9 million apps and websites are integrated into Facebook, users install 20 million applications every day, 80% of businesses are represented on Facebook and it is now the number one publisher of Ads, accumulating ad revenue of $2.16 billion in the US (Facebook, 2012b). The unprecedented success and popularity of Facebook, together with the detail-rich personally identifiable information of one billion people, has opened new avenues and challenges for researchers, marketers, regulators, educators, media and businesses. Of particular interest to this research is the launch of Facebook's social Ads programme and the associated privacy concerns which provided an opportunity to the researcher to empirically investigate the issue of privacy leakage in social networks in real-life settings so that the context-dependent conceptualisations of privacy can be identified and analysed.

3.6 Data Collection and Analysis

3.6.1 Methodological challenges in Privacy Research

A variety of data collection methods are utilised by qualitative researchers, including interviews, focus groups, observations, and documents. However, case study researchers usually adopt interview and documentary material (Myers, 1997) as do IS researchers. Though the practice is widely recognised amongst IS researchers in general, it is not without limitations, especially when researching a sensitive topic like privacy. These limitations are assessed below, keeping in mind the methodological challenges and guidelines highlighted by various researchers in privacy research.

Garde-Perik (2009) provides methodological guidelines for privacy research which among others include: (a) privacy research aiming to investigate users' privacy perceptions should include privacy sensitive people as well; (b) the researcher's interest should be disguised as it could interfere with true representation of privacy concerned people, and researcher's bias could sensitise people about privacy during the data collection process; and (c) the research should expose people to realistic privacy risks, otherwise participants may not take privacy seriously and their attitudes will not reveal their actual behaviour. Also, such risks should be context-specific so that the actual behaviour of participants can be investigated as attitudes expressed out of context are not good predictors of behaviour (Ajzen and Fishbein, 2005; Nissenbaum, 2004). The temporal nature of privacy further poses challenges for privacy researchers as current privacy perceptions do not exist in isolation, but rather are an outcome of past privacy experiences (Altman, 1975; Palen and Dourish, 2003).

Surprisingly, the review of the literature suggests that not many privacy studies in IS have adopted these guidelines and therefore their results can be questioned (Garde-Perik, 2009). For instance, the majority of privacy studies employed interviews and surveys without regard to recruiting privacy concerned participants. Also, few studies disguised the researcher's interest in privacy during the data collection process. Again, though a few studies have been able to expose participants to realistic privacy risks by using scenarios within certain contexts, the researcher argues that such scenarios cannot fully represent

real-world settings. Finally, though the literature review reported a few longitudinal privacy studies in IS, the majority of such studies were aimed at investigating the information disclosure behaviour of users. To the best of the author's knowledge, no study exists within the context of online social networks which has exposed participants to real privacy risks, employed longitudinal case study approach or included privacy sensitive participants.

The current study, in order to fill this methodological gap in privacy research and to offer new methodological insights, is designed to adhere to the aforementioned methodological guidelines to examine the issue of privacy leakage in social networks. So that these methodological challenges should be overcome, selection of appropriate data collection method is critical.

3.6.2 Rationale for Using Blogs for Data Collection

The current research gathered commentary from user weblogs (blogs in short) to examine users' perceptions of privacy leakage in social network sites. Blogs are frequently modified web pages with dated entries in reverse chronological order (Bortree, 2005; Buckingham and Willett, 2006; Schmidt, 2007). Although there are many motivations for blogging, according to a recent survey 70% of bloggers use blogs to share their expertise and experiences with other people (State Of The Blogosphere, 2011). Hence, blogs become a useful source of data collection for researchers (Jones and Alony, 2008; Hookway, 2008), particularly in cases where people share their experiences. The inherent benefits of using blogs, the nature of the research problem and the methodological challenges informed the researcher's choice to use blogs data in the context of this thesis as discussed below.

- First, blogs provide many benefits to researchers because they offer a low-cost and instant way of collecting huge amounts of publicly available data (Hookway, 2008). Accordingly, in this thesis the researcher collected publicly available opinions of SNS users' published as a reaction to the launch of Facebook Beacon and Facebook Connect and offered via blogs in almost real-time. The blogs comments were published instantly by SNS users as a reaction to leakage of their privacy. To the best of the researcher's knowledge, this is the first study which

collected these blog-opinions to investigate the phenomenon of privacy leakage in SNSs.

- Also, blogs represent a natural form of text data which does not require any tape recording and transcription (Liamputtong and Ezzy, 2005) and hence they are potentially free from the errors which can occur during transformation of data e.g. from voice to text (Razera et al., 2010). Similarly, Jones and Alony (2008) suggest blogs provide codified data with lots of ease and convenience. Thus, SNS users' commentary published on blog sites as a reaction to privacy leakage seems a highly appropriate source of data to examine privacy leakage in SNSs.

- Gruhl et al., (2005) argue that blogs give more reliable insight into public opinion. Methodologically, Thelwall and Hasler (2007) suggest that blogs are a useful means of acquiring a relevant set of opinions or attitudes towards an event or topic. Therefore, blogs were seen as the most appropriate means of gathering users' opinions about privacy leakage in SNSs.

- Moreover, according to Jones and Alony (2008) blogs are a rich source of data, given the depth of information available. The author argues that such richness and depth of data is desirable in order to investigate the underlying causes and consequences of privacy leakage in SNSs.

- Also, blogs data collection is unbiased by the research process (Jones and Alony, 2008: pp. 439-440) as researchers can disguise their interests during data collection, which occurs passively. This is one of the key motivators for adopting blogs to collect user perceptions of privacy leakage in SNSs because it enables the researcher to automatically disguise his interest in privacy research as proposed by Garde-Perik (2009). Hine (2009) also argues that unobtrusive online data collection is promising, particularly for sensitive topics, as it will reduce the burden on the participants being researched. This may be attributed to the anonymous nature of online context that enables bloggers and online users to share their opinions without any self-consciousness (Hookway, 2008), particularly on the sensitive issue of privacy leakage.

- Again, in line with the suggestion of Garde-Perik (2009), blogs data promise the ability of gathering the opinions of privacy sensitive users who shared their experiences of privacy leakage relating to the launch of Facebook social Ads programme. It may be because during this launch, the stories of their actions performed on other websites were shared with their friends on Facebook without their knowledge – thus leaking their privacy. Also, online users who comment on privacy leakage on blogs are more likely to be privacy concerned. The author believes that other methods of data collection, such as interviews, may not be adopted because they do not provide any mechanism to gather data specifically from privacy sensitive SNS users.

- Also, research suggests that users' privacy behaviour can better be predicted if they are exposed to real privacy risks (Garde-Perik, 2009) as SNS users were found disclosing their real personal information on profiles voluntarily whilst at the same time their attitudes showed they were concerned about their privacy (Acquisti and Gross,2006). This suggests that peoples' general attitudes towards privacy cannot predict actual behaviour unless people are exposed to real privacy risks (Garde-Perik , 2009). Therefore, the use of blogs data published as a reaction to the launch of Facebook's social Ads which leaked SNS users' privacy seems appropriate as it enables the researcher to collect user's perceptions of privacy when they were actually exposed to real privacy risks in a real-world setting.

- Another reason for choosing blogs data in current research is the fact that blogs were the only source of engagement between SNS service providers and SNS users. Facebook's CEO blogged his public apology and all changes made by the company in the social Ads tools were communicated via blogs as well. Accordingly, SNS users reacted almost in real-time via blogs commentary.

- Furthermore, blogs provide the researcher with an opportunity in the context of this thesis to access geographically distributed or socially removed populations in the current research (Hessler et al., 2003; Mann and Stewart, 2000), which was otherwise challenging in other methods of data collection.

- Last but not least, one of the objectives in this research is to explore how privacy perceptions of SNS users evolve overtime – which mandates collection of longitudinal data. The archived nature of blogs makes them a tempting resource for the collection of longitudinal data as they can be used to examine change in the social process overtime (Hookway, 2008). Therefore, using blogs in the current research is fitting with the objectives of this research.

3.6.3 Ethics of Using Blogs in Current Research

Like any data collection and analysis method, a researcher utilising blogs data in a research has to adhere to ethical guidelines. Importantly, due consideration was given to ethical issues related to collection and analysis of blogs postings in the current research as discussed below. The use of online data collection methods, such as blogs, has focused the attention of researchers on how to address the ethical issues relating to their use in any research. An important ethical issue is whether blogs data is public or private and if a participant's consent is required or not. Three standpoints are reported in the literature around this issue (Hookway, 2008). Researchers such as Sudweeks and Rafaeli (1995) and Walther (2002) claim that archived material (e.g. blogs) is publicly available and no consent is required. At the other extreme are the researchers like Elgesem (2002), King (1996) and Scharf (1999), who believe that online data is posted with the expectation of privacy in mind, so consent is required. Finally, researchers like Waskul and Douglas (1996: 131) argue that online data is both private and public.

Hookway (2008: p.105) whilst clearing the fog surrounding the private/public status of blogs data, builds on the 'fair game – public domain' argument, suggesting firmly that blogs data is public and that therefore the consent of the participant should not be required. Arguing further, Hookway maintains that as blogs can be defined as private by users and accessible only to their friends, the blogs which are not made private are therefore public. Moreover, Jones and Alony (2008) also view blogs as public and claim that no consent of the participant is required. Apparently, this view is dominant in the research community as more and more studies are employing online data collection methods, claiming it to be publicly available data. For instance, Money et al., (2011) collected consumer opinions on a popular e-commerce website and regarded these

opinions as public data. Similarly, Bulgurcu et al., (2010), when considering the public nature of online data, collected online comments posted by Facebook users in their research on information privacy in SNSs.

Therefore, the researcher argues for the use of blogs data as a public source of information without the need to obtain consent from the bloggers and commentators in this research. Indeed, this research adopts the 'fair game-public domain' position as advocated by Hookway (2008) in order to collect and analyse blogs data. Accordingly, in the interests of fairness, the privacy of bloggers and commentators was ensured since their personal information, such as name, was disguised by replacing it with the sequence numbers to refer to the bloggers during analysis. It is worthwhile mentioning here that very few blog commentators used their names, most of them employing pseudonyms. Even so, online pseudonyms were also concealed and replaced with the sequence numbers. Furthermore, quotes that contained any information that might identify the blogger were also disguised, as suggested by Hookway (2008:106).

Another ethical issue relating to blogs, as highlighted by Jones and Alony (2008), relates to the reliability of collected data, as many fake blogs are published on the internet. Accordingly, the researcher's data collection protocols (see details in the following section) were carefully selected, which ensured the collection of reliable blogs. For instance, a social recommendation system (Digg) was used where people rated blogs after reading the content and blogs with most recommendations were collected. Presumably, the collected blogs were reliable as they were based on people's recommendations. Also, popular news media and technology websites were used to search blogs which also ensured reliability of the collected blogs since it is highly unlikely that someone can publish a fake blog on the BBC or New York Times websites. Other checks applied are discussed in the following section.

3.6.4 Data Collection Protocols

Theoretical sampling was chosen as the most appropriate strategy to collect user blog commentary. However, the data collection and analysis overlapped, which is one of the important features of case study research as a theory building strategy (Eisenhardt, 1989). The joint data collection and analysis approach adopted in this thesis is in-line with the

grounded theory method proposed by Glaser and Strauss (1967). Also, to collect the most reliable and credible blog postings, Digg (Digg.com), social news aggregator and recommender system was used and the blogs with high user ratings were collected. Indeed, blogs with higher ratings appeared at the top of the page. To further optimise the data collection and to gather an uninterrupted up-to-date user commentary on blogs, RSS feeds commonly known as 'Rich Site Summary' or 'Real Simple Syndication (Duffy et al., 2006) were used. RSS is a standard feature which enables users to subscribe to the contents of a website/blog using newsreaders or aggregators (Duffy et al., 2006). Specifically, Google Reader (http://www.reader.google.com), a popular news aggregator tool, was used to automatically subscribe to up-to-date user opinions on the selected blogs gathered using above keyword phrases on the aforementioned blog sites. The use of an RSS feed provided automatic access to collect up-to-date user comments from the subscribed blogs sites.

In order to collect diverse opinions, maximum variation sampling strategy was followed (Kuzel, 1992). This was achieved through selecting different types of blogs published on popular news media sites, general technology sites, social media blogs sites and personal blogs. Among the popular news media sites used were: BITS at The NY Times and dot.life at BBC News) and general technology sites such as Techcrunch and PCWorld, and social blogs sites such as SociableBlog. Within the qualitative interpretive approach adopted in this thesis, qualitative evidence was collected to empirically investigate the phenomenon of privacy in social networks. Gephart (2004) suggests that qualitative data can be collected by using one or more research methods, including case studies, observation, interviews, grounded theory, and textual analysis. However, Benbasat et al. (1987) argue case study is an appropriate method to research a relatively new phenomenon and where organisational issues of IS are studied. As SNSs have emerged recently and the current research aims to investigate users' perceptions of information privacy in SNSs which are a consequence of organisational practices, the case study research method is adopted in this thesis. Data analysis is inspired by grounded theory method (GTM) because it offers systematic analysis procedures (Strauss and Corbin, 1998) which enable researchers to identify patterns in data and thus develop empirically valid theories through analysing these patterns (Glaser and Strauss, 1967; Martin and Turner, 1986).

Gephart (2004) argues that data collection and analysis are two important activities of qualitative research and researchers can use either qualitative or quantitative methods in both the data collection and data analysis stages. However, quantitative analysis on qualitative data requires data to be quantified and hence such a research is quantitative and not qualitative (Gephart, 2004). Therefore, researchers should choose research methods carefully in both data collection and data analysis steps of a research. Also, they should make sure that the research methods employed in different stages of research are consistent with the aim of research study, as well as with the underlying epistemological assumptions (Gephart, 2004).

This research is informed by the use of GTM for data analysis purposes. The choice is predominantly influenced by the established data analysis techniques of GTM (Hughes and Jones, 2003; Matavire and Brown, 2008; Orlikowski, 1993). Urquhart et al., (2010) argue that GTM has proved to be an extremely useful method in information systems to build context-based, process-oriented descriptions and explanations of information systems phenomena. Consistent with the goal of this research, GTM assists in developing the theoretical framework of the general properties of a phenomenon as well as providing deep insights from the empirical data (Martin and Turner, 1986; p.141). Also Orlikowski (1993) highlights that grounded theory is particularly useful for new areas of research which lack established theories. Undeniably, identifying and analysing the privacy perceptions of online social network users which are mainly a result of organisational practices is a new phenomenon. Therefore, the established open and axial coding techniques are adopted in this thesis. This overlap between data collection and analysis process is an important characteristic of grounded theory method, as it enables the researcher to remain close to data (see figure 3.2) Creswell (2002). Also, because this research used inductive analysis approach predefined codes were not adopted; rather, the codes were generated from user blogs commentary. Thus, the developed themes are tightly linked to data (Patton, 1990). Furthermore, constant comparison principle of grounded theory approach guarded against any bias (Corbin and Strauss, 1990) during data analysis process.

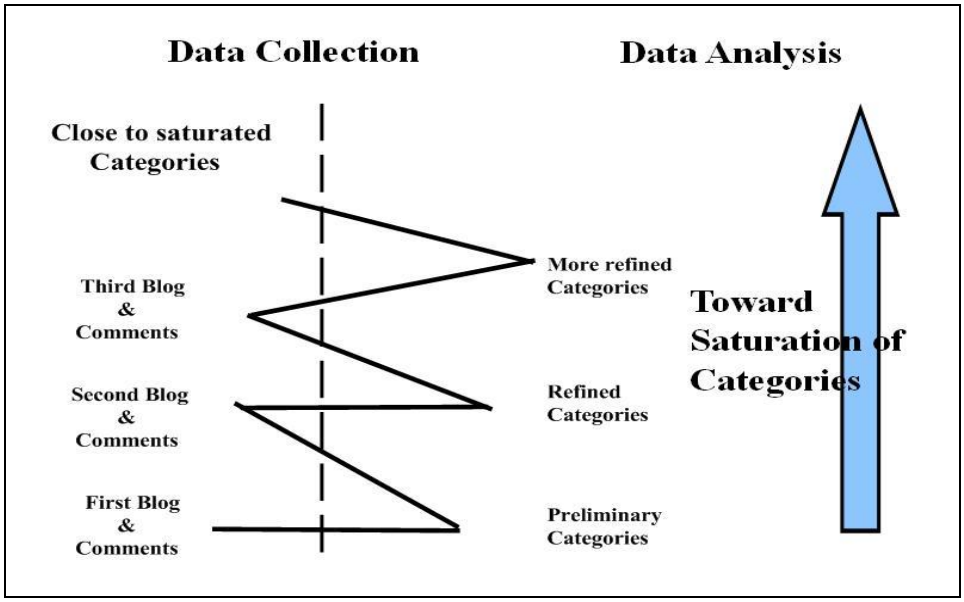


Figure 3.2: Zigzag Data Collection and Analysis Approach
(Adapted : Creswell (2002)

An important decision any researcher has to make is whether to use preconceived theory during the data analysis process. Various researchers have argued for and against this. However, the researcher agrees with the point of view of Ng and Hase (2008) who argued that it is neither possible nor desirable to enter into research with a totally clean sheet. Strauss and Corbin (1967) also argue the need for theoretical sensitivity in the sense that it improves the analytical capacity of researchers to better understand patterns and categories in the data. Therefore, the researcher conducted a preliminary literature review and identified the issue of privacy leakage in social network sites. Also, the literature review identified the general construct 'privacy concern' to measure privacy related issues in IS research, which was adapted within the context of this research and used as an analytical guide in the data analysis process. However, the researcher avoided using preconceived theories or frameworks because privacy leakage is a recent phenomenon and little is known yet, particularly about how SNS users perceive such leakage of their privacy.

3.6.5 Use of Computer Assisted Qualitative Data Analysis Software (CAQDAS)

The iterative nature of grounded theory data analysis procedures requires researchers to move between data collection and analysis, writing memos, coding, and developing models. Therefore, non-linear computer-assisted qualitative data analysis software (CAQDAS) was used to support these iterative activities (Bringer et al., 2006).

Specifically, NVIVO (version 8 and 9), a dedicated computer assisted qualitative data analysis software was used. Many qualitative research studies based on grounded theory data analysis techniques use CAQDAS. For example, studies of Bringer et al., (2006) and Hutchison et al., (2009) used NVIVO. The use of NVIVO is particularly appropriate for grounded theory analysis techniques as NVIVO support the iterative collection and analysis of data (Bringer et al., 2006; Hutchison et al., 2009). Bazeley (2007) argues that computer features such as storing, sorting, searching, and linking can be effectively utilised to enhance the data analysis process.

The researcher found some features particularly useful during the collection and analysis of data. Because user blogs commentary was gathered iteratively using RSS feed as mentioned above, NVIVO facilitated adding more data to existing already coded text. For example, the NVIVO feature 'Highlight' allowed the researcher to highlight coded text in different colours so as to distinguish which text has been coded and which has not. To apply open coding technique with a microanalysis approach was also well supported in NVIVO. For example, the free nodes feature enabled the researcher to open code data at word, sentence and in some cases at paragraph levels. Another feature which the author found really helpful during open coding was the 'In Vivo' coding facility in NVIVO. The 'In Vivo' coding allowed the researcher to code data by using the same words or terms used by the bloggers in the text so that the codes are tied to the actual data – the dominant feature of grounded theory analysis.

NVIVO provides rich features to manage codes (called nodes in NVIVO). For example, nodes can be deleted, merged, and copied easily. This feature proved invaluable during the second iteration in the open coding process as duplicate nodes were merged together easily. During this step, the 'Highlight' node feature of NVIVO was also helpful in differentiating the coded and un-coded text, ensuring the researcher did not omit any text to code. The axial coding procedure of grounded theory is also well supported in NVIVO. Tree nodes facility in NVIVO version 8 specifically facilitated the researcher to group similar nodes together and form a hierarchy of nodes. Also, the NVIVO feature 'Coding Strips' helped the researcher to group related nodes during the axial coding process as they show the nodes coded to a particular text (Hutchison et al., 2009). Overall, the

researcher maintains that the use of NVIVO in this thesis provided invaluable support to manage the overall process of data collection and analysis.

3.6.6 Data Analysis Process

The data analysis process was divided into following steps:

1. ***First level of Analysis:*** As the data collection and analysis overlapped, the blog postings are continuously imported in NVIVO so as to facilitate the iterative data analysis process. . Indeed the use of NVIVO facilitated iterative approach to data analysis (Bringer et al., 2006; Hutchison et al., 2009). The first level of analysis concluded in two iterations. In the first iteration the focus was on making sense of the data through repeated reading whilst keeping in mind at all times the research question i.e. to find privacy concerns related to privacy leakage. Open coding technique of grounded theory was embraced to analyse data such that *"data are broken down into discrete parts, closely examined, and compared for similarities and differences"* (Strauss and Corbin, 1998: p.102). Open coding analysis technique applied during the first level of analysis enabled the researcher to remain open in order to identify all potential privacy concerns and issues in the data. The coding was done under a microscopic examination of the data (mostly sentence by sentence but in some cases even word by word or paragraphs) as proposed by Strauss and Corbin (1990).

   Inductive coding analysis approach assisted the researcher in avoiding any predetermined or preconceived ideas, but rather codes were used in bloggers/commencers' own terms. Indeed the 'in-vivo' coding feature of Nvivo was invaluable towards meeting this end. The codes represent a feature of the data that is of analytic interest to the researcher and is a basic element of the raw data (Boyatzis, 1998). The analytical interest of the researcher, in this particular instance, was to identify privacy concerns of SNS users associated with the leakage of their privacy as a consequence of organisational practices. In the second iteration the codes were checked to confirm that they corresponded to the coded extracts and were not redundant. Accordingly, similar codes were merged. The second iteration also gave the researcher an

opportunity to read entire data sets again to see there was no omission of important codes.

2. ***Second Level of Analysis:*** Using the results from the first level of analysis wherein the emerged codes were not shown in any relationship or hierarchy, the second level analysis adopted axial coding procedure to group the related codes to generate more generic or abstract themes. During this stage of analysis, all the first level codes were iteratively grouped into distinct theoretical themes (Eisenhardt, 1989). All the codes were reviewed to find out any relationship between them. Accordingly, the similar codes were grouped into more abstract themes. At this stage, all the themes represented separate data and could not be further refined or eliminated.

3. ***Theoretical Framework:*** Both the first and second level of analysis identified themes and sub-themes relating to the launch of Facebook Beacon (conceptualised as Beacon privacy framework- see chapter four) and Facebook Connect (called Connect privacy framework- see chapter five). The next step of analysis proceeded to consolidate both the privacy framework to devise a comprehensive framework called a taxonomy of privacy leakage concerns of online social network users.

4. ***Evaluation via Literature Review:*** In the final stage of analysis, the taxonomy of privacy concerns is compared with already published and most cited taxonomy of privacy (see chapter six for details) so as to highlight similarities and differences in order to identify potential contributions of this study.

CHAPTER 4: LONGITUDINAL CASE STUDY FINDINGS - THE BEACON CASE

4.1 Chapter Overview

This chapter presents research findings of stage one of the longitudinal case study – the failed launch of Facebook's social Ads tool Beacon, which had to be withdrawn after its launch because of concerns related to privacy leakage of SNS users. Also, the analysis of literature in chapter two highlighted that social network users' data is a hugely tempting resource for organisations which are adopting innovative ways and practices to utilise SNS users' data for commercial gains. Recently, SNSs have started integrating with third-party websites, including advertisers, aggregators and Ads networks, raising concerns amongst SNS users over privacy leakage. This study aims to understand the nature and forms of users' privacy concerns related to such privacy leakage in SNSs as well as establishing how these concerns change overtime.

Thus, as highlighted in chapter three, a longitudinal case study spanning three years and comprising two stages – the launch of Facebook Beacon (stage one) and that of Facebook Connect (stage two) – was designed. This chapter presents findings and analysis of stage one only. Its structure is as follows. Firstly, the Beacon launch and the subsequent user backlash are discussed to contextualise stage one of this research. Next, details of collection and preparation of data are discussed. Empirical findings are presented thereafter. This is followed by analysis of the results. The chapter closes with the conclusions of stage one of the longitudinal study.

4.2 Beacon Launch and User Backlash

In November 2007, Facebook, the hugely popular SNS, launched an innovative marketing tool called Beacon. Beacon was intended to provide an innovative approach to personalized marketing by means of *"socially distributing information"* (Facebook Press, 2007). According to Facebook Press (2007), 44 leading businesses including Blockbuster, eBay, Fandango, Travelocity and Yelp participated in the Beacon launch. The central premise was to leverage social networks by enabling third-party commercial companies to

allow users to share various actions amongst their friends via automatic news feed. Such actions could involve posting an item for sale, purchasing an item such as a cinema ticket or holiday and relaying scores achieved in an online game (Jamal and Cole, 2009). Hence, whenever a user performed an action on a third-party site (e.g. buying an airline ticket on Travelocity), a Beacon alert informed the user that this 'action' had been sent to all their Facebook friends (see Figure 4.1). Users could cancel these automatic transfers by opting out, but they would have to do this for each and every action taken on a participating third-party website.

Figure 4.1: Example of an Early Beacon Alert

Figure 4.1 illustrates an early Beacon notification, which indeed was evasive since no prior notice had been given to the user about what Beacon was and what it was doing. Also, there was no opportunity to accept or reject the offer of sending the 'story' to the user's profile unless s/he clicked the button 'See More', which revealed details of the 'story' (Figure 4.2).

Figure 4-2 Beacon interface when users clicked 'See More'

Even then, the Beacon interface still failed to provide an explicit opportunity to accept or reject the offer of sending a story to a user's Facebook profile; that was introduced in the final Beacon interface (Figure 4.3). Although the final version allowed the user to universally accept publication of all stories, it failed to provide a similarly universal opt-out.

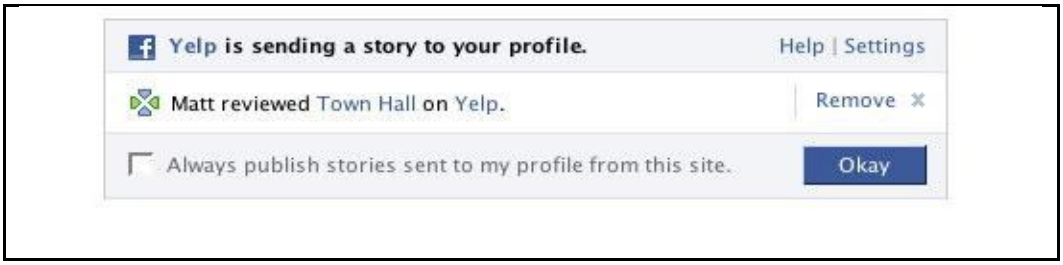


Figure 4.3: Final Version of Beacon Interface Allowing Universal Opt-in

The overall information flow is well depicted in Figure 4.4.

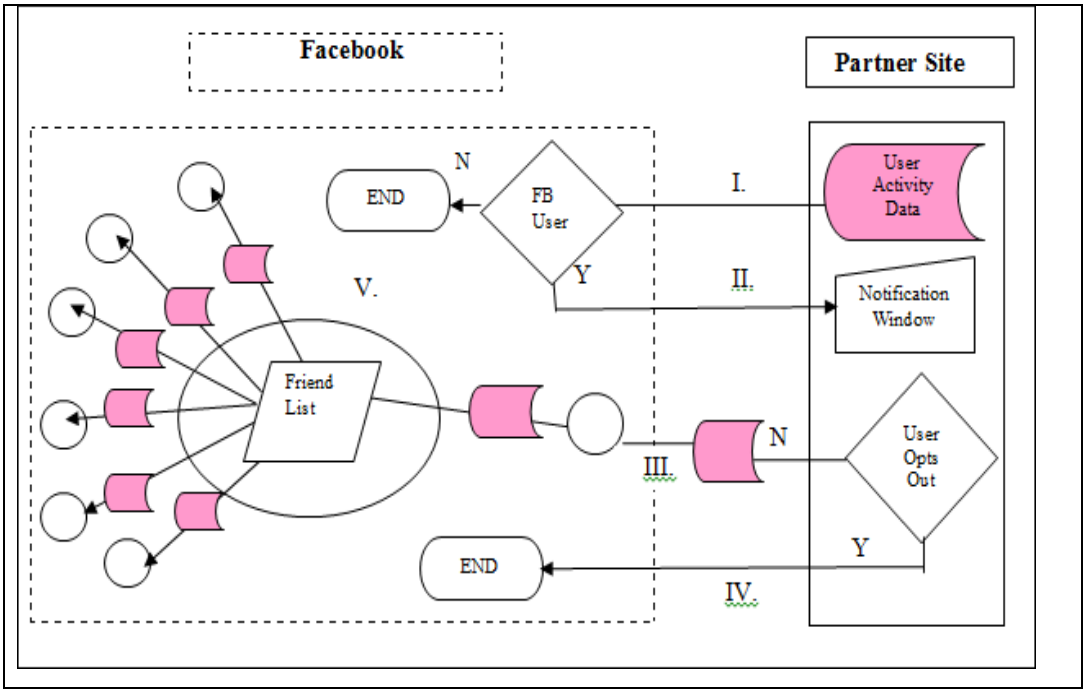


Figure 4.4: Beacon Information Flow (Compiled after Martin (2010, p.2))

From a business perspective, automatic disclosure of purchase actions by Facebook users offered tremendous opportunities to commercial organisations. As Gary Briggs, the senior vice president and chief marketing officer of eBay (North America), remarked:

> *"Beacon offers an interesting new way for us to deliver on our goal of bringing more bidders and buyers to our sellers' listings. In a marketplace where trust and reputation are crucial to success, giving sellers the ability to easily alert their network of friends – the people who already know and trust them – to an item for sale has the potential to be a powerful tool " (Facebook Press ,2007).*

However, contrary to the expectations of third-party businesses, and the owners of Facebook, attempts to leverage personal activity for commercial gain provoked outrage amongst SNS users, whose negative reactions were loudly voiced in blogs showing their concern over privacy. User blog-commentary became increasingly hostile to the perceived encroachment on personal privacy by commercial organisations. Such was the media coverage of this issue that partner businesses such as Coca-Cola, Overstock and Travelocity, concerned about their credibility and reputation, withdrew from Beacon within weeks of its launch (Schonfeld, 2007). Amongst many such examples, Kaufman (2007) reports the case of a user who wanted to give a surprise gift to his wife at Christmas, but a message disseminated without his knowledge in his Facebook newsfeed stated that: *"- - - bought 14k White Gold 1/5 ct Diamond Eternity Flower Ring from Overstock.com — last week on the social networking Web site Facebook". Consequently, the message was visible to his 500 classmates and 220 friends, including his wife who said "I was really disappointed because for me the whole fun of Christmas is surprise, I never want to know what I am getting."*

What this and many other such examples demonstrate is that SNS users were concerned about privacy of their data being automatically shared between online businesses and SNS for marketing purposes. Beacon was also criticised by domain experts. For instance, a professor of communication in the US thought such marketing practices too intrusive, as they collected unprecedented user data and built their profiles (Havenstein, 2007). Similarly, a professor of privacy law noted that Facebook seems to have had *"antiquated"* view of users' privacy because a year before [the Beacon launch] Facebook users had protested massively against a newsfeed feature which shared some profile information with others; the Beacon initiative strongly suggested that Facebook still didn't understand the privacy concerns of its users (Havenstein, 2007). In response, four weeks after the launch and amidst a storm of user protests against their lack of control over their personal information, Facebook replaced the opt-out system with an opt-in system. This was an attempt to reclaim both user trust in the Facebook platform and commercial credibility amongst the participating third-party organisations. As Facebook CEO, Zuckerberg (2007) remarked *"We've made a lot of mistakes building this feature, but we've made even more with how we've handled them. We simply did a bad job with this release"*. The problem, though, was that Beacon continued to track users who opted-out, or were not even logged into Facebook (Berteau, 2007).

This destroyed user trust completely (Jamal and Cole, 2009) and in December 2007, within one month of the launch, Facebook had to provide privacy control to users allowing them to turn off Beacon completely (Zuckerberg 2007), thus rendering it ineffective. Consequently, Beacon was shut down in September 2009 as part of the settlement of a court order against Facebook of $9.5 million, money which would be used to establish a foundation to promote awareness of online privacy, security and safety (Perez, 2009). However, Facebook's response did not seem to appease its users, who kept on joining an online petition launched by a civic action organisation MoveOn (www.MoveOn.org) as a reaction to Beacon. Soon, the number of people signing this petition had risen to 80,000. The petition stated: *"Facebook must respect my privacy. They should not tell my friends what I buy on other sites—or let companies use my name to endorse their products—without my explicit permission"*[4].

The reaction culminated in the former founder of Microsoft, Bill Gates, publicly withdrawing his support from Facebook, stating a concern with the privacy controls provided to users (Jamal and Cole, 2009). As a result, what should have been a successful innovation was badly damaged and ultimately withdrawn because the nature and form of privacy concerns in SNSs was poorly understood. Moreover, mere protests and Zuckerberg's apology proved inadequate and they subsequently filed a $9.5m law suit against Facebook and its collaborating third-party businesses for their failure to provide notice and privacy controls in the launch of Beacon (Elden, 2010). That lawsuit represents a concrete, commercial consequence of exploiting SNS users' personal information for commercial purposes – as distinct from general, online environments such as e-commerce and online shopping, offline shopping and direct marketing, and internet use. Consideration needs to be given to the scope of the use of personal information in social networks in order to avoid user backlash and ultimately undermine the commercial aspect of SNS data. The interesting question, therefore, is the extent to which the unique character of online social network data is distinct from commercial data-set. Understanding the nature and form of privacy concerns, through a qualitative examination

[4] The online petition can be found at: (http://civ.moveon.org/facebookprivacy/?rc=fb_privacy). Accessed 20 April 2008.

of user response to Beacon, will help explore the possible boundaries and different avenues of use that organisations can contemplate when seeking to use online social data for commercial purposes.

4.3 Data Collection and Analysis

4.3.1 Data collection and preparation process

As Thelwall and Hasler (2007) highlight the need for *"an appropriate blog search to yield a set of relevant posting[s]",* social recommender system Digg (Digg.com) was used in order to collect relevant and reliable blog postings. People vote for or rate a particular blog on Digg and such ratings improve our confidence in that particular blog. The highly-rated blogs appear on the top of the page, enabling the researcher to collect the most relevant and thus improve the overall quality of postings collected from blogs. For example, some sample user-ratings of postings collected were 1237, 1094, 980, 668, 661, and 274 (see figure 4.5). Indeed, such high ratings proved useful in kick-starting the search for blogs in the current thesis by overcoming the issue of information overload. Moreover, blogs collected based on people recommendation provided hyperlinks to other relevant blogs (just like snowball sampling technique). This also proved a useful feature for the gathering of relevant blog postings. Moreover, to gather an uninterrupted up-to-date user commentary on blogs, the RSS feed commonly known as 'Rich Site Summary' or 'Real Simple Syndication (Duffy et al., 2006) was used. Specifically, Google Reader (http://www.reader.google.com), a popular news aggregator tool, was used to automatically subscribe to up-to-date user opinions on the blogs selected from Digg ratings. To direct the search process, keywords/phrases and sources of blogs were carefully selected. Keywords were defined to search for and collect the most relevant blogs, specifically: *"Facebook"*; *"Beacon"* and *"Privacy"* and their combinations resulted in four search strings: *"Facebook Privacy"; "Facebook Beacon"*; *"Beacon Privacy"*, *"Facebook Beacon Privacy"*, which were finally used to search relevant blogs. Noticeably, all three search strings have either the word "Facebook" or "Beacon". This was done to restrict the result only to blog commentary relating to Facebook Beacon.

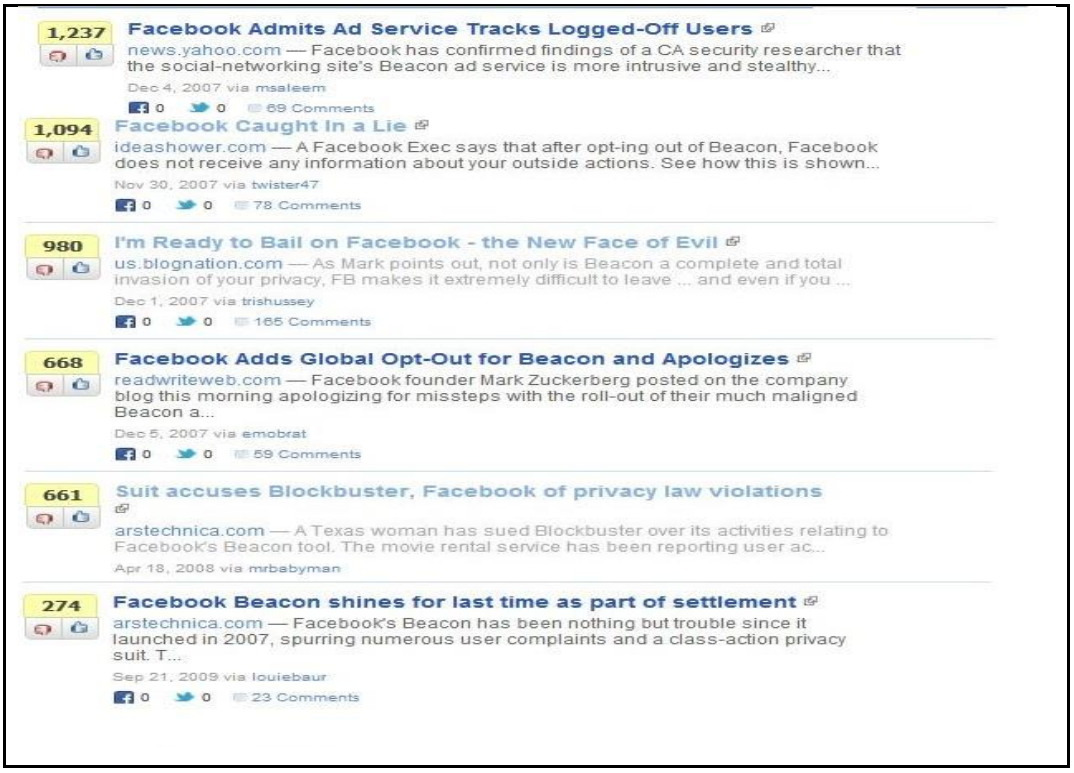


Figure 4.5: Digg rankings of Beacon-related blogs

Furthermore, to optimise the search to blogs only, blog search engines and media sites were used. However, due care was exercised in the selection of these blog-related search engines and media sites so that a wide range of blogs on the topic would be gathered. The blogs were gathered using dedicated blog search engines, namely: Google Blog Search (http://blogsearch.google.com/), Technorati (http://technorati.com/), and Bloglines (http://www.bloglines.com). In order to get wider coverage, blogs were also searched and collected from popular news media sites and technology debating sites, namely: New York Times (BITS) (http://bits.blogs.nytimes.com/), BBC News (dot.life) (http://www.bbc.co.uk/blogs/technology/) , Techcrunch (http://www.techcrunch.com/) , PC World (http://blogs.pcworld.com/staffblog/) , and Sociable Blog. To further optimise the search, LiveJournal (http://www.livejournal.com) an online weblog service was also used. Together, 45 blogs containing 1190 user comments from 844 unique users were collected. Table 4.1 gives a summary of the blog commentary collected during the launch of Beacon between 6th November 2007 (when Beacon was launched) and September 2009 (when Beacon was finally shut down by Facebook). However, the majority of blog postings were published soon after the Beacon launch and had dwindled within 3-4 months of it, reviving a little in September 2009 (when the Beacon was shut down). Theoretical sampling of blog commentary guided the researcher as to when to stop

collecting data – the point at which the codes, categories and themes became saturated and no new findings were emerging.

Type of blogs	Number of blogs collected	Total comments	% of Comments
News Media blogs	14	334	28.0
Technology blogs	9	248	21.0
Social media blogs	10	277	23.0
Personal blogs	12	331	28.0
Total	**45**	**1190**	

Table 4.1: Summary of blog comments

4.3.2 Blogs Posting Time Series Analysis

A time series analysis was conducted to show the pattern of blog comments in reaction to the Beacon launch and subsequent organisational practices related to it. Specifically, a time series of all blog comments was conducted which highlighted those posted on the first, second, third or later days during and after a blog's publication date. See table 1 in appendix A, which shows details of comments posted on each blog. The analysis shows that a large number of comments (72%) were posted on the first day of a blog's publication, followed by 17%, 6% and 5% posted on the second, third and later days respectively. This trend is depicted as a time series graph (see figure 4.6). Interestingly, blog postings followed a pattern indicating that most comments were triggered by an event or situation. For instance, the first peak (representing comments posted in November 2007) coincides with the launch of Beacon, the second (representing comments posted in December 2007) may have been influenced by Facebook's decision to change Beacon to an opt-in system instead of an opt-out system, while the third, in September 2009, was a reaction to Beacon's official shut-down by court order. In the intervals between these triggers, blog posting remains stable and dormant most of the time.

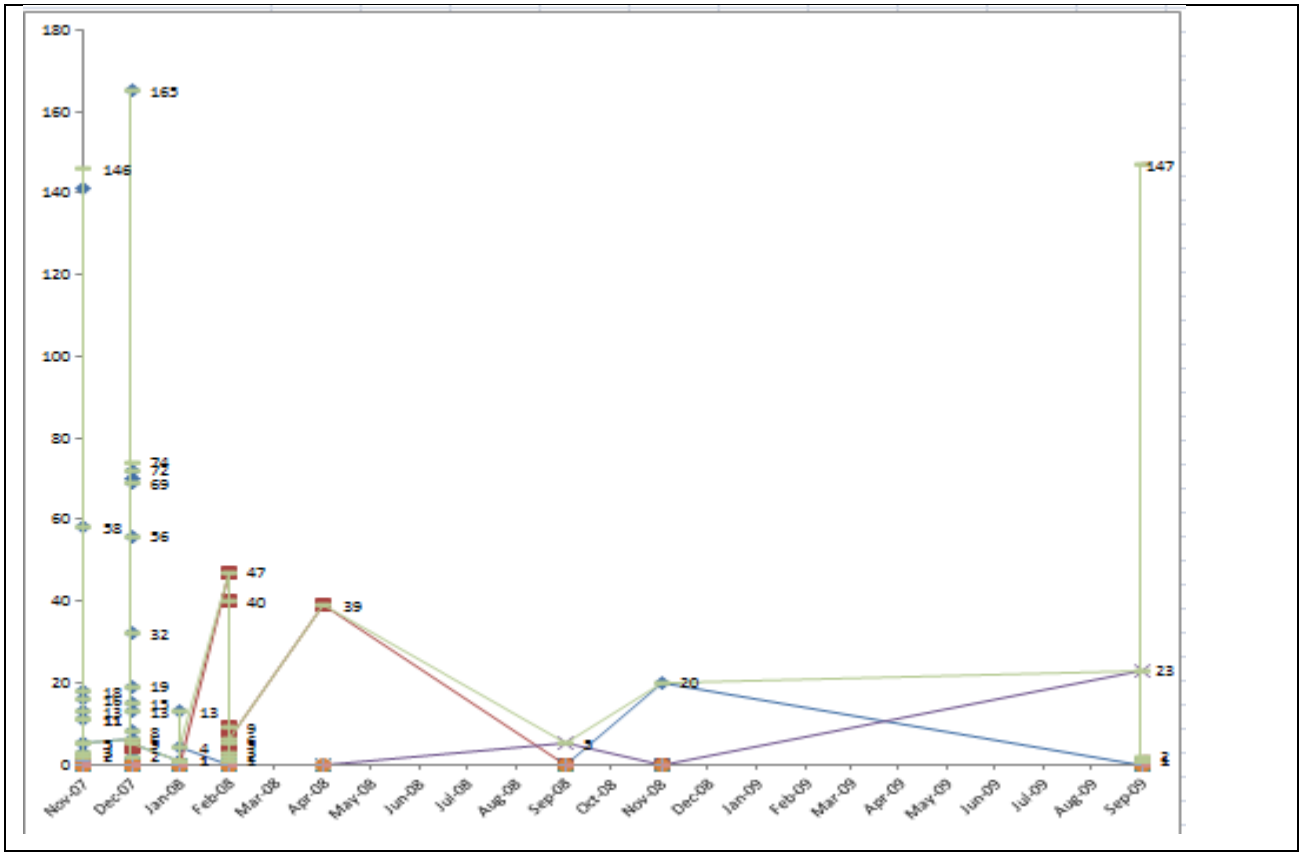


Figure 4.6: Time Series of the Incidence of Blog Comments

4.3.3 Data Analysis Process

As detailed in chapter three, this thesis adopted grounded theory procedures to analyse user blog commentary collected in two stages during the three-year period of Facebook's social Ads programme – Beacon (between November 2007 and September 2009) and Facebook Connect (between December 2008 and December 2010) respectively. The data analysis process is split into two phases. The first level of analysis applied open coding data analysis techniques to conduct microanalysis of user blog commentary. The second level of analysis adopted axial coding technique to relate and group categories identified in the first level of analysis. The overall data analysis process was discussed in chapter three. This chapter presents the first stage of the data analysis. as briefly discussed below. The rich-in-detail qualitative blog commentary comprised of 314 (A4 size) pages of text was analysed using NVIVO. As discussed in chapter three, the established grounded theory data analysis techniques such as open coding and axial coding (Strauss and Corbin, 1998) were used to guide the data analysis process. Open coding was done at a micro level in order to ensure that the entire text is checked for privacy concerns. NVIVO uses the term nodes to refer to codes, therefore both the terms nodes and codes are used in this

thesis. The 'In Vivo' coding feature of NVIVO was used and names of nodes were assigned in the same terms as commentators.

Open coding concluded in two iterations. During the first iteration, an iterative approach to coding was used through active reading and re-reading of text which helped ensure the quality of coding through feedback and refinements. Further, to improve the reliability of the coding process, 'In Vivo' nodes were created because they ensured that the nodes remain close to actual data. For example, the nodes 'Cross-pollination of information' and 'online archaeology' were created 'In Vivo'. The first iteration of open coding resulted in creation of 58 privacy themes or concerns (See Appendix A). The second iteration was done to review existing nodes as well as create new ones. Nodes were reviewed to ensure that they are not redundant and hence redundant nodes were merged. Similarly, nodes with different names but on the same semantic lines were also merged. The coding strips feature was also utilised to merge nodes which were coded at the same text. For example, the node 'Interception' was merged into 'Intrusion' node as most of the coded text of 'Interception' node was coded at 'Intrusion' node. Also, some new nodes were created as the entire data set was actively read again during the review process and the candidate text not coded in the first iteration was coded. The open coding process concluded with 39 unique privacy concerns.

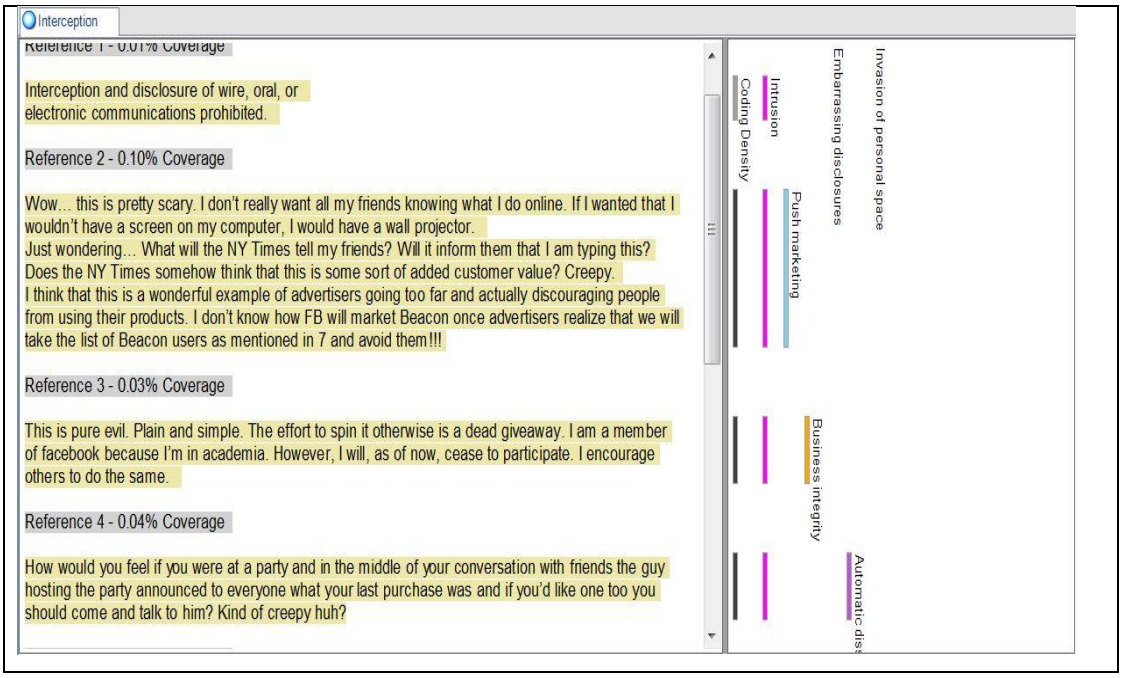


Figure 4.7: Coding Strips Interception Node

In order to move towards a richer explanatory framework (Glaser, 1978; Glaser and Strauss, 1967) of privacy concerns, the nodes were grouped in a hierarchical manner using axial coding procedures. The axial coding is an analytical process which aims to reassemble data which was divided into nodes during open coding process (Hutchison et al., 2009). Indeed, tree nodes feature of NVIVO version 8 was invaluable during axial coding process because it allowed the researcher to create major themes (called major privacy concerns) Also, the coding strips feature of NVIVO provided much help to link related nodes as they facilitate the process of comparing different nodes (Bringer et al., 2008). Accordingly, the nodes coded at similar text were grouped to form major categories (called major privacy concerns in the context of this thesis). The axial coding process resulted in six major categories which were composed of 39 sub-categories also called sub-privacy concerns (see Appendix A).

4.4 Empirical Findings

Both the first and second level of analysis identified themes and sub-themes relating to the launch of Facebook Beacon, which is conceptualised as the Beacon privacy framework and shows 39 distinct user concerns related to privacy leakage in SNSs. These are arranged in two levels of granularity, moving from the general to the specific (see figure 4.8). From within the general concerns, 6 broader categories, labelled here as major privacy concerns, emerged. These relate to user control, business integrity, transparency, information broadcast, data protection breaches and information leak. The specific detailed concerns called sub-concerns as outlined in Figure 4.8 provide greater clarity regarding the nature and form of the core concerns.

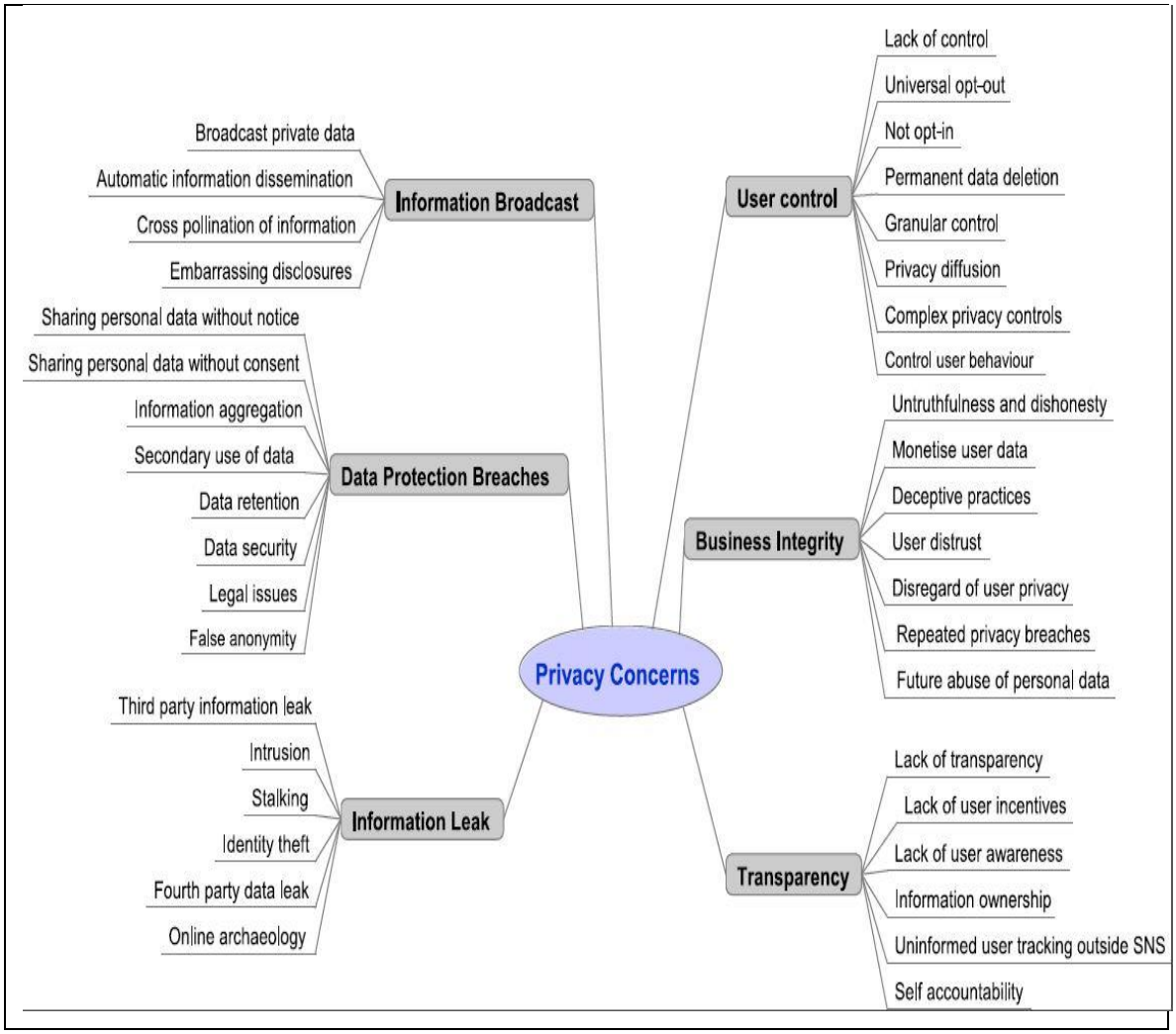


Figure 4.8: Beacon Privacy Framework

In order to rank these elements in order of severity as perceived by users, a frequency count of the major privacy concerns was conducted (see Figure 4.9). This provided insight into those that were considered mere irritants as opposed to those that represented actual boundary crossings and resulted in privacy leakage of SNS users. For example, the privacy concerns 'User Control' and 'Business Integrity' are viewed as the most severe breaches of privacy by users, representing almost half of all responses (26% and 23% respectively). Mild breaches of privacy are represented by 'Transparency' (17%), and 'Information Broadcast' (14%). In contrast, 'Data Protection Breaches' and 'Information Leak' together received 20% of total responses and were classified irritants.

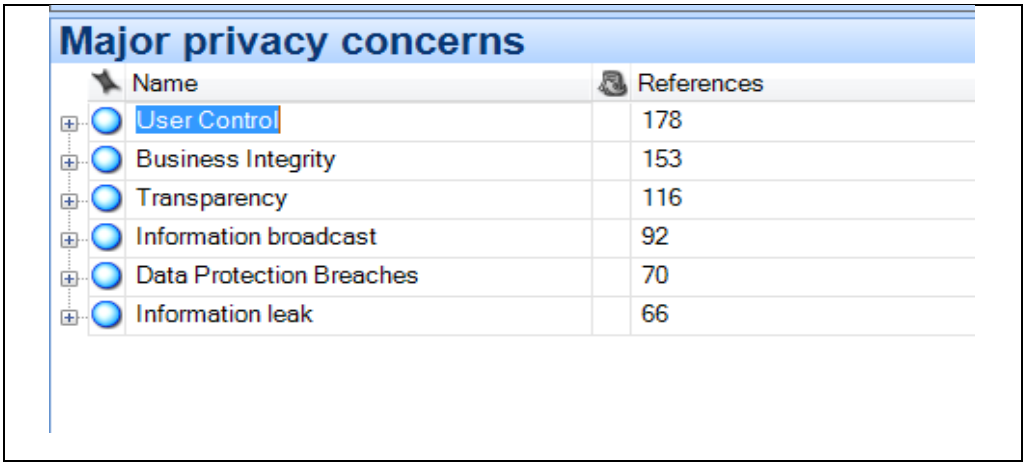


Fig 4.9: Frequency Counts of Major Privacy Concerns in Beacon

4.5 Analysis of Empirical Findings

4.5.1 User Control

The privacy concern 'User Control 'received the most counts compared with other major privacy concerns, which is not surprising as control over personal information is a recognised factor included in most information privacy studies in IS (e.g. Malhotra et al., 2004; Smith et al., 1996). Indeed, SNS users were seriously concerned about the leakage of their privacy as they were not able to control such leakage or the privacy risks associated with it. This concern divides into eight sub-concerns such as: 'Lack of control', 'Not opt-in', 'Universal opt-in', 'Granular control', 'Privacy diffusion', 'permanent data deletion', 'Control user behaviour', and 'Complex privacy controls'(see Figure 4.10 for frequency counts of each concern).

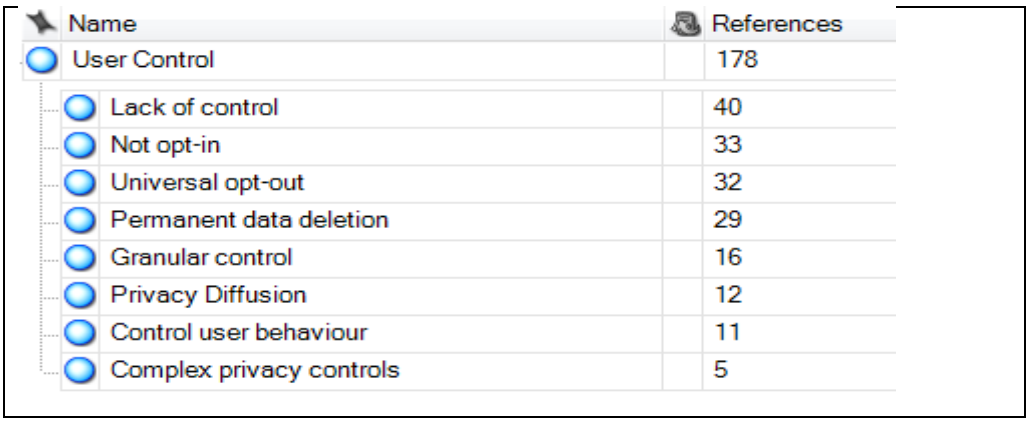


Fig 4.10: Frequency Counts of Sub-Concerns within Major Concern of User Control

Within this group, 'Lack of Control' received the highest count (40) given the intrusive nature of Beacon as perceived by SNS users and echoed in the following quotes from the blog commentary (See Appendix A for more text extracts supporting privacy themes):

> *It's the *design* principle "user in control". If the design doesn't clearly communicate what's going on, and how the user can influence it, the user can't feel in control.*

Here, Beacon's ability to automatically leak users' shopping secrets on Facebook's newsfeed without their consent or knowledge was considered a serious infringement to users' privacy. Users felt outraged because they were given no choice about what they wanted to reveal and they did not want Facebook to disseminate their private information. Certainly, no prior notice was given to Facebook users of Beacon's presence and its ability to leak the actions performed on third-party websites back to their profiles. Regarding the timing of the Beacon launch, a blogger commented:

> *With the holiday season approaching and the growing number of online shopping, friends now need only check their Facebook news feed to find out what presents they can expect to receive, taking away the best part of gift giving, the surprise.*

This quote suggests that the timing of Beacon's launch near Christmas was also a factor in fanning the flames of user's angry reaction to the leakage of their private actions performed on third-party websites: people felt more vulnerable when their Christmas shopping details, which people usually like to keep secret to surprise their friends and family, were leaked. Indeed, Beacon was perceived to have ruined the whole philosophy of gift giving, to have trespassed private boundaries and consequently people felt powerless. A user made an interesting comment whilst challenging the business philosophy to manage user information:

> *Decision makers like to feel that they're in control of client interactions. Sometimes they insist on control even when it would be rational to follow*

> *the client's lead. Where privacy is concerned, they want to decide what clients should want, rather than listening to what clients actually do want.*

In this particular instance, Facebook management was strongly criticised and challenged because they failed to listen to what social network users wanted regarding the privacy of their information but rather dictated what they should want. Such organisational thinking may be the consequence of Facebook management's assumption that, as users voluntarily share personal information on their profiles and with their friends, they are (or will be) willing to share the actions they perform on third-party websites with their Facebook friends. In fact, this was not the case, as highlighted by the aforementioned comment that users want to decide themselves what information about them should and should not be communicated. Such user thinking highlighted another concern within this group of concerns – 'Granular Control'. As a user remarked:

> *We like having the control to limit specifically who can see our profiles (including being able to say that everyone who isn't a graduate student at a particulate school is excluded, or all high school students so our campers or students can't see us… or even limiting the exposure of certain photos poster to just a select group of friends).*

This reaction suggests that SNS users expected to have sophisticated controls to manage privacy of their information, rather than just 'yes or no' based control mechanisms built into Beacon. It seems SNS users do not want absolute control over their personal information, but rather granular control to manage their privacy. This is also echoed in another blogger's comment: *"that's not being hypocritical. being able to choose what you share with your acquaintances and not being able to choose what is being shared are two different things. i'd be glad to see beacon die".* Such user expectations are also highlighted in another concern within this category, 'Universal opt-out'. The lack of ability to opt-out of Beacon altogether was perceived as an invasion of users' privacy and users' accused Facebook of employing aggressive and unfair practices. As a user notes "*Not providing a single, global opt-out of the whole program is completely unacceptable. Taking the attitude that we should just "get used to it" is in no way an excuse for not providing one".* It may be because SNS users wanted to completely opt-out of the Beacon programme so that it could not leak their private transactions on third-party sites back to

Facebook profiles. Another likely explanation is that users felt that their behaviour was being controlled by Facebook, another perceived sub-concern within the 'User Control' privacy theme. The sub-concern 'Control User Behaviour' is interesting as it highlighted previous privacy practices of Facebook whereby they encroached on the limits of users' privacy (e.g. the launch of News feed option) first and then provided slightly better privacy controls to appease consumers, though crossing previous privacy boundaries. For example, a blog user sums up this organisational behaviour in the following comment:

> *In each incident, Facebook pushed the boundaries of privacy a bit further and, when public outcry took place, retreated just a wee bit to make people feel more comfortable.*

This shows sophisticated user thinking as they were able to link and evaluate previous organisational practices with then-current practices – this further raises question about the integrity of organisations, which was perceived as a major privacy concern. Similar user sentiments were highlighted in sub-concern 'Not opt-in' where a user shows distrust of Facebook's actual practices as opposed to the claims made for them:

> *The big PR mistake was that Facebook caved on opt-in (or tried to) when they really didn't need to. opt-out isn't a cardinal sin, and opt-in isn't holy... but trying to position yourself as opt-in when you're still mostly opt-out was probably not the right call. and that's the PR issue they're dealing with right now, not that opt-out was the default.*

So, though SNS users' preferred choice when considering adopting Beacon was opt-in, to their surprise – which subsequently became their fury – Facebook positioned Beacon to be opt-in when it was actually launched as opt-out. Hence, it raised users' distrust about organisational practices which consequently called into question the integrity of online businesses.

4.5.2 Business Integrity

The finding 'business integrity' is explained by seven sub-themes – 'Untruthfulness and dishonesty', 'Monetise user data', 'Deceptive practices', 'User distrust', 'Repeated

privacy breaches', 'Disregard of user privacy' and 'Future abuse of personal data' (see figure 4.11 for frequency counts).

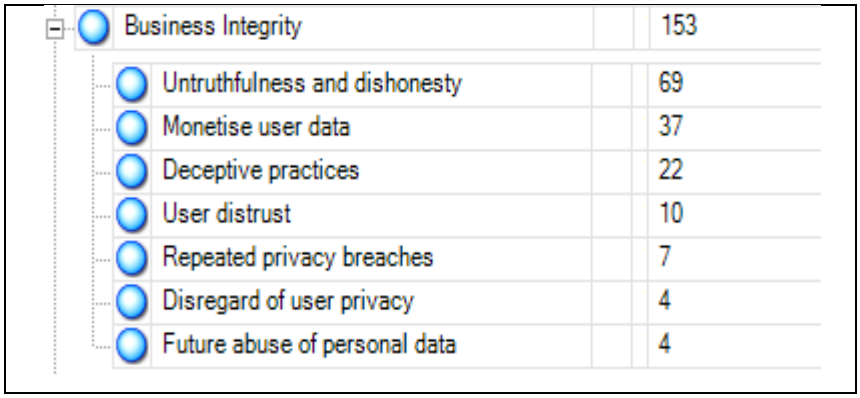


Fig 4.11: Frequency counts of sub-concerns within major concern of Business Integrity

The sub-privacy concern 'Untruthfulness and Dishonesty' received the most coverage (45%) within the group, almost twice that of the next highest sub-concern, with 69 user responses expressing outrage that the company did the opposite of what it had promised users. As a user noted:

> *It is the deceptiveness of the word "story" which really gets to me. "Sending a story" - this is a euphemism, right? When will the corporate world stop spinning profits as charity?*

Here, users perceived the insidious nature of Beacon's mendacity: as the company seemed to be offering a feature that might enhance user experience on the site, it was actually working underhand to leverage SNS user data for commercial gains. For SNS users such leakage of personal information infringed consumer privacy and freedom of choice and hence was a major reason for user backlash. Similarly, users considered organisational practices and behaviour dishonest, as a user remarked:

> *Company completely misrepresented (i.e. LIED) how beacon works. When asked if it tracks users whether they are logged-in or not, they said "no, absolutely not" Computer Associates uncovered that it was, to which they now say, "we don't do anything with that" Come on*

Again the integrity of the company was challenged by SNS users who believed Facebook was untruthful (whether intentionally or inadvertently) when they found out that the personal information of SNS users was even tracked (leaked) outside Facebook – a position which the company had consistently denied, but later admitted, as if they were not using that information for any purpose. However, the last position did little to calm down SNS users, who suspected that their data was a tempting source of revenues for organisations and hence was endangering their privacy. As a user commented:

> *No one knows when or to whom that perhaps obscure database will be sold or stolen.*

Consequently, SNS users raised the concern about monetisation of user data, which received 24% coverage of user responses within the business integrity category. Users were seen as being uncomfortable and furious about how Facebook had sold their activities to third parties. As a blogger confirmed: "*Instead we find that corporations will buy/sell/trade "private" information to anyone who pays as if there were no legal repercussion".* This concern reminds us of the central position and, for that matter, the practice of organisations: their readiness to collect and use of personal information for monetary advantage so long as they can avoid infringing data protection laws. Although the concern 'deceptive practices' did not have many responses (22), users seemed outraged by Facebook's handling and use of information. A blogger said: "*The arrogance and practices of Facebook are appalling. I can't wait for the next user revolt. I will first in line that is if I don't delete my profile first".* Here, SNS users felt helpless because they perceived organisational behaviour and practices as being aggressive and deceptive. The user comment which challenged Facebook's tardiness of response as being intentional and misleading – *"Facebook's slow response to those who did express anger in the platform showed that they were milking it for as long as they could - again* questioned the integrity of Facebook as users perceived these delaying tactics as being misleading and indeed harmful to their privacy.

4.5.3 Transparency

The major concern 'Transparency' subdivides into six sub-concerns related to the leakage of private data of SNS users (see figure 4.12). The sub-concerns are: 'Lack of

transparency', 'Lack of user incentives', 'Uninformed user tracking outside SNS', 'Lack of user awareness', 'information ownership', and 'Self Accountability'. These concerns emerged as a consequence of the leakage of personal data without the notice and knowledge of SNS users.

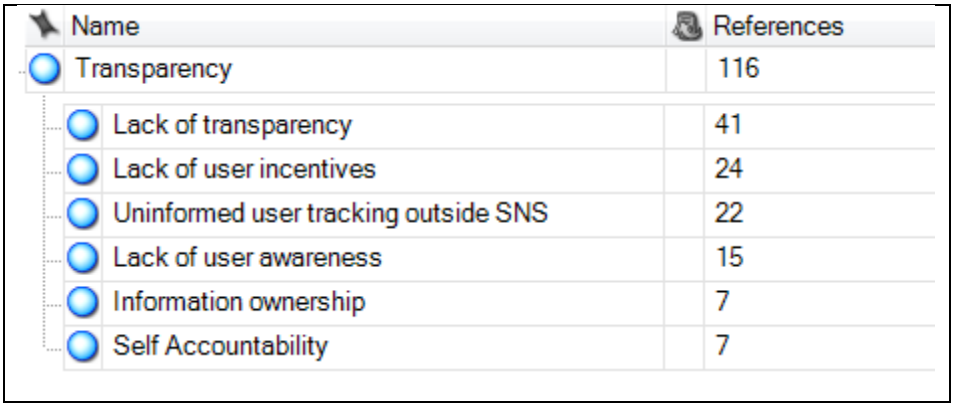


Figure 4.12: Frequency counts of sub-concerns within major concern of Transparency

Amongst these six sub-concerns, 'Lack of transparency' was considered serious by SNS users, receiving (35%) coverage within this group of privacy themes. The analysis suggests that users were upset as the purpose , visibility and presence of Beacon was not made clear to them regarding the collection and broadcast of users' actions performed on third-party websites back to their SNS profiles. What further raised users' concerns was that online businesses disseminated (leaked) personal information without users' knowledge or benefit to them. A blogger neatly highlighted the issue of lack of transparency and commented:

> *In my view, the uproar was really about making it too hard to notice you were being Beaconed. You're right that most users don't even know about all the fuss - which is why the opt-out is fairly useless if they don't know to go and check the box, and their Christmas present secrets can still be spilled...*

Here, for SNS users it was the opaque information practices of Facebook which caused concern and were considered violations of their privacy. The secret tracking of SNS users across the web by Beacon without their knowledge surprised and betrayed SNS users as they were totally clueless about what was happening and about how to resist Facebook leaking their private information

across the web. SNS users in this particular instance of the Beacon launch were demanding some discrete knowledge of how the leaked information was being used and how to delete it. Another concern within this category was the 'lack of user incentives' as SNS users seemed to conduct a risk analysis to assess and compare the benefits of such sharing of their data with the associated privacy risks. As a user remarked:

> *I'm willing to let a company use data they collect about my behavior to target content and advertising to me. That should create a better user experience for me, and doesn't expose my individual behaviors to other individuals or organizations.*

So, SNS users seem to be willing to allow Facebook to use their data for advertising within Facebook provided they were compensated by having a better user experience on the platform. However, leaking their behavioural as well as personal data to other individuals and organisations (including third-party advertisers) was not acceptable to SNS users and that's why they showed an affective negative response to leakage of their privacy. This indeed showed sophisticated thinking on the part of SNS users who did not reject completely the idea of their behavioural information being used by first parties (i.e. Facebook in this instance – as long as they were offered incentives for such sharing and use of data for advertising within Facebook), but rather showed serious concern about the leakage of such information to third parties (i.e. other individuals and organisations).

However, at the same time, SNS users did not seem to forgo the right to own their information and perceived it as another sub-concern called 'information ownership': they believed they should be able to remove their information permanently, as they own that data. SNS users therefore urged that organisations should act responsibly to protect their privacy which they believe could rebuild user trust on social networks to foster their growth. So, SNS users expected 'Self Accountability', rather than putting burden on users to think how to protect their information from being leaked. As a user commented:

> *If developers try to predict what a "bad guy" can do with the latest feature, they may completely miss a major aspect of privacy: often people seek privacy for purely sociological reasons rather than any concrete risk.*

In this particular instance, SNS users challenged the classical designer thinking which perceives loss of user privacy according to the concrete harm which may arise when their information is leaked. Instead, they wanted and expected more social responsibility from organisations, which should move beyond just defining privacy policy as a matter of legal compliance to actually employing concrete measures to respect user expectations of privacy.

4.5.4 Information Broadcast

This concern was the most obvious because of the automated nature of information dissemination embodied in Beacon. SNS users perceived this as invasive disclosure as they were neither asked nor given the choice to completely opt-out of this feature. Therefore, this concern very much relates to lack of user control. Also, users questioned the integrity of businesses due to the secret broadcast (leak) of their private data to third parties. The automatic dissemination of data and broadcast of private data were perceived as serious concerns together constituting 85% of user responses within the category. A blogger noted:

> *Do they really think that I want my purchases, and other private information automatically going to my network of friends without my permission?*

Here, SNS users questioned the automatic broadcast (leakage) of their private data on the network without their knowledge and notice. Therefore, they considered this Beacon feature a grave violation of their privacy as they were offered no choice and no freedom to control their information. Such leakage of private data may result in 'embarrassing disclosure' for SNS users, as a blogger commented:

> *How would you feel if you were at a party and in the middle of your conversation with friends the guy hosting the party announced to everyone what your last purchase was and if you'd like one too you should come and talk to him? Kind of creepy huh?*

This again reminds us that it is not so much the collection and use of data by first parties (i.e. Facebook) that disturbs SNS users, but rather its leakage to third parties (both individuals and organisations). This was also neatly summarised by an SNS user simply as 'Cross pollination of information' which was a concern for the users.

Name	References
Information broadcast	92
Broadcast private data	41
Automatic information dissemination	37
Embarrassing disclosures	12
Cross polination of information	2

Figure 4.13: Frequency count of sub-concerns within major concern of Information Broadcast

4.5.5 Information Leak

The finding 'Information Leak' is explained by six sub-concerns (see figure 4.14 for the composition and frequency count). Though similar to automatic broadcast, this privacy concern highlights the consequences of data leak which in most cases result in third-party abuse, such as 'Identity theft', 'Stalking' , 'Third party information leak', 'Intrusion' and building 'Online archaeology'. The concern 'Third party information leak' received almost 50% of responses within this privacy theme. A blogger noted:

> *I can't believe any of these merchants sites (Blockbuster, Fandango, etc) would ever agree to send out that kind of information on a purchase of mine. It's absolutely astounding. I do NOT want that kind of information sent to ANYONE, Facebook or other.*

Whereas the previous privacy concerns related mainly to the broadcast of data from Facebook to third parties and Facebook to other individuals within Facebook, the concern this time as expressed by SNS users is the leakage of data by third-party websites back to Facebook profiles. Indeed, such leakage was surprising and disturbing for SNS users who raged against such corporate invasion of privacy which was a consequence of the integration of social networks and third-party websites. So, users felt vulnerable to privacy risks such as identity theft and stalking as the companies now had archives of

their online lives which may have been used to screw them in future, particularly the youngsters who share their personal life matters on the internet without understanding the repercussions.

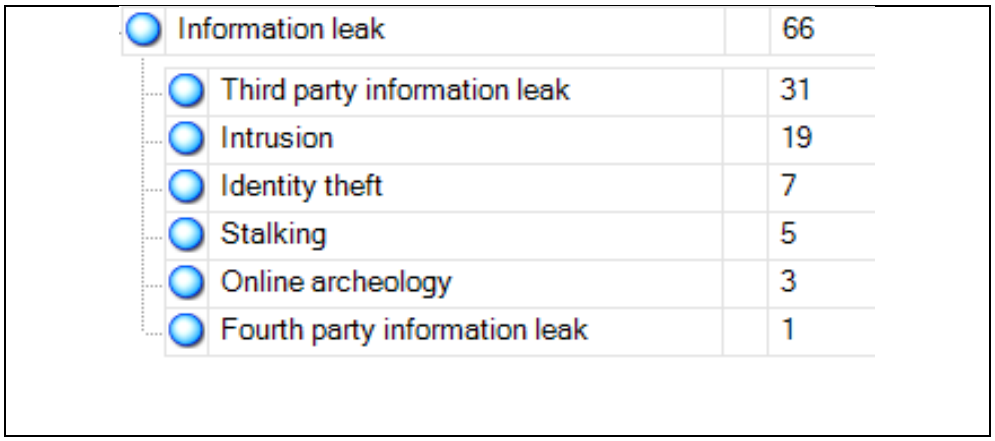


Figure 4.14: Frequency counts of sub-concerns within major concern of information leak

Even so, compared with the aforementioned major privacy concerns 'User control' and 'Business Integrity', 'Information leak' comprised of only 10% of user responses. This is mainly because all other concerns directly or indirectly related to the leakage of users private data. For example, SNS users' perceived such leakage of information a huge invasion of their privacy, betraying their trust, and they challenged the integrity of online businesses. Also, for SNS users, the issue and a challenge was not only that the majority of users have little or no awareness of such leakage but also the lack of user control. This, consequently, furthered user's concerns on the long-term implications of such leakage, as their private lives are now archived almost forever.

4.5.6 Data Protection Breaches

The privacy concern 'Data protection Breaches' did not accumulate as many responses as other privacy concerns, perhaps because most users have got used to them or perceived these as lack of control. However, this does not mean that these concerns are not important. Figure 4-15 provides a summary of this concern. Based on the frequency count, it appears that these concerns were considered mere irritants. The concerns 'Sharing of personal data without consent' and 'Sharing personal data without notice' together received more than 60% of user responses within this category. Certainly,

sharing (leakage) of users' personal without notice and consent was seen as a violation of privacy by SNS users, as one blogger commented:

> *So I'm definitely a little creeped out by the way this Beacon thing works, and don't really like the idea of notifications about my activity on the Web being broadcast to friends without my consent.*

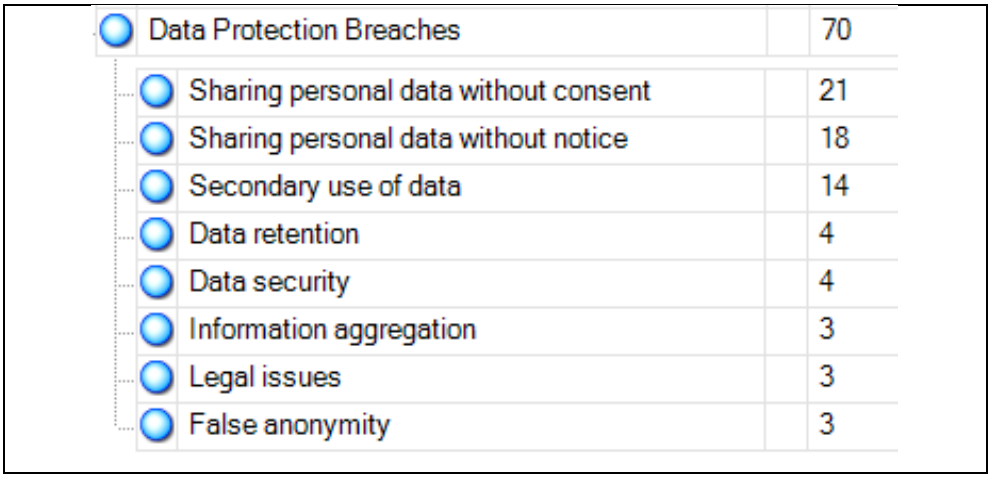


Figure 4.15: Frequency counts of sub-concerns within major concern of data protection breaches

So, Beacon not only violated SNS users' expectations of privacy but also data protection laws and hence committed data protection breaches. SNS users also questioned the limitations of the legal frameworks in their lack of ability to recognise such data protection breaches.

4.6 Conclusions longitudinal study – The Beacon Case

Beacon represented an innovative marketing tool within the burgeoning online social network environment (Jamal and Cole, 2009). What should have been a successful innovation, however, was damaged and ultimately withdrawn because the limits of the use of SNS data for social advertising in social networks were not well understood (Jamal and Cole, 2009). The qualitative investigation of user blog commentary was collected throughout the existence of Beacon during November 2007 (when it was launched) and September 2009 (when it was shut down). Subsequently, the blog commentary was analysed using grounded theory data analysis techniques of open and axial coding. The use of dedicated data analysis tool NVIVO facilitated the analysis process. The resulting

framework of privacy concerns – called Beacon privacy framework – identified six major privacy themes related to the leakage of SNS users' data. These include 'User control', 'Business Integrity', 'Transparency', 'Information Broadcast', 'Information Leak' and 'Data Protection Breaches'. SNS users display a sophisticated understanding that it is not the capture of information itself that is the concern – and therefore the limit – but how that information is leaked, combined, used and reused.

The Beacon privacy framework shows three levels of privacy concerns in relation to the scope and use of social ads tools in social network sites. The top level of concern represents serious boundary limits to the leakage and use of SNS data for business marketing in social networks and requires organisations to consider their business integrity and elements of user control. The second level highlights aspects that are considered moderate concerns. Here, the focus is more on the nature and type of metrics that should be created rather than the scope of use. The lowest level represents privacy irritations rather than concern. Interestingly, most business practice is centred on this privacy concern – data protection – and is therefore focused on safeguarding the consumer database rather than upholding the data mining practices of companies. The framework of privacy concerns coupled with the discussion offers organisations a concrete way of conceptualising the SNS business landscape, especially with regard to better understanding the limits of use and acceptance of social advertising tools causing privacy leakage in social networks.

CHAPTER 5: LONGITUDINAL CASE STUDY FINDINGS - THE CONNECT CASE

5.1 Chapter Overview

This chapter presents the research findings of stage two of the longitudinal case study – the launch of Facebook Connect, a tool which allowed users to take their online identity (Facebook profile) with them to third-party partner sites and share actions they performed on those sites back to their Facebook profile. However, soon after its launch social network users criticised the tool because of privacy leakage and considered it a grave violation of their privacy. SNS users' reaction published on blogs was collected in order to understand the unique nature and form of privacy concerns of SNS users. What should have been a successful innovation, however, was damaged because the nature and form of privacy concerns related to privacy leakage in SNS was poorly understood. First, the launch of Facebook Connect and the subsequent user backlash is discussed to contextualise this study. Next, the details of data collection and analysis are discussed. Empirical findings are presented thereafter, followed by the analysis of the results.

5.2 Study Background: The Launch of Facebook Connect

On 4th December 2008, the Facebook CEO announced the launch of Facebook Connect (hereafter called Connect) when he published a blog entitled *"Facebook Across the Web"* (Facebook, 2008). Connect allowed users to use their online identity (i.e. Facebook profile) across the web and share with their friends what they do online and get up-to-date information on what their friends are doing online. Also, it gave Facebook users the ability to take their Facebook privacy settings or preferences with them to other websites – a concept Facebook referred to as Dynamic privacy (Facebook, 2008). Like Beacon, which was withdrawn due to the hugely negative user backlash over privacy leakage, Connect also leaked user actions performed on third-party sites to Facebook friends (Stone, 2008). However, learning from the experience of the failed launch of Beacon, Facebook management were exceptionally discreet, introducing Connect gradually and pitching it as a privacy tool since it empowered users to take Facebook privacy preferences with them to third-party sites (Stone, 2008), unlike Beacon, which had not

offered any such privacy controls. Specifically, Facebook management took the following precautions in the launch and implementation of Connect.

- the early briefing of the civic advocacy group www.MoveOn.org – which launched a massive movement against Beacon over privacy (Stone, 2008).
- a careful and critical review of the information use practices of the partner companies before authorising them to use Connect.
- giving users control of up-dating privacy preferences and information they kept on Facebook across all partner websites (Moran, 2008), thus positioning Connect as a privacy enabling tool.

5.2.1 How Connect Works

Initially, Connect was partnered with 24 companies, among them Citysearch, CNN's The Forum, CBS' The Inside, TechCrunch, Govit, Howcast, and VLan. Figure 5.1 shows the Connect prompt suggesting that a Facebook user has visited a website, Citysearch, which then establishes a connection with Facebook so that the user can interact with Facebook friends and share stories through his/her Facebook wall and friends' News Feeds. Although Facebook had originally claimed to be giving full control to users to manage sharing of information between partner websites and the Facebook profile, the alert announced that the Citysearch website would be able to automatically post online stories back to Facebook – a worrying development indeed for SNS users.

Figure 5.1: Example of early Facebook Connect alert

An independent Webbmedia knowledge management group, (http://www.webbmediagroup.com), also confirmed this; they discovered that the user

comments posted on StumbleUpon and Citysearch, for example, were reposted automatically back to their Facebook news feeds without even seeking users' approval (Webbmedia, 2009). However, they later (10th February 2009) found that Citysearch did not post users' comments automatically without their approval. They summarised their analysis of Facebook Connect as *"while some users may not mind every comment being reposted on Facebook, the feature may alienate other users who desire better privacy" (p.3).* So, Facebook's claim that "[*They*]*'re shaping and defining what the internet is going to look like in the next couple of years"* (Tsotsis, 2010*)* stands in contrast to what they had said during the launch of Facebook Connect early in 2008 – that users would have full control to manage sharing their actions on third-party partner websites with their Facebook profiles. It is indeed users' behaviour and not Facebook which should redefine the future of the web – and users make the best guides when they are fully in control of managing the sharing of their information on the web.

Furthermore, the early Facebook Connect warning (see Figure 5.1) was in fact misleading, since no prior notice about Facebook Connect had been given to users – rather, users were ignorant of the tool's existence and of what it was going to do. Additionally, once the user was connected to Facebook, the second alert (see Figure 5.2) asked them whether to publish this story (e.g. a user review on Citysearch) to their Facebook profile. Although, Connect allowed users to universally accept publishing all stories, it still failed to provide users with a universal opt-out control.

Figure 5.2 Example of a Facebook Connect Alert to Publish a Story to Facebook Profile

Another annoyance for users was the inability to permanently delete activity data on Facebook servers (containing databases of users' actions on third-party sites), though Facebook did allow users to delete activity data on third-party websites. Whilst Facebook users' concerns about Connect's ability to share (leak) their identities with third-party websites and repost the actions they performed thereon back to their Facebook profiles were valid and significant, another important aspect of the problem related to users of third-party websites who were required to register in order to use the services. More specifically, Facebook Connect was given the ability to harvest third-party users' personal as well as behavioural data even when they were not Facebook users and had not consented. Alarmed by such leakage of privacy, technology users reacted strongly, and expressed their concerns via blog commentary. In order to understand the nature and form of privacy concerns, user blog comments were gathered and subsequently analysed.

5.3 Data Collection and Analysis

5.3.1 Data Collection

User blog commentary was collected between December 2008 and December 2010 for a two-year period, during which various businesses launched Facebook Connect in partnership with Facebook. This subsequently triggered user reactions which were published on blogs sites. As discussed in chapter two, Digg's recommender system was initially utilised to select the most rated blogs. Also, to gather an uninterrupted, up-to-date user commentary on blogs, the RSS feed commonly known as 'Rich Site Summary' or 'Real Simple Syndication (Duffy et al., 2006) was used. RSS is a standard feature which enables users to subscribe to the contents of a website/blog using newsreaders or aggregators (Duffy et al., 2006). Specifically, Google Reader (http://www.reader.google.com), a popular news aggregator tool, was used to automatically subscribe to up-to-date user opinions on the selected blogs gathered using the above keyword phrases on aforementioned blog sites. The use of the RSS feed provided much flexibility and ease to the researcher as all the updated comments were accessed automatically and consequently included in the data analysis.

Figure 5.3: Digg Rankings of Connect-related Blogs

Furthermore, in order to collect relevant blog comments, snowball sampling technique was also adopted, which allowed the researcher to collect the most relevant blog commentary. Figure 5.3 is a snip of Digg page showing selected blogs and corresponding ratings of blog readers. User ratings of some selected blogs were 442, 361, 325, 174, 151, 143 and 119. Blogs with higher ratings appeared on top of the page. Together, 35 blogs comprising 1014 user comments were gathered from various blog sites as shown in table 5.1. For example, 28.5% of blogs were collected each from news media sites and personal blogger sites, whereas 23% and 20% were collected from technology sites and social media sites respectively.

Type of blogs	Number of blogs	Percentage (%) to Total	Number of comments
News Media blogs	10	28.5	269
Technology blogs	8	23.0	242
Social media blogs	7	20.0	246
Personal blogs	10	28.5	257
Total	**35**	**100**	**1014**

Table 5.1: Details of blog commentary

5.2.2 Blogs Posting Time Series Analysis

To gain a better insight into the evolution of blog postings published as a reaction to Facebook Connect launch and the subsequent related data handling practices and events, a

time series of all blog comments was conducted which highlighted those posted on the first, second, third or later days during and after a blog's publication date. See Table 1 in Appendix B which shows the details of the comments posted on each blog. A predominantly large number of comments (90.24%) were posted on the first day of a blog's publication, followed by 5.82%, 1.58% and 2.36% on the second, third and later days respectively. This trend is depicted as a time series graph (see Figure 5.4). Interestingly, blog postings followed a pattern indicating that most comments were posted in reaction to an event or a trigger condition.

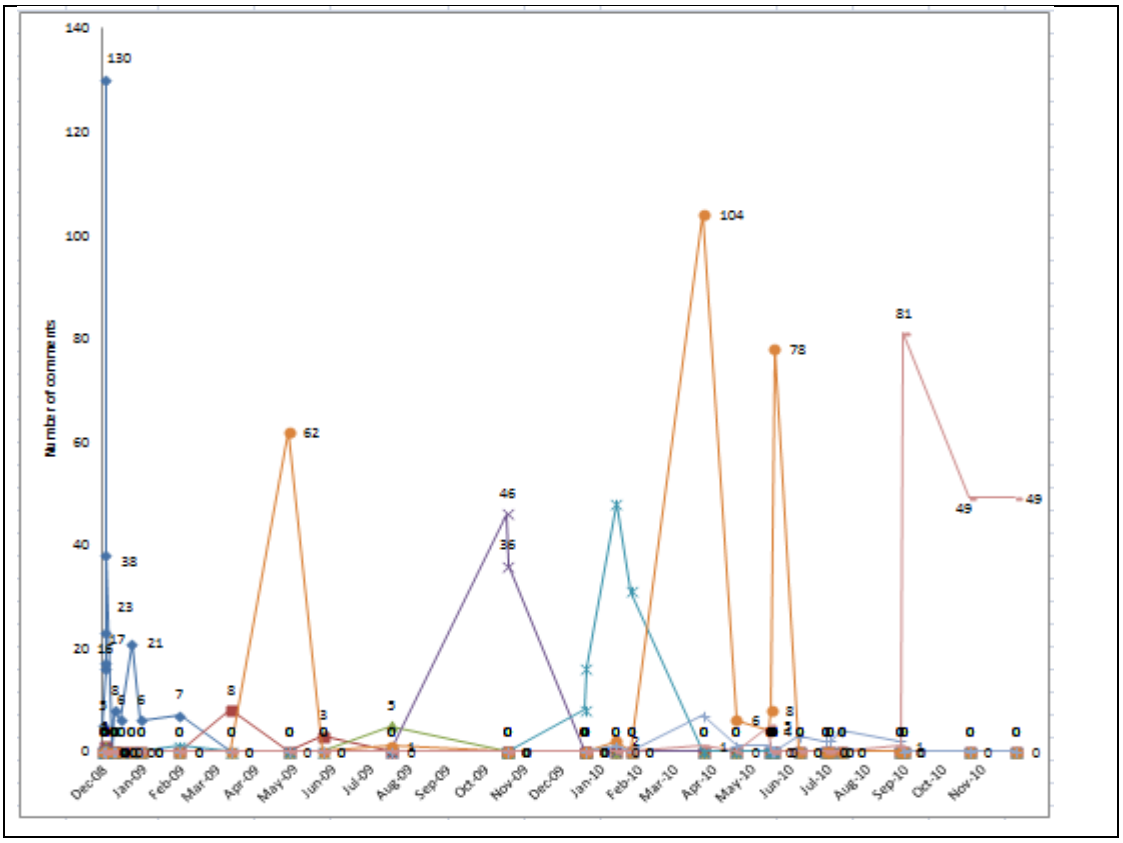


Figure 5.4: Time series Showing Incidence of Blog Comments

For example, the first and highest peak (see figure 5.4) shows the maximum number of comments posted in December 2008 when Facebook Connect was launched. However, posts dwindled between January 2009 and April 2009, to be followed by peaks of 62, 46, 48,104, and 81 respectively, mainly because of the trigger conditions such as the adoption of Facebook by a particular partner business site, changes made by Facebook management to the way Facebook Connect worked and others.

5.3.3 Data Analysis

As detailed in chapter three, this research adopted grounded theory data analysis procedures to analyse user blog postings related to the launch of Facebook Connect between December 2008 and December 2010. The use of grounded theory analysis procedures enabled the researcher to overlap the data collection and analysis procedure. Overall, the data analysis process is split into two phases. Open coding was used during first level of analysis to support microanalysis of user blog commentary. Axial coding analysis technique was used during the second level of analysis in order to relate and group categories identified in the first level of analysis. The rich qualitative blog commentary comprising 317 (A4 size) pages of text was analysed by using NVIVO. Open coding completed in two iterations. During the first iteration, an iterative approach to coding was used through active reading and re-reading of text which enabled the researcher to accommodate feedback in further refinements. Moreover, 'In Vivo' nodes were used as they allowed the nodes to remain close to actual data. For instance, the node 'Cross posting of personal information' is 'In Vivo'. The first iteration of open coding resulted in creation of 67 privacy themes or concerns (See appendix B).

In order to avoid redundancy and to double-check that any potentially relevant text had not been overlooked, the entire blog commentary was analysed again during the second round of coding. Accordingly, similar codes were consequently merged, resulting in a total of 43 nodes expressing the unique privacy concerns of SNS users. In the next stage of analysis, axial coding technique was used so that the privacy concerns could be grouped in a hierarchical manner to help identify relationships between nodes. These nodes are called tree nodes in NVIVO. The tree nodes represented major privacy concerns and their child nodes are referred to as sub-privacy concerns. Figure 5.5 shows six major privacy concerns along with the frequency counts. In order to determine which of these elements were perceived more important or severe by users, a frequency count of the major privacy concerns was also conducted as shown in figure 5.5. The frequency counts provided good insight into those elements that users considered mere irritants and those considered serious concerns about the leakage of privacy of SNS users.

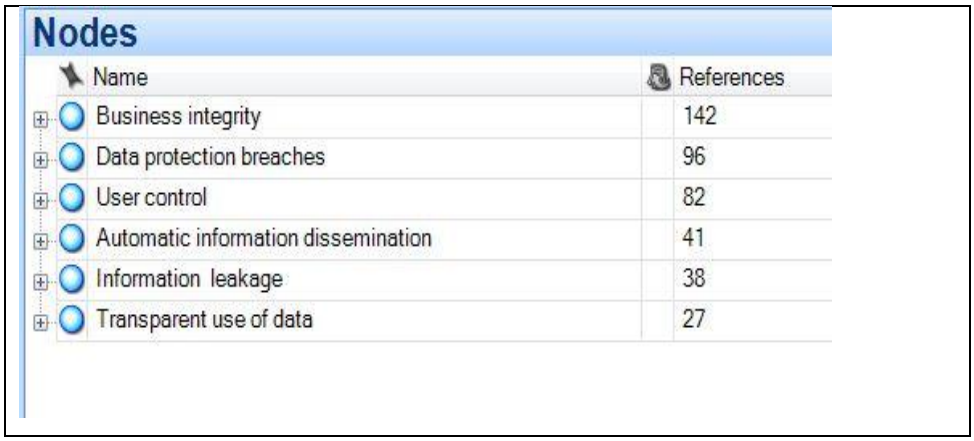


Figure 5.5: Frequency Counts of Major Privacy Concerns in Connect

5.4 Empirical findings

5.4.1 Connect Privacy Framework

The final analysis resulted in the development of a privacy framework called Connect privacy concerns. The framework shows 43 distinct concerns relating to the privacy leakage of SNS users. These are arranged in two levels of granularity moving from the general to specific (see Figure 5.6). From within the specific detailed concerns, 6 broader categories, labelled here as major privacy concerns, emerged. These relate to business integrity, data protection breaches, user control, automatic information dissemination, information leakage and transparent use of data. The sub-concerns outlined in Figure 5.6 provide greater clarity regarding the nature and form of the core concerns.

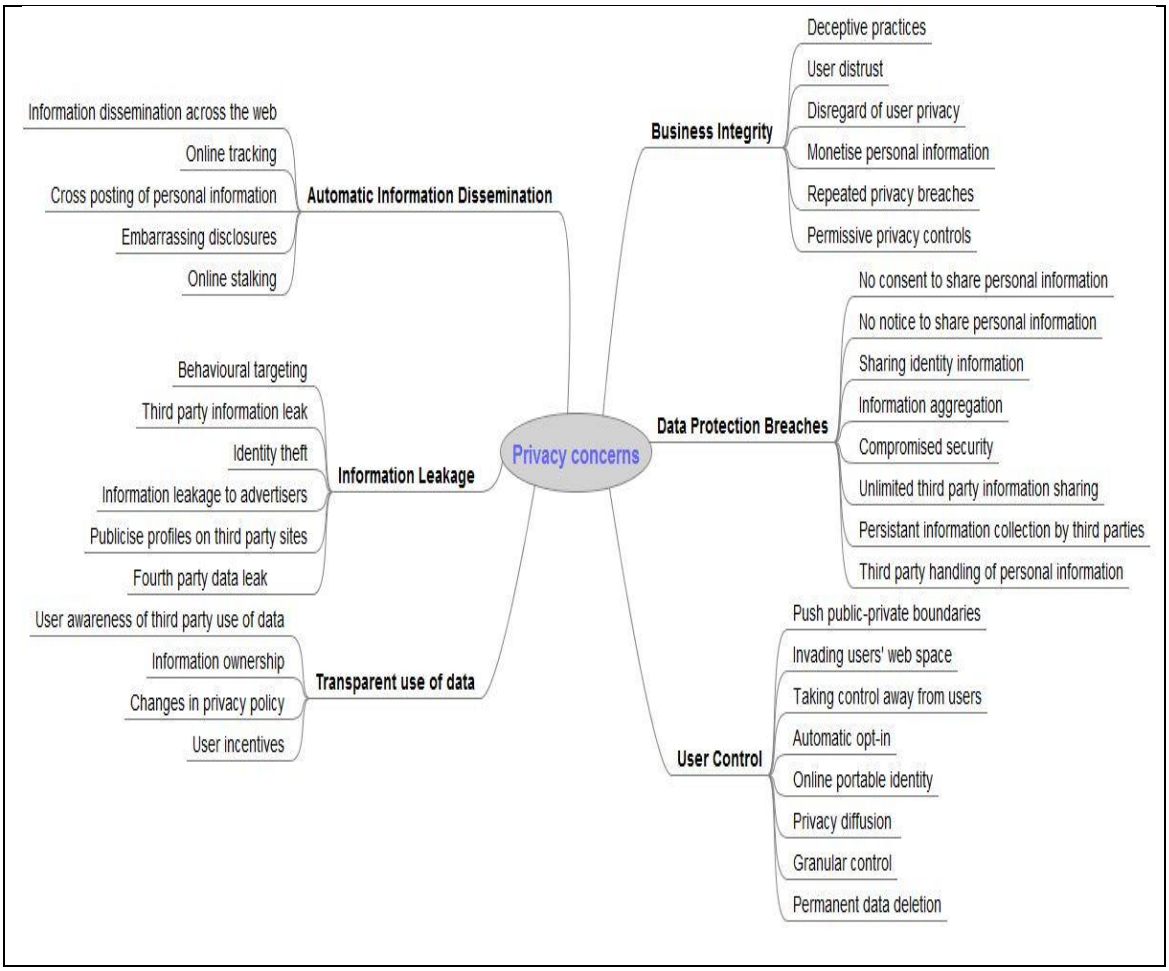


Figure 5.6: Connect Privacy Framework

For example, the privacy concerns 'business integrity' , 'data protection breaches' and 'user control' are viewed as the most severe breaches of privacy by SNS users representing almost 75 % of all responses (33% , 23% and 19% respectively). Mild form of privacy leakage is represented by 'automatic information dissemination' (10%) and 'information leakage' (9%). In contrast, 'transparent use of data' only received 6% of total responses and has been classified as an irritant.

5.5 Analysis of Empirical Findings

5.5.1 Business Integrity

The finding 'business integrity' is explained by six sub-themes – 'deceptive practices', 'user distrust', 'disregard of user privacy', 'monetise personal information', 'repeated privacy breaches' and 'permissive privacy controls' (see figure 5-7 for frequency counts).

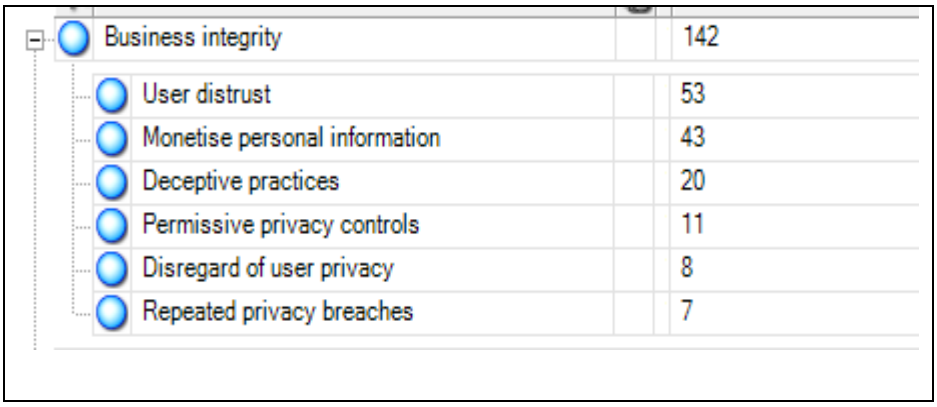


Figure 5.7: Frequency Counts of Sub-Concerns within Major Concern of Business Integrity

The sub-privacy concern 'user distrust' was perceived as the most severe of the concerns within this group as it covers (37%) responses representing 53 users who distrust Facebook's business practices. This distrust emerged not just as a consequence of a single breach of users' privacy, but rather due to repeated privacy breaches by Facebook. One user questioned Facebook's integrity and revealed distrust in the following remarks:

> *The company bringing this web to us is Facebook, the same people who had to be told by their users why Beacon was a huge mistake. Do you trust Facebook to control the next iteration of the web?*

This is interesting as it shows that at any one time SNS users' current privacy perceptions are dependent not only on contemporary issues but also on their past privacy experiences. This finding consequently highlights and confirms Altman's (1975) theory of privacy – also supported by Paul and Dourish (2003) – that privacy is a dynamic and temporal concept. This is also confirmed by a betrayed SNS user who put it bluntly: *"Sorry. I have no sympathies for Facebook. They have repeatedly breached their users' trust".* Moreover, a user perceived leakage of personal information as a grave invasion to users' privacy and remarked:

> *Leaking of the kind of personal info that facebook holds, including real names and in the case of unsophisticated users, friend lists, possibly even addresses, whereabouts etc is not just a breach of confidence.*

So, here for SNS users, leakage of their private data (which can identify a person) by Facebook goes beyond just a lack of trust as such leakage may have long-term implications for the privacy of SNS users. For example, the large numbers of teenagers who use Facebook and leave an online trail of their private lives (often sharing of their personal life secrets) are vulnerable to privacy harm throughout their lifetimes as they have no control to remove their personal information permanently. Indeed, SNS users were revealed to feel uncomfortable and outraged by how Facebook leaked their personal information to third-party sites outside SNSs. The concern 'monetise personal information' received 43 responses, being the second most severe concern in this category. This is understandable given the nature and purpose of Facebook, which was to leverage SNS data for social marketing. Criticising the commercial aspect of Facebook Connect, a user remarked:

> *Because this is their one chance of building a monetization engine. And just as it is becoming the underpinning of a brand new money-making scheme*

The quote suggests that SNS users were aware of the hidden commercial potential and purpose of Facebook Connect; hence they challenged it on privacy grounds as they believed the leaked information could get into the hands of third-party advertisers and aggregators. Another concern within the 'business integrity' privacy theme is 'deceptive practices', which covered 20% of the total responses in this category. This concern seems typical given the repeated privacy breaches as experienced and highlighted by Facebook users. A user sums it neatly as:

> *If they show one version to users initially, and another when they are criticized, then it only proves an intent to mislead.*

SNS users in this particular instance showed a neat understanding of the deceptive practices of Facebook because they observed the company initially use their personal information for commercial gains and thereafter cancel few features or introduce new privacy controls to appease users, thus pretending they care about users' privacy when actually they do not.

5.5.2 Data Protection Breaches

This concern subdivides into eight sub-concerns – 'no consent to share personal information', 'no notice to share personal information', 'sharing identity information', 'information aggregation', 'compromised security', 'unlimited third party information sharing', 'persistent information collection by third parties' and 'third party handling of personal information' (see Figure 5.8 for frequency counts).

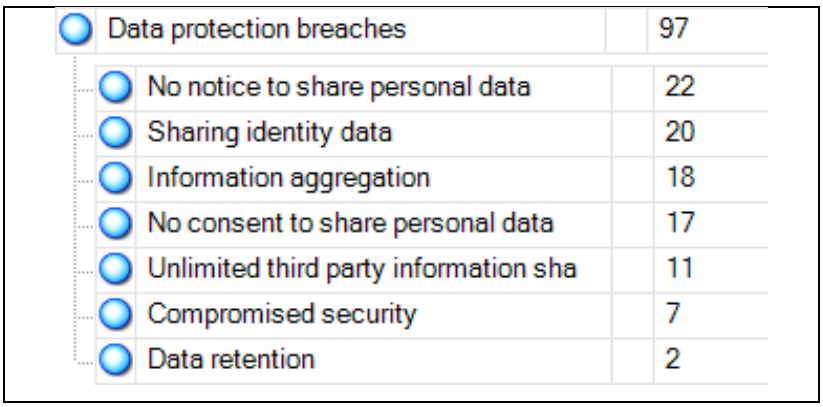


Figure 5:8. Frequency Counts of Sub-Concerns within Major Concern of Data Protection Breaches

The privacy concern 'data protection breaches 'received the most counts after 'business integrity'. Amongst these sub-concerns, the concern 'no notice to share personal information' received the most (23%) responses. This concern is highlighted by a blogger who commented:

> *I am feeling just like Pat #2, BETRAYED and VIOLATED to the hilt!! IF, I had been asked, maybe I would have been ok with it, MAYBE!! But at least I would have been given the option!! But I WASN'T!! I could have PREPARED all of my accounts so that my privacy wasn't in jeopardy!!!*

Here, the user shows flexibility and a willingness to check privacy control so as to protect the privacy of information, provided the company had consulted the user. In this particular instance, the user seems to treat privacy as a relative concept which depends on the availability of choice. Notice and choice are an integral part of all fair information practices (FIPS) which aim to safeguard users' privacy through availability of choice and consent, amongst other principles. Surprisingly for SNS users, Facebook in the particular instance of the launch of Connect, as previously with Beacon, failed to offer such

fundamental principles to safeguard SNS user privacy. Another concern which is perceived as serious by SNS users is the leakage of personally identifiable data. As a technology user opined:

> *The data shared includes names, user IDs, and other information sufficient to enable ad companies such as the Google-owned DoubleClick to identify distinct user profiles.*

Unlike the sharing of behavioural data, when personally identifiable data is shared with third-party advertisers or Ads networks (e.g. DoubleClick), they gain the ability to identify a particular person. This further raises concerns for SNS users as the availability of personal data to third parties is not without privacy threats, such as identity theft and stalking. Furthermore, another concern 'information aggregation' is also considered a grave violation of users' privacy as leaked behavioural information can be combined with personal information giving Facebook and third parties the ability to personally identify a an SNS user. Hence, a user points out:

> *What concerns me about Facebook's attempts to link all of your web activity to your profile is that so much of what you do online can be directly attributed to you.*

Here, SNS users seem to feel helpless regarding the aggregation of their web activity with the personal information on their Facebook profile in order to identify them. People indeed felt threatened as they believed in the private nature of the internet, whereas such linking of their private web actions with their profile information changed completely the fabric of the internet as a medium to support individual's private actions; hence, people feel betrayed and threatened by such privacy leakage.

5.5.3 User Control

This concern was found due to the aggressive information processing practices of online companies that result in users' feeling not in control of their personal information. Specifically, the major concern 'user control' divides into eight sub-concerns such as – 'Taking control away from users', 'Invading users' web space',

'Automatic opt-in', 'online portable identity', 'Privacy diffusion', 'Granular control', 'Permanent data deletion of data' and 'No universal opt-out' (see figure 5-9 for frequency count and breakdown into sub-concerns). Users felt betrayed as the invasive information processing practices that cause leakage of users' privacy took control away from them. A blogger expressed the sentiments as:

> *There is no way to effectively control the flow of information once it leaves Facebook or to vet the companies it goes to, or those they then sell it on to.*

SNS users felt powerless as they have no means to control who can do what with their information leaked to third parties since the information can be stored, aggregated, used, and re-used in entirely new ways. Thus, users perceived it as an invasion of their privacy. Another, user commented on such leakage:

> *Privacy controls allow you to manage that (with great and wholy unjustifiable difficulty in Facebook's case), but leaking info to advertisers renders the controls completely ineffective. You can bet that if we allow this to continue organized crime WILL find a way to get hold of some of that data and exploit it.*

In this particular instance, Facebook Connect's ability to track and leak SNS users' data across the web is strongly criticised by users as unjustified and complicated as it takes control away from users which feel threatened since third parties can exploit their data indefinitely even without their knowledge and notice.

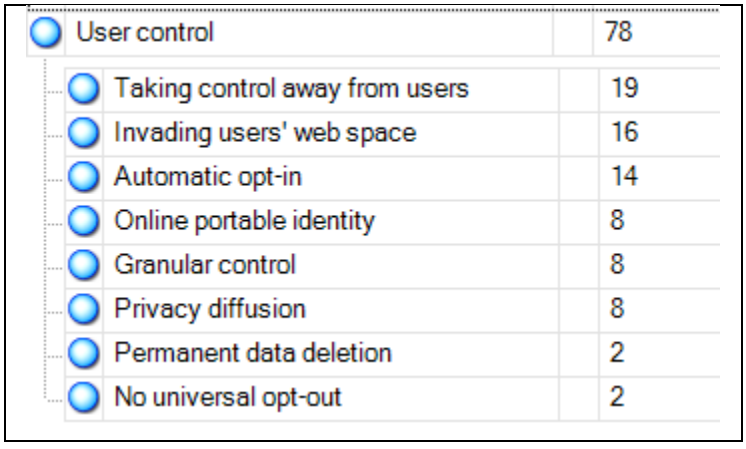


Figure 5.9: Frequency counts of sub-concerns within major concern of user control

Another blogger remarked whilst highlighting and comparing the leakage of data in online and offline world:

> *You wouldn't give control of your identity to someone in real life (at least with major regulation), so you sure as hell shouldn't do it on the Internet where everything is indexed and persistent. When will people wake up?*

So, for SNS users, such leakage of their privacy can have long term implications given the persistent nature of digital information and that they have no control over how to remove their personal information permanently. This is also echoed by another user as:

> *Even after you remove information from your profile or delete your account, copies of that information may remain viewable elsewhere to the extent it has been shared with others*

Another concern which agitated SNS users was 'Automatic opt-in' as they were not consulted, but were automatically opted in to share their private actions outside Facebook to third-party businesses. As a user pointed out:

> *That's so awesome that FB will automatically opt all of us in to share our data with the mysterious and unidentified group of approved third party sites.*

Because SNS users were not given a choice but rather opted in automatically, they have no clue who will do what with their personal information and felt threatened. Another concern, 'privacy diffusion', becomes relevant here because the leaked user data is a source of privacy diffusion as third parties (both individuals and organisations) may share (leak) this data, thus causing further harm to users' privacy.

5.5.4 Automatic Information Dissemination

This concern was the most obvious one because of the automated nature of information dissemination embodied in Facebook Connect. Specifically, this concern divides into sub-concerns - 'information dissemination across the web', 'cross posting of personal information', 'embarrassing disclosures' and 'push public-private

boundaries' (see figure 5.10 for frequency count and breakdown into sub-concerns). Indeed, SNS users perceived 'information dissemination across the web' as a serious threat to their personal privacy. As a user remarked:

> *However, the linking of Pogo and facebook provides way too much information between the two.*

Since Facebook Connect gave the ability to third-party sites (e.g. Pogo in this case) to leak users' data, users feared such integration was dangerous as it allowed both companies to hold and use unlimited data about them. This consequently caused concerns to their privacy as they had limited or no control to protect their privacy. Therefore, this concern very much relates to lack of user control.

The concern 'push public-private boundaries' was questioned by an SNS user as:

> *But, in the end, if the head of the company doesn't want his private posts/photos made public, why should I let them mine my data for advertising.*

The SNS user here directly challenged Facebook action which allowed third-party advertisers to use their private data publicly, which they had never liked as they consider it an infringement of their privacy. SNS users believe and expect that their Facebook data should not be leaked to third parties and also that their browsing on other websites should be kept separate to maintain the private-public boundary.

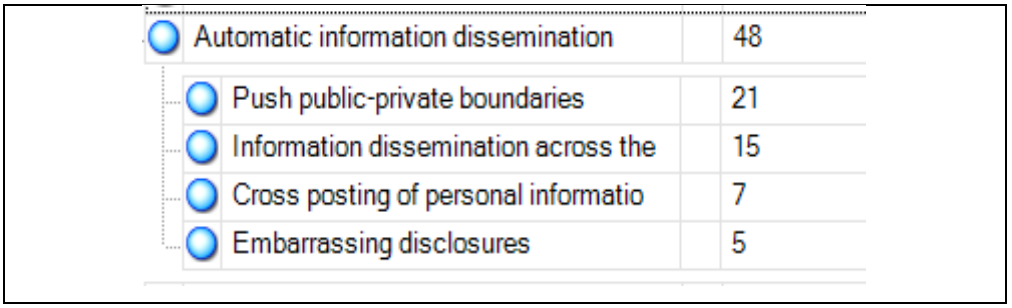


Figure 5.10: Frequency counts of sub-concerns within major concern automatic information dissemination

5.5.5 Information Leakage

The finding 'information leakage' seems most obvious given the very nature of Facebook Connect was to disseminate users' private actions across the web. This finding is explained by six sub-concerns (see Figure 5.11 for the composition and frequency). Though similar to automatic information dissemination, this privacy concern highlights the consequences of information leakage across the web, which in most cases creates business opportunities, such as linking users' web activity with personal profile information for targeted advertisement and even selling user data to third-party companies. Consequently, such business practices cause concern amongst SNS users and expose them to privacy harm. For instance, a user pointed out:

> *In the past few weeks, Facebook applications like FarmVille reportedly shared confidential personal information about users obtained through Facebook Connect with numerous different partners such as advertising networks.*

The leaking of SNS users' private information to third-party advertisers through Facebook Connect is perceived as a grave violation of users' privacy as it may become a source of further damage if such identifiable information is stolen or used by stalkers. This concern therefore has further implications for SNS users' privacy given their inability to control the leaked information. As a user expressing concerns over the use of leaked information by stalkers blogged:

> *The problem with Facebook, Twitter, and other businesses that allow you to track what people are doing online is that they appeal to a very small segment of our society: stalkers. These stalkers might actually be willing to pay Facebook or Twitter to stalk people one day.*

The tracking of SNS users across the web could create a digital dossier of online lives and hence become a tempting source of material for stalkers wanting to take advantage of such rich online archives of private lives. Indeed, this does threaten SNS user's privacy.

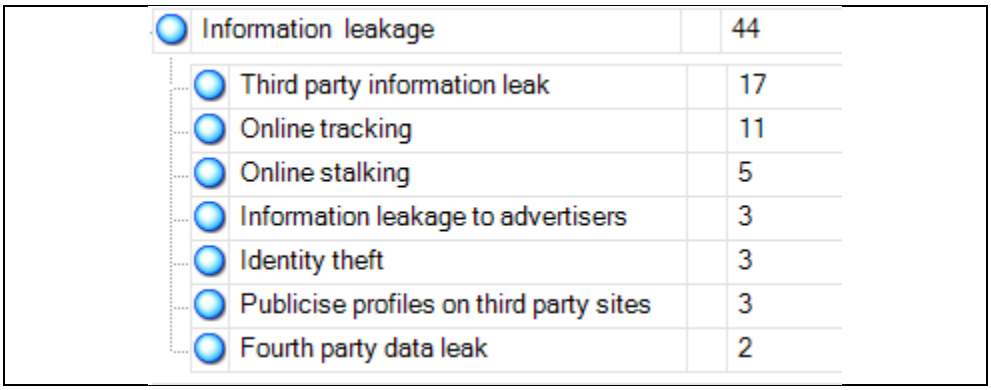


Figure 5.11: Frequency counts of sub-concerns within major concern of information leakage

5.5.6 Transparency

The privacy concern 'transparency' did not have many responses compared with others, perhaps because most users either do not have much knowledge of the risks and potential damage of privacy leakage or they might not be aware that such privacy leakage happens at all. This, however, does not mean that these concerns are unimportant. Figure 5.12 provides a summary of transparency issues which, based on the frequency count, appear to be considered mere irritants. For example, the concern 'user awareness of third party use of data' is neatly summarised by a blogger who pointed out:

> *As more people utilize Facebook to stay connected, more and more people are posting personal information without realizing the information is not always private*

Here, the SNS user showed concern to warn other users that they are not sharing their information only with their friends on social networking sites, but rather that SNS data is constantly being leaked to third party advertisers for commercial gains. Another concern, related to the 'push public-private boundaries', is 'changes in privacy policy' which is explained well by a user who commented:

> *Trouble is, with Facebook's latest revision to their privacy policy, there's a bunch of information considered "public" now that wasn't before. Things like your name, friends list, pages you "like"... all very much personal information that you no longer have any control over.*

Figure 5.12: Frequency counts of sub-concerns within major concern of Transparency

It is highly unlikely that users' could ever keep track of the changes Facebook makes in its privacy policy, and the low number of user responses explains this phenomenon well. However, the above user quote summarises neatly the habit businesses have of changing their information practices to get hold of even more data to exploit for marketing. So, SNS users view the launch of Connect as an endeavour by Facebook to exploit more SNS data for social marketing.

CHAPTER 6: EVALUATION AND DISCUSSION OF THE FINDINGS

6.1 Chapter Overview

This chapter elaborates on the findings of stage one and stage two of the longitudinal study. Specifically, privacy leakage frameworks developed during stage one and stage two are compared and consolidated into a taxonomy of privacy leakage concerns of Facebook users. This is then evaluated in the light of relevant literature. Finally, the taxonomy is applied to evaluate the response of LinkedIn to user backlash received when it launched social marketing tool. The taxonomy of privacy leakage concerns and subsequent analysis offer organisations a concrete way of conceptualising the SNS business landscape, especially with regard to better understanding the limits of the use of personal information for commercial purposes in a social network such as Facebook.

6.2 Delineation of research findings

In Chapter four, the empirical findings of stage one of the longitudinal study were conceptualised as the Beacon privacy framework, which improved our understanding of the fluid nature and form of privacy concerns of Facebook users. To recap, the stage one findings related to the launch of Beacon, a personalised marketing tool of Facebook whose features had to be radically revised and subsequently shut down on privacy grounds.

Chapter five presented the empirical findings of stage two of the longitudinal study conceptualised as the Connect privacy framework which provided further insights into the scope and limits of the broadcast of personal information by businesses for commercial gains. To reiterate, the Facebook Connect tool allowed third-party websites to broadcast users' actions performed thereon back to Facebook profiles. In order to provide a consolidated and a developmental view of the findings of stage one and two of this research, the Beacon and Connect privacy frameworks are compared and then combined to devise a comprehensive privacy framework (hereafter called taxonomy of privacy leakage concerns of Facebook users). The privacy taxonomy

aims to provide a cohesive fabric of the privacy violations. The author argues that this approach aligns well with the nature of the current longitudinal research which aimed to study users' privacy behaviour throughout a three-year period. To recap, this longitudinal study was designed and operationalised as two epoch case studies analysed in two staged research. Also, as one of the objectives of the current research was to investigate how privacy concerns emerge and subsequently evolve, the current research mandates a holistic view of the findings of both epoch cases. The resulting taxonomy of privacy concerns provides a concrete means for organisations to unpack the fluid nature and form of privacy leakage concerns of Facebook users.

6.2.1 Comparing Privacy Frameworks

To aid the consolidation process, the first logical option seems to be comparison of the privacy frameworks so that similarities and differences can be highlighted. Accordingly, Beacon privacy framework and Connect privacy framework were compared. The comparison is depicted well in Figure 6.1, which shows similarities and differences between the frameworks. For example, some privacy concerns such as 'Granular control', 'Deceptive practices', 'User distrust', 'Disregard of user privacy', 'User incentives', 'Data retention', 'Privacy diffusion', 'Identity theft' and others were common (i.e. not mutually exclusive) in both privacy frameworks. Hence, these were shown within the overlapping area of figure 6-1. However, many privacy concerns were unique (i.e. mutually exclusive) to one or other of these privacy frameworks. 'Complex privacy controls', 'Control over user behaviour', 'Untruthfulness and dishonesty',' Future abuse of personal data', 'Self accountability', 'Automatic information dissemination', 'Secondary use of data', 'Legal issues', 'Online archaeology', and 'Fourth party information leak' were found to be distinctly connected with the Beacon privacy framework. Similarly, the unique privacy concerns related to the Connect privacy framework included 'Invading users' web space', 'Online portable identity', 'Automatic opt-in', 'Changes in privacy policy', 'Information dissemination across the web', 'Unlimited third party information sharing', 'Sharing identity data', and 'Publicise profiles on third party sites'. However, it is worthwhile to mention here that some privacy concerns were semantically similar although different names were used to code them according to users' comments. For instance, the concern 'Information dissemination across the

web' is similar in meaning with the concern 'Automatic information dissemination', as both relate to the broadcast of personal data across the web.

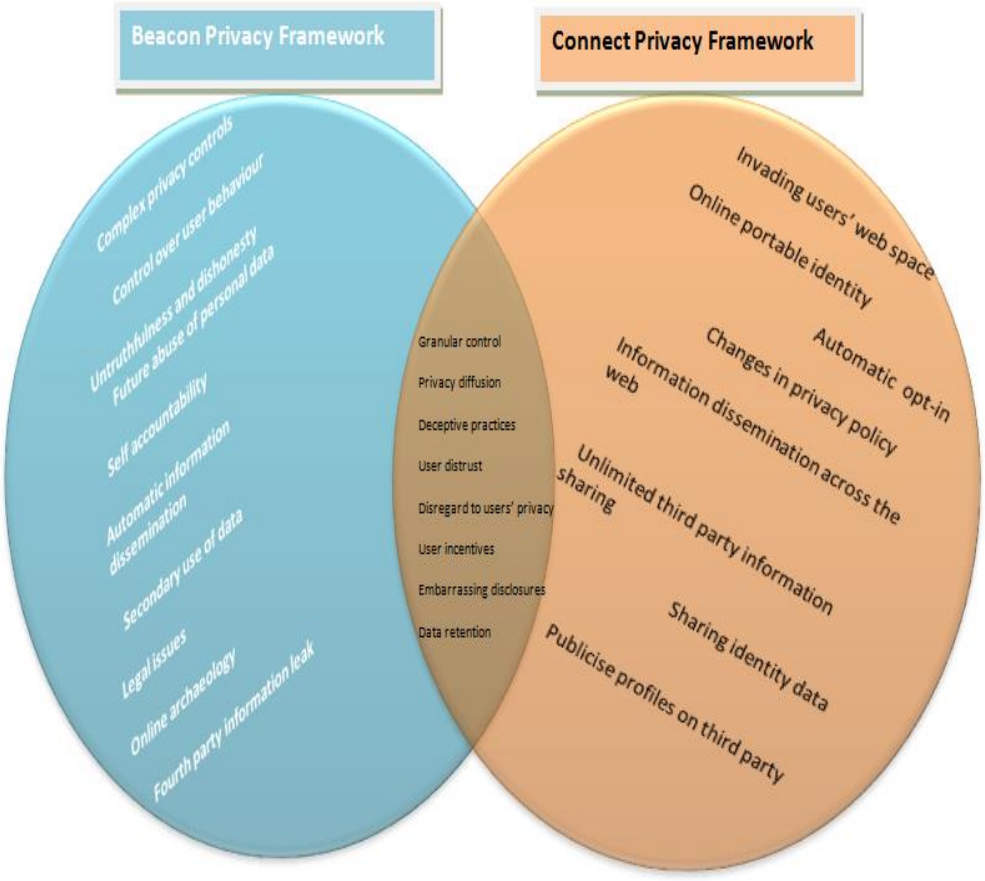


Figure 6.1: Beacon vs Connect Privacy Frameworks

The identified similarities and differences between the privacy frameworks enabled the researcher to develop a consolidated account of privacy leakage concerns of Facebook users, conceptualised as taxonomy of privacy leakage concerns. Indeed, the most sensible approach to achieving this was to include all distinct as well common privacy concerns of both privacy frameworks.

6.2.2 Consolidating Empirical Findings

The findings of stage one and stage two of this research study are compared and synthesised in order to devise a comprehensive framework of privacy concerns of Facebook users called a taxonomy of privacy leakage concerns. As mentioned previously, the author initiated the consolidation process by analysing the similarities and differences between the activities causing privacy harm to online users as identified in both the

privacy frameworks. Although, the approach seems basic, it was significantly analytic as the author had to go back to the actual data extract coded under these privacy themes in order to judge whether two concerns with similar names were actually similar and to what extent before deciding whether they should be combined as a single concern in the final taxonomy to avoid any redundancy. This was done with every single privacy leakage concern defined in both frameworks (i.e. 39 and 43 respectively in Beacon and Connect cases). Nonetheless, this process was demanding and time-consuming, yet rewarding as it culminated in a comprehensive taxonomy of privacy leakage concerns, which hopefully dispels any doubts as to what organisational activities are harmful to the privacy of online users and why and how these cause privacy problems. Indeed, the use of NVIVO was invaluable to aid the process of synthesis.

Table 6.1 details the consolidation process. The privacy concerns were arranged as major concerns and sub-concerns. Coincidentally, both the Beacon and Connect privacy frameworks had the same six major privacy concerns namely: 'User Control', 'Business Integrity', 'Transparency', 'Automatic Information Broadcast', 'Data Protection Breaches', and 'Information Leak'. Accordingly, all six major privacy concerns were kept in the final privacy framework. On the contrary, sub-concerns were both distinct and common in both privacy frameworks, as depicted in table 6.1. Nonetheless, some concerns did have different names but similar meanings and hence were merged as a single concern in the final taxonomy. For instance, sub-concern 'Lack of control' in the Beacon privacy framework, might seem different (at least in name) from the Connect sub-concern 'Taking control away from user'; however, review of the actual data extract of both concerns revealed very similar responses from users, so the author combined them as a single concern, 'Lack of control' in the final taxonomy. As a user remarked (in response to the Beacon launch) coded as 'Lack of control':

> *I personally control beacon by first logging out of my facebook session and removing all cookies before going shopping online. Of course, you can save those clicks by using a Firefox extension. This is of course not the ultimate solution. If the name/address you provide when purchasing online are checked against facebook member's database, beacon can be activated on you even without cookies.*

Similarly, a user showed concern (in response to the Connect launch) coded as 'Taking control away from user' as:

> *There is no way to effectively control the flow of information once it leaves Facebook or to vet the companies it goes to, or those they then sell it on to.*

Indeed, both the aforementioned coded extracts suggest that users were concerned about not having control over their data and hence were merged as a single concern in the final taxonomy called 'Lack of control'. This analytical procedure was maintained throughout the consolidation process. On the other hand, concerns with the same names were not automatically joined; their actual data extracts were also reviewed to make sure they had the same meaning before deciding whether they should be combined or not. For instance, sub-concern 'Granular control' from the Beacon privacy framework was reviewed through actual users' data extracts before being merged with the concern 'Granular control' from the Connect privacy framework. The sample data extracts coded at this concern in both frameworks confirm the notion that they are the one and the same and hence were merged as a single concern in the final framework.

A user showed concern (over the Beacon launch) as:

> *We like having the control to limit specifically who can see our profiles (including being able to say that everyone who isn't a graduate student at a particulate school is excluded, or all high school students so our campers or students can't see us… or even limiting the exposure of certain photos poster to just a select group of friends).*

Another user commented (in response to Connect launch):

> *Unfortunately, because you can't modify privacy controls for a Facebook Connect app, this means I can either show actions to all my friends (my profile is friends-only by default) or none of them.*

Finally, and interestingly, some privacy concerns in both frameworks had some common as well as distinct properties and hence could not be combined in a straightforward

manner. Therefore, these concerns were combined in such a way that all distinct as well as common properties were included to devise a final concern. As an example, sub-concerns 'Automatic information dissemination' (related to the Beacon launch) and 'Information dissemination across the web' (related to the Connect launch) have common properties such as 'information dissemination'. Nevertheless, each has distinct properties in addition, such as 'automatic' (Beacon) and 'across the web' (Connect). Thus, a new privacy concern was defined by taking common as well as unique properties in both the concerns and was named 'Automatic information dissemination across the web'. This ensured the addition of properties of both concerns and also avoided redundancy. Of course, this was accomplished after carefully reviewing actual data coded against these concerns. A user showed the concern (related to the Beacon launch) over automatic information dissemination as:

> *Do they really think that I want my purchases, and other private information automatically going to my network of friends without my permission?*

Similarly, another user registered concern (related to Connect launch) over information dissemination across the web as:

> *The intent is to bring people to the web so that when I am engaging on a site that is not Facebook, and I make a comment on someone's blog or make a post in a forum, my friends can get notified.*

To conclude, each and every privacy concern was reviewed carefully as already mentioned during the consolidation process and a taxonomy of privacy leakage concerns of Facebook users was devised, which is discussed below.

Beacon privacy framework		**Connect privacy framework**		**Taxonomy of privacy concerns**	
Major concerns	**Sub-concerns**	**Major concerns**	**Sub-concerns**	**Major concerns**	**Sub-concerns**
User Control	- Lack of control - Universal opt-out - Not opt-in - Permanent data deletion - Granular control - Privacy diffusion - Control over user behaviour - Complex privacy controls	**User Control**	- Taking control away from user - No universal opt-out - Automatic opt-in - Permanent data deletion - Granular control - Privacy diffusion - Invading users' web space - Online portable identity	**User Control**	- Lack of user control - No universal opt-out - Automatic opt-in - Permanent data deletion - Granular control - Privacy diffusion - Invasion of users' web space - Online portable identity - Complex privacy controls - Control over user behaviour
Business Integrity	- Untruthfulness and dishonesty - Monetise user data - Deceptive practices - User distrust - Repeated privacy breaches - Disregard of user privacy - Future abuse of personal data	**Business Integrity**	- Monetise personal information - Deceptive practices - User distrust - Disregard of user privacy - Repeated privacy breaches - Permissive privacy controls	**Business Integrity**	- Untruthfulness and dishonesty - Monetise user data - Deceptive practices - User distrust - Disregard of user privacy - Repeated privacy breaches - Future abuse of personal data - Permissive privacy controls
Transparency	- Lack of transparency - Lack of user incentives - Uninformed user tracking outside SNS - Lack of user awareness - Information ownership - Self accountability	**Transparency**	- No transparency - User incentives - User awareness of third party use of data - Information ownership - Changes in privacy policy	**Transparency**	- Lack of transparency - Lack of user incentives - Lack of user awareness - Information ownership - Uninformed user tracking outside SNS - Self accountability - Changes in privacy policy
Information Broadcast	- Broadcast private data - Automatic information dissemination - Embarrassing disclosures - Cross pollination of information	**Automatic Information Broadcast**	- Push public-private boundaries - Information dissemination across the web - Cross posting of personal information - Embarrassing disclosures	**Automatic Information Broadcast**	- Broadcast private data - Automatic information dissemination across the web - Cross posting of personal information - Embarrassing disclosures

Data Protection Breaches	- Sharing personal data without consent - Sharing personal data without notice - Secondary use of data - Data retention - Data security - Information aggregation - Legal issues - False anonymity	**Data Protection Breaches**	- No notice to share personal data - Sharing identity data - Information aggregation - No consent to share personal data - Unlimited third party information sharing - Compromised security - Data retention	**Data Protection Breaches**	- Sharing personal data without notice - Sharing personal data without consent - Sharing identity data - Information aggregation - Unlimited third party information sharing - Secondary use of data - Data retention - Data security - Legal issues - False anonymity
Information Leak	- Third party information leak - Intrusion - Identity theft - Stalking - Online archaeology - Fourth party information leak	**Information Leak**	- Third party information leak - Online tracking - Online stalking - Information leak to advertisers - Identity theft - Publicise profiles on third party sites - Fourth party data leak	**Information Leak**	- Third party information leak - Online tracking - Stalking - Identity theft - Publicise profiles on third party sites - Fourth party data leak - Online archaeology

Table 6.1: Consolidation Process of Privacy Frameworks

6.2.3 The Taxonomy of Privacy Leakage Concerns

The outcome of the analytical consolidation process is a taxonomy of privacy leakage concerns (shown in Figure 6.2). The taxonomical representation of privacy leakage concerns provided a useful means to classify and categorise privacy concerns. Admittedly, the taxonomy enabled the author to show the hierarchical structure of privacy concerns as well as the interrelationship between them.

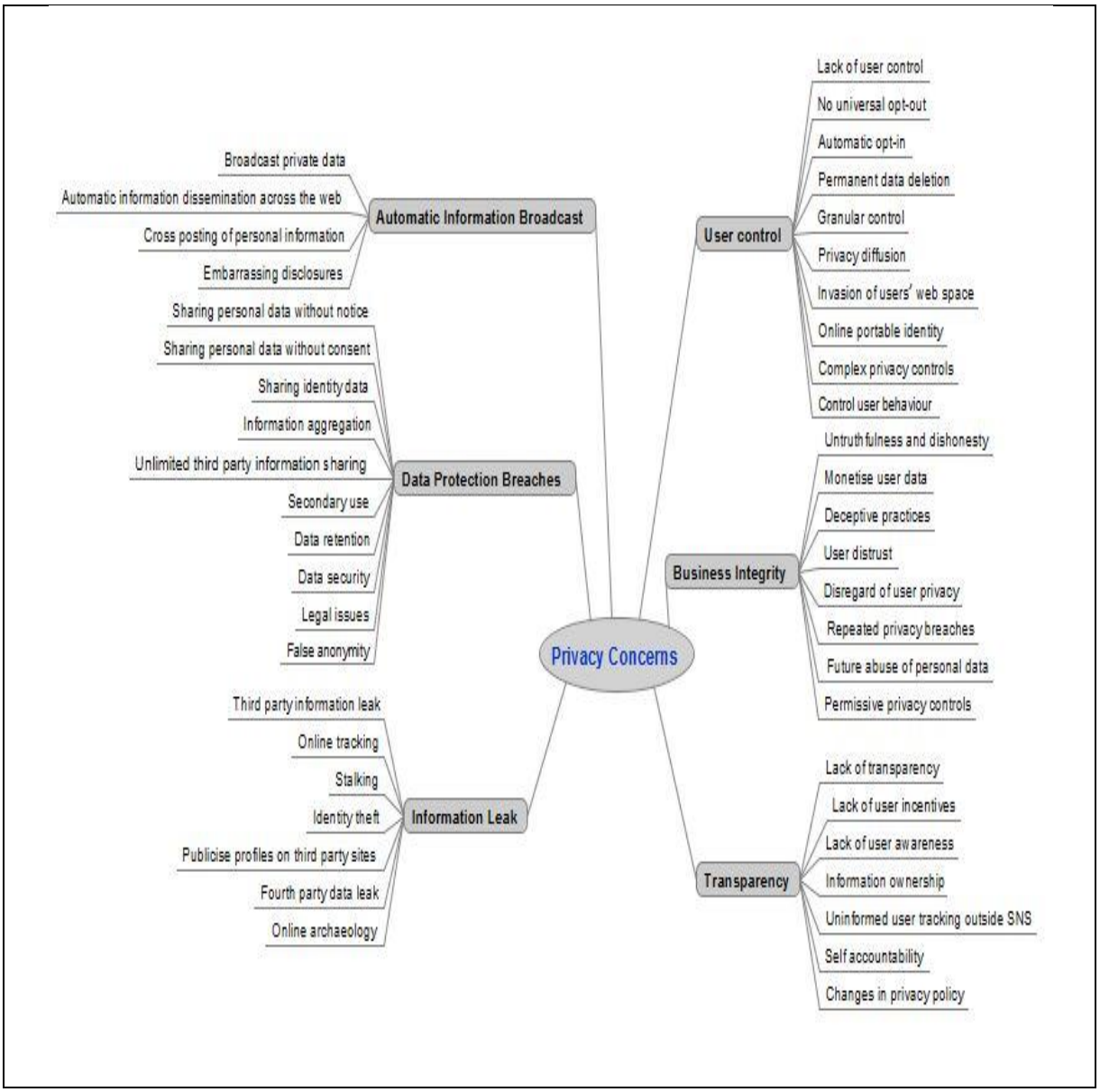


Figure 6.2: Taxonomy of Privacy Leakage Concerns

The resulting taxonomy shows 46 distinct user concerns related to privacy leakage. These are arranged in two levels of granularity moving from the general (abstract) to the

specific (see Figure 6.2). From within the abstract concerns, 6 broader categories, labelled here as major privacy concerns, emerged. These relate to user control, business integrity, transparency, automatic information broadcast, data protection breaches and information leak. The specific detailed concerns called sub-concerns as outlined in figure 6-2 provide greater clarity regarding the nature and form of the core concerns. Sub-concerns relate to the activities causing harm to users' privacy. These activities either encompass organisational processes and behaviour (e.g. disregard of user privacy or repeated privacy breaches, etc.), users' inability to control personal information (e.g. automatic opt-in or no universal opt-out, etc.) or legal issues (e.g. sharing of personal information without notice or consent, etc.).

Adopting a similar approach to that followed in chapters four and five, a cumulative frequency count of the sub-concerns was conducted in order to determine the perceived severity of these concerns. However, here the frequency counts of the concerns highlighted in both Beacon and Connect privacy frameworks were added in order to determine the cumulative counts of these concerns (see figure 6-3). This provided better insight into those elements that users considered mere irritants compared with those that represented actual information boundary transgressions.

For example, the privacy concerns 'business integrity and 'user control' are viewed as the most severe breaches of privacy by users, representing almost half of all responses (27% and 22% respectively). Mild breaches of privacy are represented by 'data protection breaches' (15%),'transparency' (14%) and 'automatic information broadcast' (13%). In contrast, 'information leak' only received 10% of total responses and has been classified as an irritant.

Major privacy concerns	Sub-privacy concerns	Frequency
Business Integrity	Monetise user data	80
	Untruthfulness and dishonesty	69
	User distrust	63
	Deceptive practices	42
	Repeated privacy breaches	17
	Disregard of user privacy	12
	Permissive privacy controls	11
	Future abuse of personal data	4
	Total	**298**
User Control	Lack of user control	62
	Automatic opt-in	45
	No universal opt-out	34
	Permanent data deletion	31
	Granular control	23
	Privacy Diffusion	20
	Invasion of users' web space	16
	Online portable identity	8
	Control over user behaviour	7
	Complex privacy controls	2
	Total	**248**
Transparency	Lack of transparency	46
	Lack of user awareness	30
	Lack of user incentives	29
	Uninformed user tracking outside SNS	22
	Information ownership	13
	Self Accountability	6
	Changes in privacy policy	4
	Total	**150**
Data Protection Breaches	Sharing personal data without notice	40
	Sharing personal data without consent	38
	Information aggregation	21
	Sharing identity data	20
	Secondary use of data	14
	Data security	11
	Unlimited third party information sharing	11
	Data retention	6
	Legal issues	3
	False anonymity	3
	Total	**167**
Automatic Information Broadcast	Broadcast private data	62
	Automatic information dissemination across the web	52
	Embarrassing disclosures	17
	Cross posting of personal information	9
	Total	**140**
Information Leak	Third party information leak	48
	Online tracking	30
	Identity theft	10
	Stalking	10
	Publicise profiles on third party sites	3
	Online archeology	3
	Fourth party data leak	3
	Total	**107**

Figure 6.3: Cumulative Frequency Counts of Major and Sub-Privacy Concerns

6.3 Discussion of Empirical Findings

Three primary stakeholders are involved in privacy protection – organisations, users and regulators. Therefore, the taxonomy of privacy leakage concerns presented earlier provides guidance at three levels: *organisational*, *user* and *legal*. This means that the leakage concerns that occur as a result of organisational behaviour and practices are labelled as organisational level privacy leakage concerns. Similarly, user and legal related concerns are termed as user level concerns and legal level concerns respectively. However, there is a constant interplay between these three level privacy leakage concerns as shown in Figure 6.4 which categorise privacy concerns in three levels.

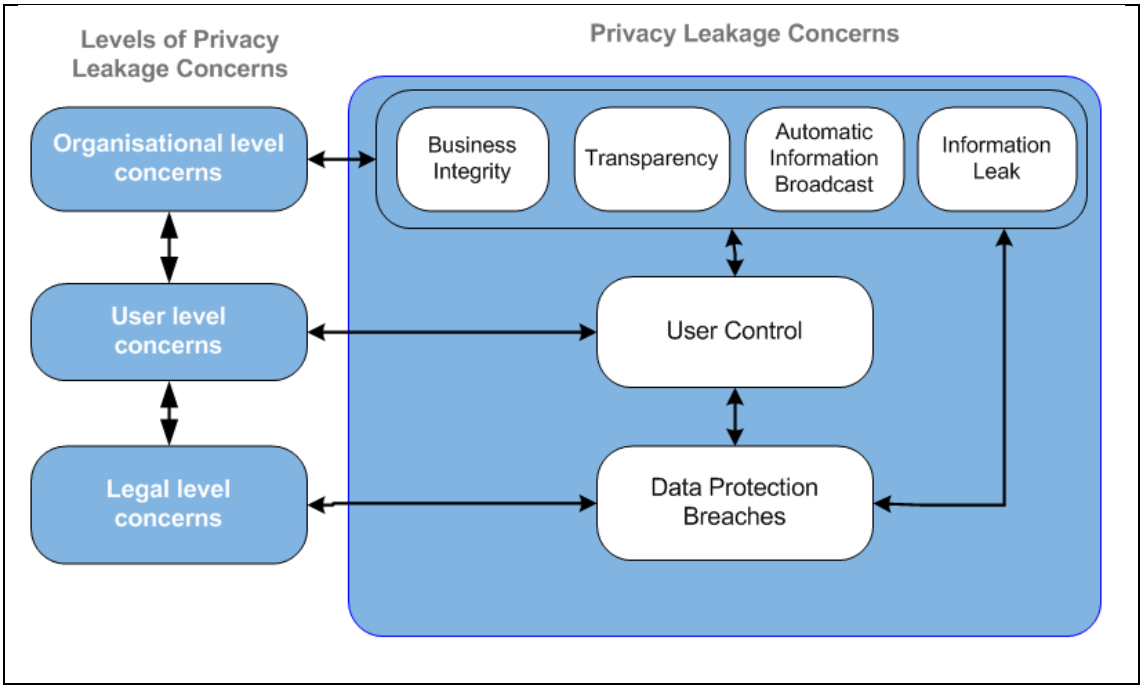


Figure 6.4: Levels of Privacy Leakage Concerns

To recap, the organisational level concerns encompass organisational behaviour as well as practices. Indeed, all three levels of privacy concerns are interrelated as shown in the above figure as they seem to trigger each other. For example, a legal level concern *'Sharing of personal information without consent'* invoke an organisational level concern such as *'Disregard of user privacy'* as users believed the company breach their privacy when it failed to obtain their explicit consent. This consequently, raised another concern *'User distrust'* as users perceived Facebook was being untruthful and dishonest during the launch of personalised marketing programme. Similarly, the organisational level concern

'Repeated privacy breaches' caused another concern '*User distrust'* which subsequently raised the issue of *'Self accountability'* as users believed the company did not seem to put in place any accountability procedures but rather held individual users to be accountable for their own privacy. Therefore, organisational level concerns raise legal concerns as users questioned the legal frameworks being inadequate in safeguarding users' privacy. Similarly, another organisational level concern *'Permissive privacy control'* prompted the concern '*User distrust'* which subsequently questioned the integrity of the business.

Thus, privacy leakage concerns are interrelated though the core issue remains invasive organisational practices and behaviour. Another example illustrating the inter-relatedness of concerns is the organisational practice 'automatic information broadcast' highlighting a user level concern 'permanent data deletion' as well as a legal concern 'Data retention'. When Facebook shared users' data with third parties without their consent and users were not given any control to permanently delete data, there were unclear data protection guidelines regarding data retention. Understandably, these three concerns were categorised at organisational, user and legal levels, although all three were connected. For greater clarity however, all three types of privacy leakage concerns are discussed separately below.

1. ***Organisational level Concerns:*** The first and most serious of the organisational level concerns relates to the integrity of business organisations. As Clarke (2006) noted in his reflections on the slow growth in e-commerce, *"consumer marketing is still characterised by aggression and dominance, no sensitivity to customer needs.....Instead of generating trust, marketers prefer to wield power"*. Because the visibility, purpose and presence of Beacon as well as Connect was not made clear to the user regarding the analysis and re-use of their browsing actions, the integrity of online businesses to safeguard the interests of SNS users was brought into question. Specifically, SNS user comments highlight the strong negative affective response created when it was perceived (rightly or wrongly) that a business was acting in an underhand or dishonest manner. Conceivably, that is why a simple apology by Mark Zuckerberg did not seem to appease Facebook users, who subsequently filed a $9.5m law suit against Facebook and its collaborating third-party businesses for their failure to provide notice and privacy controls in the launch of Beacon (Elden, 2010).

The case of Facebook Connect was not different as the company management had to retreat to provide better privacy controls soon after its launch because of the user outcry reflected in blog commentary. Indeed, the concern 'repeated privacy breaches' once again questioned the integrity of online businesses for their failure to protect the privacy of online users. Thus, both user backlash and the lawsuit represent a concrete commercial consequence of using personal information for business purposes in online social networks – as distinct from general e-commerce environments. Consideration needs to be given to the scope of information disclosure and use, especially the potential for legitimate trend analysis to create user backlash and ultimately undermine the commercial reputation of an organisation.

Another organisational level concern which accused Facebook of being involved in 'Deceptive practices' was when users found that Facebook continued to track, collect and share users' data even when they opted-out of the Beacon programme. This raised questions about the integrity of Facebook as users perceived the company was working in an underhand manner and hence user trust was betrayed. For organisations to build user trust and protect user privacy, it is vital that they should make their practices transparent, fair and accountable. Thus, rather than putting the whole burden of protecting their privacy on users via notice and consent, organisations should be held accountable for their actions and behaviour which cause breaches of users' privacy (Culnan, 2011). The developed taxonomy also highlighted 'Self accountability' as an organisational level concern wherein users expressed their inability to manage privacy of their personal information given rich features of social software and integration of SNSs with third parties. Instead, they required organisations to take responsibility for users' privacy and behave fairly.

Although slightly different in focus, both automatic dissemination and information leak reflect a user's concern with the consequences to them of unauthorised information sharing by organisations. Where automatic disclosure highlights aggressive data collection techniques and personal exposure through routine involuntary disclosures, information leak emphasizes the consequences of

unwelcome approaches. For example, studies show that SNS users were concerned about the selling of their personal information to online advertisers without their permission (Krishnamurthy and Wills, 2010). Similarly, Klingsheim and Hole (2008) note that information leak increases the chances of identity theft in online systems. Disturbingly for users, a study by Krishnamurthy and Wills (2010) found that most SNS, including Facebook, leak personally identifiable information either intentionally or inadvertently. Creating metrics that can protect the dignity and identity of a user is a novel approach to the design of social marketing tools and is also needed to support confidence and information sharing in the online social environment.

2. ***User Level Concerns:*** The second level of concern that businesses must be aware of is user control. This is a theme which is consistent with the work of Westin (1967) and Malhotra et al. (2004) who found that user control is an important determinant of privacy, which can often be exercised via approval, modification, and the choice to opt-in or opt-out (Malhotra et al., 2004). In both the current cases (i.e. the launches of Beacon and Connect), users were not given control to determine when, how and to what extent their personal information was communicated to others.

   This poses an interesting challenge – and opportunity – for using social network data for business marketing purposes. Existing research has simply shown that the majority of users want to have the ability to limit the use of personal information by third parties (Phelps et al. 2000). However, this research suggests that user control is a more nuanced and complex concept. The issue of concern is not simply the absolute amount of information held by the user (or organisation) but the complex interplay between context and content. And since the context of personal information is constantly changing, so the desire for user control is also constantly influx. The resulting challenge is an opportunity for designers and social network service providers to create more agile, streamlined metrics able to accommodate the constantly changing nature of personal information in ways that support the use and dissemination of personal data without infringing on user sensibilities.

Organisational level concerns such as transparency, automatic information broadcast and information leak also seem to interconnect with user level concerns. For example, the lack of transparency (of data use) was challenged by SNS users as unfair because organisational practices were opaque. The privacy violations by online social networks have attracted mass media coverage, such that privacy and data protection regulations have come under scrutiny by policy makers. Privacy advocates have also been active on pressing policy makers to rethink government policies. Consequently, government policies that support and encourage fair information practice principles (FIPS) in online environments are slowly emerging. This involves businesses agreeing to a) provide notices to consumers regarding how their personal data is collected, stored and shared with third parties and b) gathering user consent for such use of their personal data (FTC, 2012).

The taxonomy presented earlier suggests that opaque information practices are considered violations of trust by SNS users and negatively impact on the perceived credibility of an organisation's overall information practices (Harris Interactive, 2002). This has significant implications for the use of personal data for commercial gain as organisational practices are perceived as invasive by social network users.

Specifically, the results show that SNS users are concerned in two ways. First, they are concerned about the type of information collected. They have serious reservations about organisations' ability a) to create a complete and richly detailed profile, and b) to relate this profile to an actual person. When data is collected and combined in this way, there is little opportunity for anonymity and this makes a person feel extremely vulnerable. Secondly, SNS users are concerned about how this information will then be used. Few blog comments expressed optimism that personal user information would be used for the advantage of the SNS user. Rather, it was felt that the information would be used by commercial organisations to exploit SNS users in some way: that patterns of data use would emerge that would rob a SNS user's freedom of choice and action. Consequently, it is not sufficient for companies to publish privacy statements regarding the transparency of data use in the context of using personal information for commercial purposes. For this to be accepted by SNS users, the

businesses needs to assure and demonstrate there is some intrinsic benefit to SNS users – a benefit that may incorporate commercial interests but is able to go beyond them.

3. ***Legal Levels Concerns:*** Legal issues were not perceived as serious privacy concerns by users. This is an interesting finding because it strongly challenges existing business thinking and practice that data protection is the most effective means of reassuring users about the security of online personal information. Another reason may be that SNS users challenged organisational behaviour and practices more than they did the law as they think organisations have a social responsibility to protect their privacy in the first place. However, 15% of user responses as reflected in the Beacon and Connect blog commentaries show perceived user concern relating to data protection breaches. For SNS users, the issue is not so much safeguarding the database containing personal user information but more defending people against the misuse of such data (i.e. business integrity) as well providing more control over collected data (i.e. ability to delete data permanently). Consequently, social marketing tools built and used on the pillar of data protection and associated management theory needs to be rethought in the context of online social networks. Interestingly, the European Union (EU) has proposed new data protection laws which consist of a rule called the "right to be forgotten" – empowering online users to request deletion of their personal data (e.g. embarrassing , inaccurate or any other form of data) from the internet and company databases permanently (Warman, 2012). The author believes that such new legal frameworks would help revive user trust on the internet in general and social network sites in particular. However, at the same time they may pose a challenge for online businesses since providing such controls to online users cannot be accomplished without extra cost.

6.4 Evaluating the Privacy Leakage Taxonomy

Following the methodology outlined in Chapter three, the key findings of this research are evaluated against relevant literature to better understand the extent to which the current empirical findings fit with related theory. Solove's (2006) privacy taxonomy was found to

be the most relevant and comprehensive theoretical framework for evaluation purposes (as discussed in literature review chapter). Thus, the privacy concerns of the present taxonomy are compared and contrasted with Solove's (2006) privacy violations. There was some similarity between the research questions of the current study and those of the Solove (2006) study as both aimed to find privacy violations, though the current study was specifically designed to investigate privacy concerns of online users whereas Solove's (2006) was conducted to illuminate a general notion of privacy. Also, unlike Solove's (2006) study which only identified activities causing privacy violations, this research investigated privacy leakage concerns which emerge due to organisational practices as well as to organisational behaviour. Another important difference between both studies was the data collection protocols: the research herein was designed to understand the privacy concerns of online users through empirical data, whereas Solove's (2006) was designed to investigate privacy violations through existing literature in social sciences and law.

The methodology adopted by the author for comparison consisted of comparing each privacy violation in the Solove taxonomy with each and every privacy concern in the present taxonomy to determine if they were related or not (based on their definitions and properties). Figure 6.5 indicates where privacy concerns of the current study were echoed in Solove's (2006) analysis of privacy violations. Interestingly, the studies had notable similarities in their findings. In particular, privacy themes of data protection breaches and information leak in the present taxonomy are matched to a great extent with Solove's taxonomy. Both studies found that the information processing, collection and dissemination practices of organisations usually cause breaches related to data protection and information leak. For example, Solove (2006) identified that privacy vulnerabilities such as secondary use, identification, aggregation, and insecurity relate to information processing. Likewise, the current study indicates that users are concerned about the sharing of personal information (with third parties) without consent, sharing of identity data and the aggregation of behavioural data with personal data. Online users showed concern about identity theft which was analysed by Solove (2006) within insecurity privacy violations which result due to breach and abuse of data. Surveillance vulnerability within the broad theme of information collection in Solove's taxonomy is mapped well with the privacy concerns online tracking and stalking as Solove (2006) argues that surveillance resembles interrogation as both relate to data collection without user consent.

Online social network users were also outraged by the unauthorised involuntary tracking of their web activities.

Solove (2006) argues that disclosure of personal information sometimes damages people's reputations. The current research highlighted a similar concern, herein called embarrassing disclosure, which refers to online users' unease with the disclosure of personal information, as they believed it damaged their reputation. Nonetheless, some findings of the current study were entirely new and were not conceptualised in the Solove (2006) privacy taxonomy as such. For instance, the broad privacy theme business integrity, which was considered the most serious concern by online users, was ignored in the Solove taxonomy. This may be because in the new digital age led by social media technologies, people disclose (either voluntarily or involuntarily due to the power of tracking and aggregation technologies) rich personally identifiable information which is a tempting source of revenues for online businesses, which have been found consistently to use personal information for commercial purposes, which indeed betrays user trust. Users perceive this as especially dishonest when privacy breaches are repeated and business practices deemed underhand (deceptive).

	Privacy concerns	Privacy violations (Solove, 2006)
Business Integrity	Monetise user data	✓
	Untruthfulness and dishonesty	
	User distrust	✓
	Deceptive practices	
	Repeated privacy breaches	
	Disregard of user privacy	
	Permissive privacy controls	
	Future abuse of personal data	✓
User Control	Lack of user control	✓
	Automatic opt-in	
	No universal opt-out	
	Permanent data deletion	
	Granular control	
	Privacy Diffusion	✓
	Invasion of users' web space	✓
	Online portable identity	
	Control over user behaviour	✓
	Complex privacy controls	
Transparency	Lack of transparency	✓
	Lack of user awareness	
	Lack of user incentives	
	Uninformed user tracking outside SNS	✓
	Information ownership	✓
	Self Accountability	✓
	Changes in privacy policy	

	Privacy concerns	Privacy violations (Solove, 2006)
Data Protection Breaches	Sharing personal data without notice	✓
	Sharing personal data without consent	✓
	Information aggregation	✓
	Sharing identity data	✓
	Secondary use of data	✓
	Data security	✓
	Unlimited third party information sharing	✓
	Data retention	✓
	Legal issues	✓
	False anonymity	
Automatic Information Broadcast	Broadcast private data	
	Automatic information dissemination across the web	
	Embarrassing disclosures	✓
	Cross posting of personal information	
Information Leak	Third party information leak	✓
	Online tracking	✓
	Identity theft	✓
	Stalking	✓
	Publicise profiles on third party sites	✓
	Online archeology	✓
	Fourth party data leak	✓

Figure 6.5: Comparison of Findings with Solove's Privacy Taxonomy (2006)

6.5 Application of Taxonomy of Privacy Leakage Concerns – The Case of LinkedIn's Social Advertising Programme

The taxonomy of privacy leakage concerns of Facebook users presented earlier offers a concrete means to organisations of understanding the nature and form of privacy concerns. Such understanding may help businesses avoid breaching users' privacy and build trust which is vital for their success (Wu et al., 2012). Below is the application of the said taxonomy to a professional networking site – LinkedIn, which experienced severe user backlash on privacy grounds when it launched a social advertising programme like Facebook Beacon. Such user backlash is similar to that received by Facebook when it launched Beacon some years ago. This corroborates the current research finding that the fundamental privacy perceptions of online users do not change as users' value privacy and expect organisations to provide better user control as well as to remain transparent, fair and respectful with regard to using personal data for commercial gains. Therefore, the author believes that the developed taxonomy of privacy leakage concerns would have provided informed guidance to LinkedIn on how to address these concerns ahead of the launch of the social marketing programme so that user trust could be built.

LinkedIn – the largest professional network with over 200 million users (LinkedIn press, 2013) – launched a social advertising initiative in late June 2011 to leverage users' names, photos and network activity in third-party advertising (Ducklin, 2011; Evan, 2011; Shaughnessy, 2011). The idea was to use LinkedIn users' names and photos in third-party advertisements when they recommended a product or followed a company on third-party websites. Also, when users' performed such actions on LinkedIn their names and photos could show up in LinkedIn advertisements, as shown in Figure 6.6.

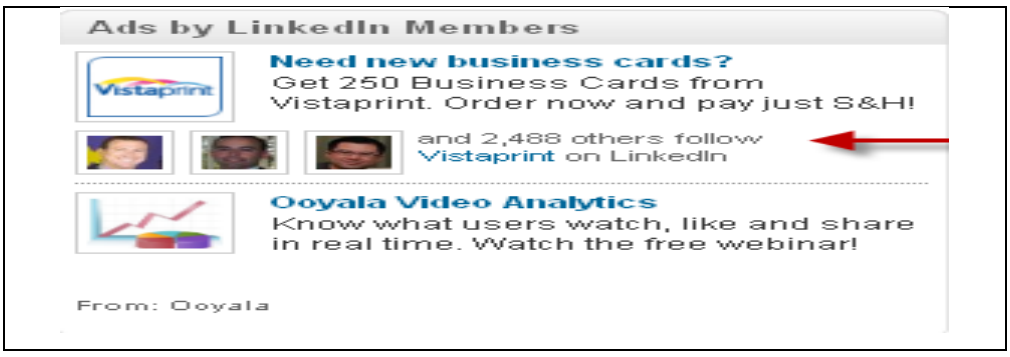


Figure 6.6: Example of LinkedIn Ads *(Source: Marketaire, 2011)*

However, LinkedIn failed to give any notice to users regarding the launch of this social advertising programme, but rather changed their privacy policy quietly without users' knowledge. Consequently, the company shared users' names, photos and network activity with third-party advertisers without explicit user consent, which provoked privacy concerns amongst SNS users. Ominously from the users' point of view, the default setting was automatically set to opt-in to share user's name and photo in social advertising (see figure 6.7). This resembled Facebook's controversial launch of Beacon in November 2007 which was also insidious in nature in the way it was designed to track users' behaviour within and outside the social network without their explicit consent.

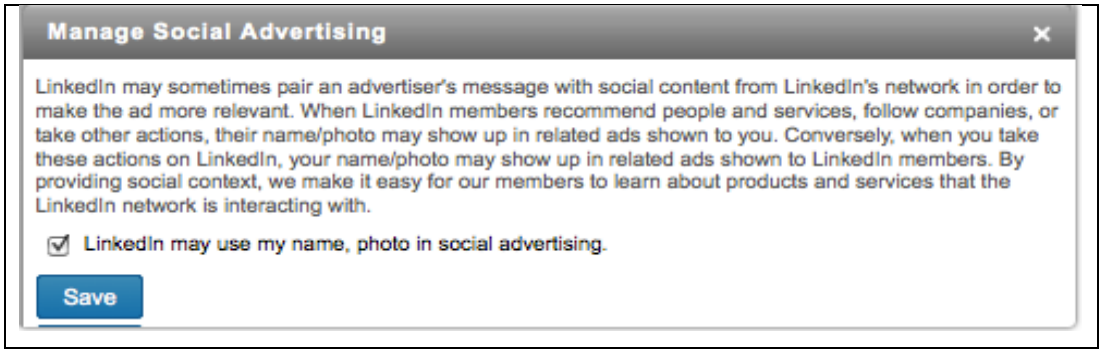


Figure 6.7: Privacy Policy of LinkedIn that Automatically opt-in Users for Social Advertising ***(Source: Ducklin, 2011)***

Surprisingly, LinkedIn repeated similar mistakes which Facebook had made some years before and users' reactions were also similar in that they did not like the company to use and share their personal data with third-parties in their advertisements. User reaction on the launch of social advertising by LinkedIn was more prominent and visible as evidenced in user blog postings. For instance, one user clearly perceived it as a privacy intrusion and remarked:

> *I can't find "Manage Social Advertising" anywhere. It doesn't even show up in their online help. They either have quietly removed this "privacy intrusion" or do not allow users to change this setting anymore, in which case I will close my account.*

Indeed, the user was helpless and outraged on this occasion because of the lack of control and the inability to turn off social advertising, believing it was a grave violation of user

privacy. LinkedIn users who were automatically opted-in had no knowledge that their names, photos and network activity data were being tracked and used in advertisements and hence they could hardly opt-out. The researcher argues that the developed taxonomy of privacy leakage concerns of Facebook users could have provided guidance to understanding the complex issue of using personal data for social marketing so as to avoid privacy violation during the launch of LinkedIn's social marketing programme. To recap, the taxonomy presented earlier highlighted lack of user control as a serious encroachment on user privacy and the fact that users value and like being given control. For instance, the taxonomy regarded automatic opt-in, lack of granular control and complex privacy controls as important elements of user control which were not taken into consideration by LinkedIn during the launch of its social marketing tool, as is evident from the aforementioned user reaction. Another user highlighted the need to have more granular control, commenting:

> *No way do I want my photo showing up in a competitor's ad!!!!!! What were they thinking?*

Here, surely, the user's concern was not that they did not want to publish their data in marketing at all, but rather the unavailability of granular control so that they could manage their actions on an individual basis. This again highlighted that control is a more nuanced and complex concept and that organisations and designers need to understand this and should provide a matrix of options/controls to users which is acceptable to them. This seems challenging for designers, but for these tools to succeed they have to move towards providing granular control to users. Hence, organisations like LinkedIn should build and implement these tools in such a manner that users feel in control of their personal information. This will also help build user trust in the integrity of online businesses. Another user questioned the integrity of LinkedIn and remarked:

> *Changes should be opt-in - the only reason these changes are opt-out is because they know full well that nobody in their right mind would *ever* opt-in to them so the only way they can deliver any people to the advertisers is by hoping users don't notice. Underhand and dishonest.*

Here, the user doubted the integrity of LinkedIn and perceived the company's practice as being dishonest. Interestingly, similar user sentiments were observed during the launch of Facebook's Beacon that was conceptualised as, for example, 'deceptive practices' and

'untruthfulness and dishonesty' in the taxonomy of privacy leakage concerns presented earlier. Indeed, the taxonomy offers guidelines to organisations to be more transparent in their information management practices to avoid such user distrust in future.

Privacy Concern	LinkedIn Actions Causing Violation of Privacy and Possible Responses to Address these Concerns
Business Integrity	**Monetise user data -**LinkedIn started monetising user data as they used users' personal information e.g. name and photo in advertisements without user consent and reward. Instead, LinkedIn could have offered some rewards to users to create a win-win situation. **Deceptive Practices-** The covert launch of social marketing and sharing of users' data with third party advertisers was challenged by users being deceptive. LinkedIn could have avoided this concern if they had notified users and obtained their consent. **User Distrust -** LinkedIn's launch of social marketing without notice to users betrayed user trust as they believed the company practice insidious. The company could improve transparency of information practices and provide choice to user to build user trust. **Future Abuse of Personal Data-** LinkedIn action to publish users' name and photograph in advertisements as well as sharing it with third party advertisers was susceptible to future abuse e.g. employers may use this against employees or third party advertisers can use it for other purposes. LinkedIn could address this concern if they disclose terms of contract with third parties to LinkedIn users to revive their trust.
User Control	**Lack of user control**- LinkedIn's launch of social marketing tool as an automatic opt-in or opt-out without any granular control to users was challenged by users as they thought the company did not provide control to manage their personal data. LinkedIn could avoid this privacy concern if they launch this tool as opt -in and provide better granular controls which enable users to choose which advertisements could use their names and pictures and which cannot. It cannot be a fit for all purpose - that once a user is automatically opt-in then LinkedIn can use users' data in all advertisements. Also, LinkedIn should provide a control that enable users to delete their information permanently so that third parties cannot use a user's name and photo indefinitely.
Transparency	**Lack of transparency-** LinkedIn launched social marketing initiative without any notice and knowledge of users and failed to create user awareness. This was perceived unfair by users due to lack to transparency. LinkedIn should make data management practices more transparent by notifying members of any change in privacy policy as well as launching awareness campaigns to educate users.
Automatic Information Broadcast	LinkedIn's action to automatically disseminate users' personal data across the web caused embarrassment to users as their names and photos were published in advertising brands. Indeed users did not want to make these actions public which was evident from negative user commentary. Therefore, LinkedIn should have provided granular control to choose which actions to share with advertisers as well as seek users' consent rather than automatically disseminating user information publicly.
Data Protection Breaches	**Sharing of personal data without notice and consent-** LinkedIn failed to provide any direct notice to users as well as obtain consent, thus violating the pillars of data protection principles. The company should have provided notice and obtained user consent to avoid breaching users' privacy. **Data Retention and secondary use**– LinkedIn's action also violated data retention and secondary use data protection guidelines because users were not given any control to delete the shared data permanently as well as their data was used for the other purposes for which it was collected. LinkedIn could avoid such data breaches if they allowed users to delete data permanently and obtained explicit consent to use for social marketing.

Table 6-2: Application of the taxonomy of privacy leakage concerns to LinkedIn case

See Table 6.2 for detailed application of the taxonomy of privacy leakage concerns in the case of the launch of LinkedIn's social marketing tool.

To conclude, the taxonomy of privacy leakage concerns proved a useful tool to evaluate LinkedIn's actions during the failed launch of its social marketing programme which had to be recalled because of user criticism on grounds of privacy. Therefore, what should have been an innovative tool and a source of revenue for LinkedIn resulted in tarnished consumer trust. What this evaluation has shown is that if organisations such as LinkedIn understand the nature and form of privacy leakage concerns they can be more articulate when launching these tools and ensure consumer privacy and trust are protected. Therefore, organisations should take the lead and implement these tools as an opt-in to provide more control and choice to users as well as make organisational practices more transparent and rewarding for users. Although, this may hit their revenues in the short run, the strategy may help meet revenue targets in the long run as well as build consumer trust.

Chapter 7: Conclusions and Future Recommendations

7.1 Chapter Overview

Overall, this chapter summarises the research conclusions and outlines recommendations for future research. Firstly, the research and its findings are recapitulated. Then, the author reflects on the findings and provides a summary of the conclusions drawn from the current research. Thereafter, the author discusses the research contributions. Finally, the research limitations and recommendations for further research are discussed.

7.2 Recapitulation of Research and Findings

To recap, the current research aims to provide a better understanding of the nature and form of concerns relating to the leakage of privacy of online social network users. A two-stage longitudinal case study was designed to direct the research. The overall research is organised into seven chapters. The author finds it appropriate to summarise this research along with its findings according to what was discussed and accomplished in each chapter.

Starting with **Chapter 1**, the author explored the main motivations for carrying out this research. An initial review of the literature revealed that information privacy has become a matter of increasing concern to consumers, businesses, policy makers, researchers and privacy activists (Smith et al., 2011). As revealed by consumer opinion polls, privacy was one of the greatest concerns for consumers (Smith et al., 2011). For instance, a consumer poll in 2008 revealed that 72% of consumers were concerned by the online tracking and behavioural profiling of their information by companies (Consumers-Union, 2008). Also, realising the potential of consumers' personal information, organisations are increasingly seeking to exploit consumer data for commercial benefits. Consequently, such tracking and commercial use of personal information has raised concerns regarding the privacy leakage of online users. This motivated the researcher to further explore the issue of privacy leakage. However,

unlike many researchers who largely depend on conventional literature sources (e.g. journal articles, published conference papers, and books), the author explored other additional sources of information, such as news media sites, blog sites, and technology sites. This was undertaken mainly to identify contemporary information privacy issues – almost in real time and especially from the point of the view of online consumers (users). The author found these social media resources useful because they provide a technological platform which encourages people participation and engagement, making it an invaluable forum for the discussion and debate of the very social issue of information privacy. Nonetheless, conventional literature sources enabled the author to have a multidisciplinary, chronological view of the issues surrounding information privacy research, as well highlighting the methodological challenges in privacy research.

Accordingly, the literature identified the complex (Solove, 2006) and fluid nature of privacy as it is highly context-dependent (Ajzen and Fishbein, 2005; Nissenbaum, 2004) and temporal in nature (Palen and Dourish, 2003). Consequently, it becomes impossible to predict what could cause a breach of privacy since privacy means different things to different people given different times and contexts. In the digital age led by social networking sites, the situation is exacerbated as people have to share personal information to use these free services, while businesses use pervasive technologies to find innovative opportunities to make use of personal data for commercial gain. This consequently leads to the leakage of privacy of online social network sites. The literature also reported such privacy leakage recently. For example, Krishnamurthy and Wills (2008, 2010) and Bonneau (2009a, 2009b) found SNSs are leaking users' information. However, the literature review identified a gap in current knowledge – no study had investigated users' perceptions of privacy leakage in SNSs. Therefore, given the importance of this contemporary issue and to fill the gap, the author was motivated to undertake the current research to provide a better understanding of the nature and form of concerns related to such privacy leakage and devise a cohesive framework which deconstructs privacy leakage by identifying what specific organisational activities and practices cause concerns amongst SNS user and why and how these cause privacy harm.

Keeping in view the aims and objectives of the current research and the methodological challenges in privacy research, the author also proposed the research approach to be adopted herein. Because of the fluid and temporal nature of privacy, and the lack of research in privacy leakage in SNSs, the author was motivated to conduct an empirical longitudinal case study of the launch of Facebook social Ads programme (Chapter 3 discussed these grounds in detail).

Chapter 2 provided the review of the literature relating to information privacy in general and privacy leakage in particular. Privacy concerns of SNS users are also discussed. Within information privacy literature, the author reviewed information privacy frameworks particularly related to information systems. Of such information privacy frameworks, concern for information privacy (CFIP) proposed by Smith et al., (1996) was the first within the information system discipline to conceptualise privacy as a multidimensional construct; earlier research, such as Westin (1967) and Stone et al., (1983), had viewed privacy as an ability to control personal information (i.e. a single attribute construct). CFIP aimed to deconstruct privacy in terms of four factors - collection, unauthorized secondary use, improper access, and errors. However, CFIP was used mostly in an offline or direct marketing context (Malhotra et al., 2004), where communication is usually one directional. Perhaps that is why Malhotra et al. (2004) devised a privacy framework called internet users information privacy concerns (IUIPC) which enabled two-way communication between companies and users. IUIPC conceptualised privacy as a combination of three factors – collection, control and awareness. However, Malhotra et al., (2004) applied this construct in the environment of e-commerce, where information disclosure is mandatory (Nov and Wattal, 2009), unlike social networking environments which are characterised by voluntary information disclosure and information broadcast principles. Consequently, Dwyer (2007a,2007b) argued that privacy within SNSs is often not expected or is undefined, with the result that it is often impossible to predict what could cause a privacy breach, because 'privacy' means different things to different people. Solove (2006) developed a taxonomy of privacy which conceptualised privacy in terms of activities causing privacy problems. Whilst, the Solove (2006) taxonomy cleared some fog around privacy in general, it was a literature-led study conducted without much focus on the online environments characterised by pervasive use of technologies and the new digital age of social

networking services designed to support widespread information dissemination (Rosenblum, 2007). Therefore, the author conducted this empirical study to investigate perceptions of SNS users related to leakage of their privacy in SNS which often occurs due to organisational activities and behaviour.

Chapter 3 elaborated the interpretive research approach adopted in this thesis to investigate the social phenomenon of privacy leakage from users' perspective. A case study research method was employed and qualitative longitudinal data was collected from user blog commentary between November 2007 and October 2010 spread across two cases i.e. the launches of Beacon and Facebook Connect. Grounded theory data analysis procedures (i.e. open and axial coding) were used to analyse user blogs comments. QSR NVIVO – a dedicated qualitative data analysis tool was used to guide the analysis.

Chapter 4 reported the finding and analysis of stage one of the longitudinal case study, i.e. the launch of Facebook Beacon, a social marketing tool aimed to manipulate social network data for business marketing. However, Beacon had to be revised as an opt-in tool soon after its launch due to huge user backlash over third-party leakage of personal data and was subsequently shut down by court order to settle a law suit of $9.5 million. The empirical findings were conceptualised as the Beacon privacy framework which proved a first step towards improving our understanding of the fluid nature and form of privacy concerns relating to the leakage of privacy of online social network site (SNS) users.

Chapter 5 presented the empirical findings and analysis of stage two of the longitudinal case study – the launch of Facebook Connect, which enabled third-party websites to leak users' actions performed on these sites back to Facebook profiles. The Connect privacy framework was devised based on the empirical findings, thus providing further insights into the nature and form of concerns of SNS users relating to the leakage of their personal information in SNSs.

In **Chapter 6**, in order to provide a consolidated and a developmental view of the findings of stages one and two of this research, both the Beacon privacy and Connect privacy frameworks were compared and then combined. This enabled the author to

devise a comprehensive privacy framework called taxonomy of privacy concerns of online users. The author hopes that the privacy taxonomy may provide a cohesive fabric for the understanding of privacy leakage in SNSs. Also, as one of the objectives of the current research was to investigate how privacy concerns emerge and subsequently evolve, the synthesised research findings helped towards accomplishment of this objective. The taxonomy grouped general privacy concerns into six categories namely: 'User Control', 'Business Integrity', 'Transparency', 'Automatic Information Broadcast', 'Data Protection Breaches', and 'Information Leak'. Thus, the taxonomy provided a concrete means to improve our understanding of the nature and form of privacy leakage of SNS users and to appreciate how and why such leakage caused harm to users' privacy.

The current **Chapter 7** began with the summary of the research and its findings. In the next section, the author while building on the empirical findings of this research provides a summary of the current research. Afterwards, research contributions are highlighted, followed by the discussion of the research limitations. Finally, recommendations for future research are presented.

7.3 Conclusions

This research focused on understanding the nature of concerns relating to privacy leakage as perceived by online social network users. The substantial empirical evidence collected through user blog commentary suggests that online social network site users were concerned about the leakage of their privacy. This research demonstrated that SNS users displayed a sophisticated understanding that it is not the capture of information that is the issue but rather how that information is leaked, combined, reused and broadcasted. Specifically, the study findings revealed that there are three different types of activities causing concerns relating to the leakage of privacy of SNS users: organisational, user and legal. The following high level privacy concerns were highlighted within these three boundary levels.

- Business Integrity
- User Control
- Transparency

- Data Protection Breaches
- Automatic Information Broadcast
- Information Leak

These privacy concerns provided insights into the nature and form of privacy leakage in social network sites. The organisational activities causing such concerns are termed the organisational boundaries. The most serious limit relates to the integrity of business organisations. Users questioned the integrity of companies to safeguard their interests because the visibility, purpose and presence of marketing tools (i.e. Beacon and Connect) was not made clear to them regarding the use and dissemination (leakage) of their browsing actions. In particular, the strong affective response to such secrecy, as evidenced from SNS user comments, indicate that the business was perceived to be acting in an underhand or dishonest manner.

The user level boundary as identified in this research suggests that user control is a more nuanced and complex concept, contrary to the findings of most existing research, which has simply shown user control as an ability to limit the use of personal information by third parties. The research findings of this study revealed that the issue of concern is not simply the absolute amount of information held by the user (or organisation) but the complex interplay between context and content. As the context of personal information is constantly changing, so the desire for user control is also constantly influx. However, available privacy controls fail to provide such level of contextualisation – thus posing a challenge, as well as an opportunity for designers and social network service providers to strike the right balance between intrusion and engagement.

The results of this research indicate that opaque information practices are considered invasive violations of trust by social network users, which consequently limits the potential to use SNS data for business marketing. Thus, in order for social network service providers to build trust and harness the potential SNS data for commercial purposes, their information processing practices need to be transparent as well as rewarding to SNS users. For instance, this research revealed that users were upset that businesses use their personal data for commercial reasons and in return offer no compensation or reward. Although one could argue that their ‘reward’ is the free use of these services, this research shows that that was not perceived enough by SNS users, as

they think the risks associated with the dissemination and use of their personal data outweigh the benefits.

Interestingly, the legal issues as perceived by online SNS users pertained to the use, reuse (e.g. associating users' web actions with real identity data) and dissemination of personal information across the web, rather than to the capture of personal data which is the cornerstone of the available data protection laws. So, for SNS users the real issue was the leakage of their private information and concerns associated with such leakage. Specifically, for online SNS users, it is not so much a matter of safeguarding the database containing personal user information but of defending people against the misuse of such data (i.e. business integrity), as well of providing more control over collected data (i.e. ability to delete data permanently). Consequently, social marketing tools built and used on the pillar of existing data protection and associated management theory needs to be rethought in the context of online social networks. A recent interesting development is that European Union (EU) has proposed new data protection laws that may take more than two years to enforce and include a rule called "right to be forgotten", empowering online users to request deletion of their personal data from the internet and company databases permanently (Warman, 2012). The author believes that such new legal frameworks would not only help revive user trust in online businesses and social network sites, but also provide better user control and accountability of information processing practices of online businesses. However, their implementation may pose a challenge for online businesses since providing such sophisticated controls to the huge online user base is not possible without extra costs. Alternatively, online businesses may need to provide better user controls and demonstrate self-accountability at their end.

7.4 Research Contributions

This thesis addressed the critical issue of information privacy with a particular focus on understanding the nature of privacy leakage in the new digital age led by social networking sites. The research contributions of this thesis covered multiple facets such as theory, practice and methodology.

7.4.1 Theoretical Contributions

Theoretical contributions derived from this thesis are multifold. The first and foremost contribution of this research is the development of a taxonomy of privacy concerns of online social network users. The developed taxonomy identified three boundary levels of information dissemination and use such as organisational (behaviour), user and legal. Such distinctions between privacy levels are vital in order to analyse privacy leakage systematically. An interesting and surprising finding and contribution derived from this thesis is that this research for first time identified a relationship between privacy concerns and organisational behaviour. For example, within the organisational behaviour boundary, the most serious concern relates to business integrity, which was questioned by online users due to a number of worries including deceptive business practices, disregard of user privacy and repeated privacy breaches. Consequently, users perceive organisational behaviour as quite insidious. Particularly interesting was the revelation that users not only questioned the integrity of the SNSs but also of third-party businesses who partnered with SNSs. This loss of user trust can be damaging for the entire online business landscape, which is usually built on consumer trust. Therefore, organisations need to open up and provide more transparency for their actions and information dissemination practices to assuage users' privacy concerns and to rebuild user trust. This, consequently, will help improve the integrity and credibility of organisations.

The user level boundary indicted that the lack of user control was perceived to be a major concern for online users. Furthermore, the analysis revealed that user control is a more nuanced and complex concept as the issue of concern is not simply the absolute amount of information held by the user (or organisation) but the complex interplay between context and content. And since the context of personal information is constantly

changing, so the desire for user control is also constantly influx. This insight indeed provides a challenge and an opportunity for organisations to consider providing a more agile, streamlined metrics able to reconcile the constantly changing nature of personal information in ways that support dissemination of personal data without infringing user privacy. An interesting finding was that users were not as concerned about the collection of personal data and protection of databases as they were about the aggregation of behavioural data with identity data (available on their SNS profiles), which was perceived to be a serious privacy concern. This strongly challenges existing business thinking and practice that data protection is the most effective means of reassuring users about the privacy of their personal information.

Another contribution derived from this thesis relates to the taxonomical representation of privacy leakage concerns of Facebook users. To the best of the author's knowledge, to date no study exists which investigates the privacy leakage concerns of Facebook users. The taxonomy provides a comprehensive framework for researchers and practitioners to better understand the nature and form of privacy leakage by offering a systematic way to identify, classify and critically analyse privacy problems of online social network users. Overall, the taxonomy consists of 46 distinct privacy concerns of online SNS users covering almost all diverse activities (or violations) which cause privacy concerns. The research problem identified in Chapter 1 related to understanding the complex, fluid and context-dependent nature of information privacy in the new digital world dominated by social networking sites. Unlike many privacy studies in IS (e.g. Malhotra et al.,2004; Smith et al.,1996) that used privacy concerns to measure general privacy problems, the current research adapted context-based perspective of privacy concerns as advocated by Margulis (2003), Solove (2006, 2008) and Xu et al. (2011). Specifically, this thesis adapted the theory of privacy concerns in the context of online social network sites and defined privacy concerns as privacy loss perceived by online social network users as a consequence of the information dissemination practices of organisations (i.e. social network sites and third-party websites). Such context-specific conceptualisation of privacy concerns proved a useful construct to unpack privacy in terms of not only identifying what activities cause privacy concern but also how and why these activities cause privacy violations for online social network users.

The taxonomical representation of privacy concerns provided a useful means of classifying and categorising the privacy concerns of online social network users. The developed taxonomy of privacy concerns has provided a better understanding of the unique characteristics of online social network data as distinct from commercial data-set. Specifically, the taxonomy may help us to understand the nature, form and the root causes of privacy leakage in online social network and to explore possible boundaries and to contemplate different avenues of use when seeking to use and disseminate online social data.

Another theoretical contribution derived from this thesis relates to the finding that fundamental privacy beliefs of online social network users do not evolve much and mostly remain consistent over time, as the analysis of user blog comments shows little change in users' privacy perceptions over a three-year period. The qualitative evidence gathered between November 2007 and December 2010 showed that users valued privacy and did not seem to allow organisations to intrude into their private-public boundaries. Rather, online users perceived privacy in the new digital age as being of equal importance to privacy in the offline world where it is respected as an established social norm. Although the severity of some privacy concerns changed, the fundamental privacy perceptions remained largely unchanged. This is an interesting contribution which challenges the common business belief that voluntary disclosure of personal information by online users gives businesses an automatic right to explore ways on how to exploit personal data for commercial purposes. Of course, both Beacon and Connect launches (innovative tools to leverage SNS data for commercial purposes) demonstrated such business thinking. However, the analysis of user reaction on both these launches (reported in Chapters 4 and 5) revealed that users challenged such business thinking, believing that the aggregation and dissemination of their personal data for commercial purposes was unauthorised and an infringement to their privacy. User blog commentary unveiled an interesting user perception that the web is (or should be) a place where their private actions (such as purchases, browsing etc.) ought to remain private unless they themselves allow otherwise. Therefore, the designers of social marketing tools and social network service providers should appreciate the socio-technical context of SNS and the need for a high level of sensitivity about core human values and concerns.

The developed taxonomy of privacy concerns and subsequent discussion and analysis of the findings enabled the author to propose fair information broadcast guidelines, another contribution. The guidelines may provide further insights for organisations, users and policy makers into how to strike the right balance between invasion of and need for privacy of online social network users.

7.4.2 Practical Contributions

The empirical evaluation, analysis and discussion of Facebook Beacon and Facebook Connect conducted in this thesis (in Chapters 4, 5 and 6) provide useful insights into practice. The critical analysis of these two tools suggest that the designers and developers of business marketing tools should appreciate the socio-technical context of social networks and the need for a high level of sensitivity to user privacy. What should have been successful innovations, however, were damaged and ultimately withdrawn (e.g. Beacon) because the limits of the use of personal data in social networks were not well understood. Therefore, specific privacy concerns identified in the developed taxonomy provide concrete means for designers and developers of such tools to understand the boundaries of disclosure and use of SNS data. The author argues that such understanding of privacy concerns is critical for designers and developers of social marketing tools as *"many IT professionals have common-sense notions about privacy that can turn out to be inaccurate"* (Iachello and Hong, 2007:p.3).

The analysis and subsequent discussion of the findings of this research show that user control is a more nuanced and complex concept and is dependent on the complex interplay between context and content. Also, because the context of personal information is constantly changing, the desire for user control is also constantly inflow. This insight indeed provides a challenge and an opportunity for designers and developers of social marketing tools to consider providing a more agile, streamlined metrics able to accommodate the constantly changing nature of personal information so as to successfully design and implement such tools. Therefore, the research insights provide informed direction to designers and developers on how to design and develop social marketing tools that strike a right balance between user control and organisational data use needs.

Another finding of this research suggests that social marketing tools (e.g. Beacon) failed to strike a correct balance between user control and user experience. Beacon was designed as a lightweight tool that would not interrupt users' web browsing, as Mark Zuckerberg remarked whilst reflecting one month after the Beacon launch (Facebook Blog, 2007), and consequently its interface was made minimal and without much consideration of providing users with choices and control for informed consent. This study shows that users were outraged because neither the visibility and purpose of Beacon was made clear to them nor were they offered more choices to control their privacy. Rather, the lightweight interface of Beacon was strongly criticised by users since they perceived the tool was working in an underhand manner. So, for online social network users, perceived ease of use of such tools (i.e. third-party data sharing) was not the issue; their priority was having more choices and controls to safeguard their privacy. This consequently challenges the classical thinking outlined in the popular technology acceptance model that perceived ease of use is the major factor determining user intention to use a system (Davis et al., 1989). This finding and contribution has implications for practice as it highlighted the critical need to incorporate the issue of user choice and control (to safeguard users' privacy) in the design of systems and interfaces rather than disposing of privacy as a matter of policy.

The author acknowledges that it is often not possible to devise a recipe for designers and practitioners on how to balance privacy requirements with better user experience. However, specific privacy concerns highlighted in the privacy taxonomy offer insights into designers and social network service providers about the nature and form of privacy leakage – which enable them to incorporate user sensitivities into the design of such technologies whilst keeping the system as easy to use as possible.

7.4.3 Methodological Contributions

This thesis provided valuable methodological insights which can be incorporated into future empirical privacy research. Contributions to methodology have been accomplished by overcoming the methodological challenges in privacy research (as discussed in chapter 3). Therefore, the author exploited the challenges in privacy research as opportunity.

- The first methodological challenge overcome by this research relates to selecting the right construct to measure the complex and fluid nature of privacy. Many research studies have used a vague concept of privacy since privacy means different things to different people at different times. However, the current research adapted privacy 'concern' as a construct to measure privacy. Many IS studies have used privacy 'concern' as a construct to measure privacy, but the majority focused on exploring the relationship between privacy concerns and behavioural variables such as information disclosure and intention to transact (Xu et al.,2011). However, the current research adapted privacy concern to measure the root cause of privacy violations as suggested by Phelps (2000). Also, in contrast to most IS privacy studies which used pre-defined variables (deductive approach) and were aimed at theory testing, the current research took the inductive approach because the aim was to understand the root causes of privacy problems as experienced by online social network users. Therefore, building a new theory (i.e. taxonomy of privacy concerns) based on empirical data was very promising. However, the construct of privacy concern was used as a sensitising device to guide the investigation as suggested by Gregor (2006).

- Another methodological challenge in privacy research is the context-specific nature of privacy (Margulis, 2003; Solove, 2006; Xu et al., 2011) and attitudes expressed outside a specific context are not good predictors of behaviour (Ajzen and Fishbein, 2005). This posed a challenge as well as an opportunity for the author on how to study the context-dependent nature of privacy. Whilst a majority of IS studies adopted the general notion of privacy concerns (e.g. Malhotra et al., 2004 and Smith et al., 1996), this thesis conceptualised privacy as a context-based concern as advocated by Margulis (2003), Solove (2006, 2008) and Xu et al., (2011). The selection of context-specific environments and case(s) where personal information of users were exploited for business purposes is another contribution this thesis has made. Specifically, the selection of the contemporary cases of the Beacon and Connect launches by Facebook is an important contribution as user reactions were collected almost in real time for the first time ever in any empirical study investigating privacy concerns of online social network users.

- The third contribution to methodology relates to understanding the temporal nature of privacy (Palen and Dourish, 2003). The current research collected qualitative longitudinal data to understand not only what the specific privacy concerns are but also how privacy concerns evolve overtime. Accordingly, three years' worth of evidence was collected relating to the launch of Beacon and Connect which provided insights into how and if privacy concerns evolve over time – a useful methodological guideline for future privacy researchers.

- Lastly, is the contribution derived from the method of data collection employed in the current research. User blog opinions were collected almost in real time soon after the launch of social marketing tools (i.e. Beacon and Connect) to direct the investigation. To the best of author's knowledge, no study has used user blog commentary to investigate the privacy perceptions of online social network users. Compared with traditional data collection methods, such as interviews and focus groups, blogs data provided many benefits as the data was collected in real-time, in an inexpensive and ready-made manner (already transcribed) as well as the data collection was free from bias by the research process (Jones and Alony,2008). Also blogs data is well suited to measure social trend overtime (Hookway, 2008). More importantly, data collection through blogs opinions enabled the author to discover the privacy perceptions of those concerned, unlike the use of conventional data collection methods which cannot guarantee collecting data from privacy-concerned people. Obviously, the author employed various checks to ensure the quality and validity of data such as: the careful selection of blogs sites (e.g. popular news and technology blogs sites such as BBC, New York and Techcrunch), the use of a social news recommendation system (e.g. Digg), and the use of RSS web feeds to receive automatic feeds through RSS reader (e.g. Google Reader). Thus, the data collection method and protocols adopted in this research may guide researchers in devising their research methodology.

7.5 Research Limitations

Like all research studies, this research also has some limitations. One limitation is that the findings of the current research cannot be generalised as it followed a qualitative case study research approach to investigate privacy leakage concerns of Facebook users. In

order to ensure that the research findings are not idiosyncratic, the researcher selected a failure case (i.e. Beacon) as well as a neutral case (i.e. Connect) and evaluated the developed taxonomy of privacy leakage concerns with a more generic taxonomy of privacy (Solove, 2006). Since the evaluation of the present taxonomy with Solove's (2006) taxonomy reported similarities as well as differences, the author suggests the need for further empirical studies to generalise the findings of current research. Also, as the findings of the current research are context-dependent, they are more appropriate to application within Facebook. Although, the taxonomy has been applied to inform the actions of LinkedIn when they launched a social advertising programme similar to Facebook, yet it is not appropriate to conclude that these findings can be applied to the wider context of all social network sites as well as general internet users.

Furthermore, data triangulation is considered a useful approach to study privacy (Garde-Perik, 2009) because privacy is a complex concept and a variety of methods while complementing each other could overcome the weaknesses one of the other. However, the current research could not benefit from such triangulation of data collection methods, because one of the objectives of this research was to find out how privacy concerns evolve over time, it was not possible to adopt other methods such as interviews; collecting longitudinal data through interviews would have been very problematic since recruiting the same participants over a long period of time (i.e. three years) would have been difficult, not to say unfeasible given the available resources (i.e. time, funding). Also, blog data offered many benefits (e.g. measuring a social trend overtime) which outweighed the advantages of triangulating data in the current research.

Another limitation and challenge faced in this research relates to accommodating demographic and cultural factors: the demographic make-up of online users who commented on blogs could not be collected since most bloggers are anonymous. Similarly, only comments published in the English language were included in the final data set because of language barriers. However, some studies suggest that gender can influence privacy perceptions because women were found to be more concerned about their privacy than men (Sheehan, 1999). Also, research suggests that Italian society shows a low level of privacy concerns as they have a different concept of privacy (Dinev et al., 2006). But since many languages (e.g. Russian, French, Italian) use the English word

'privacy', not having their own, interpretive studies will be overwhelmingly complex because of the limited ability to collect interpretive data (Smith et al., 2011).

7.6 Recommendations for Future Research

This thesis derived findings and contributions related to theory, practice and methodology as well as identifying limitations of this research. Some of these findings, contributions and limitations pose challenging opportunities to future researchers. The finding that organisational behaviour is related to the privacy leakage concerns of Facebook users, offers an opportunity for researchers to further investigate this phenomenon from an organisational perspective. The resulting challenge is an opportunity for developers of business marketing tools to create more agile, streamlined metrics able to accommodate the constantly changing nature of personal information without infringing user sensibilities. As the current research examined perceptions of privacy leakage of online social network users without regard to their gender and culture (nationality), as cited as a limitation, further research is needed to evaluate whether gender and nationalities have an impact on the nature and form of privacy concerns as highlighted by this research.

References

Ackerman, M. S., Cranor, L. F., & Reagle, J. (1999). Privacy in e-commerce: Examining user scenarios and privacy preferences. In Proceedings of the 1st ACM conference on electronic commerce EC '99 (pp. 1-8). New York: ACM Press.

Acquisti, A., & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In P. Golle & G. Danezis (Eds.), Proceedings of 6th Workshop on Privacy Enhancing Technologies (pp. 36–58). Cambridge, UK. .

Acquisti, A. and Gross, R. (2009). Predicting Social Security numbers from public data. Proceedings of the National Academy of Sciences,106 (27),10975-10980

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), The handbook of attitudes (pp. 173-221). Mahwah, NJ: Erlbaum.

Alexa. Top Sites, (2012). Accessed on Oct 18, 2012 from: http://www.alexa.com/topsites

Altman, I. (1975). The Environment and Social Behaviour: Privacy, Personal Space, Territory and Crowding. Monterey, CA: Brooks/Cole Pub. Co., Inc.

Agarwal, S. and Mital, M.(2009). An Exploratory Study of Indian University Students' Use of Social Networking Web Sites: Implications for the Workplace. Business Communication Quarterly 2009 72: 105-110

Angst, C. M., and Agarwal, R. (2009). Adoption of Electronic Health Records in The Presence Of Privacy Concerns: The Elaboration Likelihood Model and Individual Persuasion, MIS Quarterly (33:2), pp. 339-370.

Angwin, J. (2010). The Web's New Gold Mine: Your Secrets. Available at: http://www.agriculturedefensecoalition.org/sites/default/files/file/constitution_1/1G_2010_The_WEB_New_Goldmine_Your_Personal_Information_The_Wall_Street_Journal_July_30_2010.pdf. Retrieved on 20 Aug 2010.

Arrington, M. (2009). Friendster Valued At Just $26.4 Million In Sale. Accessed on Dec 20, 2009 from: http://www.techcrunch.com/2009/12/15/friendster-valued-at-just-26-4-million-in-sale/

Avison, D. and Pries-Heje, J. (2005) Research in Information Systems: A handbook for research supervisors and their students, Elsevier Ltd, Oxford.

Awad, N. F., and Krishnan, M. S. (2006). The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization, MIS Quarterly (30:1), pp. 13-28.

Barnes, S. (2006). A privacy paradox: Social networking in the United States. First Monday,11(9). Accessed on September 12, 2008 at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1394/1312

Battelle, J. (2010). The database of intentions is far larger than I thought. Accessed on 10 October 2011. http://battellemedia.com/archives/2010/03/the_database_of_intentions_is_far_larger_than_i_thought.php

Benbasat, I., Goldstein, K. D. and Mead, M. (1987).The Case Research Strategy in Studies of Information Systems', *MIS Quarterly*, **11**(3): 369-386.

Benbasat, I. and Weber, R. (1996). Research Commentary: Rethinking Diversity in Information Systems Research," Information Systems Research (7), December 1996, pp. 389-399.

Belluck, P.(2006).Young People's Web Postings Worry Summer Camp Directors. The New York Times, 22 June 2006; Accessed on Dec 21 , 2008 from : www.nytimes.com/2006/06/22/technology/22camp.html.

Beynon-Davies, P. (1994). Information Management in the British National Health Service: The Pragmatics of Strategic Data Planning. International Journal of Information Management, 14, pp. 84-94.

Bonneau, J., Anderson, J. and Danezis, G., (2009a).Prying Data Out of a Social Network. International Conference on Advances in Social Network Analysis and Mining, 2009.

Bonneau, J., Anderson, J., Anderson, R. and Stajano, F., (2009b).Eight Friends are Enough: Social Graph Approximation via Public Listings. In proceedings of the Second ACM Workshop on Social Network Systems

Bortree, D. S. (2005). Presentation of self on the Web: An ethnographic study of teenage girls' weblogs. Education, Communication, and Information, 5(1), 25 - 39.

boyd, d. (2004). Friendster and publicly articulated social networks. Proceedings of ACM Conference on Human Factors in Computing Systems (pp. 1279-1282). New York: ACM Press.

boyd, d. (2006). Friendster lost steam. Is MySpace just a fad? Apophenia Blog. Accessed Nov 25, 2009 from: http://www.danah.org/papers/FriendsterMySpaceEssay.html

boyd, d. (2007). Social Network Sites: Public, Private, or What?.

boyd, d., and Ellison, N. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1), article 11. Retrieved June 08, 2008 from: http://jcmc.indiana.edu/vol13/issue1/Boyd.ellison.html

boyd, d. (2008). Face book's privacy train wreck: Exposure, invasion, and social convergence. Convergence, 14(1), 13-20.

boyd, D. and Hargittai, E. (2010). Facebook Privacy Settings: Who Cares?. First Monday, 15 (8).

boyd, D. and Marwick, Alice. (2011). Social Privacy in Networked Publics: Teens' Attitudes, Practices, and Strategies. A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society.

Boyle, M., and Greenberg, S. (2005). The language of privacy: Learning from video media space analysis and design. ACM Transactions, 12 (2), 328-370.

Bringer, J.D., Johnston, L.H., and Brackenridge, C.H. (2006). Using Computer-Assisted Qualitative Data Analysis Software to Develop a Grounded Theory Project. Field Methods 18: 245-266, doi:10.1177/1525822X06287602.

Bulgurcu, B., Cavusoglu, H. and Benbasat, I. (2010). Understanding emergence and outcomes of information privacy concerns: a case of Facebook. ICIS 2010 Proceedings. Paper 230.

Bussen, W. and Myers, M.D. (1997). Executive Information Systems Failure: A New Zealand Case Study. Journal of Information Technology Vol. 12, No.2, pp. 145-153.

Campbell, A. J. (1997). Relationship marketing in consumer markets: A comparison of managerial and consumer attitudes about information privacy. J. Direct Marketing 11(3) 44–57.

Campbell, J., Sherman, R.C., Kraan, E., & Birchmeier, Z.(2001). Internet Privacy Awareness and Concerns among College Students. Paper presented to APS, Toronto. June 2001. http://www.users.muohio.edu/shermarc/ aps01.htm.

Catanese, S., De Meo, P., Ferrara, E., Fiumara, G., and Provetti, A. (2011). Crawling Facebook for Social Network Analysis Purposes. Proceedings of the International Conference on Web Intelligence, Mining and Semantics. (pp. 52:1-8). ACM.

Cassidy, J. (2006). Me media. The New Yorker (May 15), 50–59.

Cavaye, A.L.M. (1996). Case study research: a multi-faceted research approach for IS. Information Systems Journal (6:3), pp. 227-242.

Cavaye, A.L.M. & Cragg, P.B. (1995). Factors contributing to the success of customer oriented interorganisational systems. Journal of Strategic Information Systems, (4:1), pp. 13-30.

Chellappa, R. K., and Sin, R. (2005). Personalization Versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. Information Technology and Management (6:2), pp.181-202.

Chua, W.F. (1986). Radical Developments in Accounting Thought. The Accounting Review (61), 1986, pp. 601-632.

Consumers-Union. (2008). Consumer Reports Poll: Americans Extremely Concerned About Internet Privacy, September 25 (http://www.consumersunion.org/pub/core_telecom_and_utilities/006189.html). Accessed on 25 Jan 2011.

Corbin, J. and Strauss, A. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. Qualitative Sociology, Vol. 13, No.1, 1990

Creswell, J. (1998). Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Saddle River, NJ: Prentice Hall

Creswell, J.W. (2002). Research Design: Qualitative, Quantitative, and Mixed Method Approaches, 2nd edition, London: Sage.

Culnan, M.J.(1993). How Did They Get My Name? An Exploratory Investigation of Consumer Attitudes toward Secondary Information Use. MIS Quarterly, 17, 3 (1993), 341-364.

Culnan, M.J. (2011). Accountability as the Basis for Regulating Privacy: Can Information Security Regulations Inform Privacy Policy?. Privacy Law Scholars Conference, Berkeley, CA., June 2-3, 2011.

Debatin, B., Lovejoy, J.P., Horn, A.-K., and Hughes, B.N.(2009). Facebook and Online Privacy: Attitudes, Behaviours, and Unintended Consequences. Journal of Computer-Mediated Communication,15 (1),83 – 108

Dinev, T., and Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. Information Systems Research,17, 1 (2006), 61-80.

Donath, J. (2007). Signals in social supernets. Journal of Computer-Mediated Communication, 13(1), article 12. http://jcmc.indiana.edu/vol13/issue1/donath.html

Donath, J. and boyd, D. (2004). Public displays of connection. BT Technology Journal, 2, 4, pp.71-82.

Dube, R. and Adomaitis, M., B., P.,(2009). Characteristics of Social Networks. Accessed on Dec 04, 2009 from: http://socialnetworking.lovetoknow.com/Characteristics_of_Social_Networks. (Their work is based on the work of Mislove, A., Marcon, M., Gummadi, K., P., Druschel, P., and Bhattacharjee, B.,(2007). Measurement and Analysis of Online Social Networks. In the proceedings of IMC'07, October 24-26, 2007, San Diego, California, USA. Retrieved on Dec 04, 2009 from: http://www.imconf.net/imc-2007/papers/imc170.pdf).

Ducklin, P. (2011). LinkedIn 'does a Facebook' - your name and photo used in ads by default. Accessed on January 03,2013 from: http://nakedsecurity.sophos.com/2011/08/11/linkedin-copies-facebook-does-a-privacy-bait-and-switch/

Duffy, Peter and Bruns, Axel (2006). The Use of Blogs, Wikis and RSS in Education: A Conversation of Possibilities. In Proceedings Online Learning and Teaching Conference 2006, pages pp. 31-38, Brisbane.

Dutton, J. E., & Dukerich, J. M.(1991). Keeping an eye on the mirror: The role of image and identity in organizational adaptation. Academy of Management Journal, 34: 517–554.

Dwyer, C. (2007a). Digital Relationships in the 'MySpace' Generation: Results from a Qualitative Study. 40th Hawaii International Conference on System Sciences (HICSS). Waikoloa, HI

Dwyer, C. (2007b). Behavioural Targeting: A Case Study of Consumer Tracking on Levis.com. Proceedings of the Fifteenth Americas Conference on Information Systems.

Eisenhardt, K. M. (1989). Building theories from case study research. Academy of Management Review, 14: 532–550.

Eisenhardt, K. M. (1991). Better stories and better constructs: The case for rigor and comparative logic. Academy of Management Review, 16: 620–627.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. Academy of management journal, 50(1), 25.

Ellison, N., Heino, R. and Gibbs, J. (2006). Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. Journal of Computer Mediated Communication, 11, 2.

Ellison, N.B., Steinfield, C. and Lampe, C. (2007). The Benefits of Facebook ‘‘Friends’’ Social capital and college students’ use of online social network sites, Journal of Computer-Mediated Communication 12(4).

Evan, J. (2011). How to manage your LinkedIn Social Advertising privacy. Accessed 03 Jan 2013 from: http://www.julianevansblog.com/2011/08/how-to-manage-your-linkedin-social-advertising-privacy.html

Etzioni, A.(1999). The limits of privacy. New York: Basic Books.

EU Directive (95). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Accessed on Oct 20, 2009 from: http://www.cdt.org/privacy/eudirective/EU_Directive_.html#HD_NM_28

Facebook (2007). http://www.facebook.com/business/?beacon. Retrieved on April 28, 2008.

Facebook Blog (2008). Facebook Across the Web. Retrieved 10 Dec 2008 http://blog.facebook.com/blog.php?post=41735647130.

Facebook (2009). Facebook Advertising. Accessed on Oct 21, 2009 from: http://www.facebook.com/advertising/

Facebook (2009a). An Open Letter from Facebook Founder Mark Zuckerberg Accessed on Dec 04 , 2009 from : http://blog.facebook.com/blog.php?post=190423927130

Facebook (2010). Privacy Policy. Accessed on 15 Jan,2010 from : http://www.facebook.com/policy.php

Facebook (2010a) Facebook statistics. Accessed on Feb 20, 2010 from : http://www.facebook.com/press/info.php?statistics

Facebook (2013) Newsroom Retrieved 27 May 2013http://newsroom.fb.com/Key-Facts.

Felt, A. and Evans, D.(2008).Privacy Protection for Social Networking Platforms. Workshop on Web 2.0 Security and Privacy,2008.

Festa, P. (2003). Investors snub Friendster in patent grab. CNet News. Accessed Dec 05, 2009 from: http://news.com.com/2100-1032_3-5106136.html

Fogel , J. and Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. Computers in Human Behaviour, 25, 153-160.

FTC (2012). Fair Information Practice Principles. Retrieved in Nov ,2012 from : http://ftc.gov/reports/privacy3/fairinfo.shtm .

Galliers, R. (1992). Information Systems Research: Issues, Methods and Practical Guidelines. Blackwell, Oxford, U.K.

Garde-Perik, V.D. E., Markopoulos, P., Ruyter,B.D., Eggen, B. and Ijsselsteijn, W. (2008) Investigating Privacy Attitudes and Behavior in Relation to Personalization. Social Science Computer Review, Volume 26 Number 1, p.20-43.

Garde-Perik, V.D.E. (2009). Ambient Intelligence & Personalization: People's Perspectives on Information Privacy (PhD Thesis).

Gomez, J., Pinnick, T., and Soltani, A. (2009). Know Privacy: The Current State of Web Privacy, Data Collection, and Information Sharing, School of Information, University of California Berkeley (http://www.knowprivacy.org/).

Goldberg, S. (2007). Analysis: Friendster is doing just fine. Digital Media Wire. Accessed Dec 05, 2009 from: http://www.dmwmedia.com/news/2007/05/14/analysis-friendster-is-doing-just-fine.

Govani, T. and Pashley, H. (2005). Student awareness of the privacy implications when using Facebook. Paper presented at the Privacy Poster Fair at Carnegie Mellon University School of Library and Information Science. Retrieved on Aug 10, 2009 from: DOI=http://lorrie.cranor.org/courses/fa05/tubzhlp.pdf

Grinter, R., and Palen, L. (2002). "Instant Messaging in Teen Life," in Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, New Orleans, Louisiana, pp. 21-30.

Gross, R., & Acquisti, A. (2005).Information revelation and privacy in online social networks. Proceedings of WPES'05 (pp. 71–80).

Guba, E. G. and Lincoln, Y. S. (1994). Competing paradigms in qualitative research, In N. K. Denzin and Y S. Lincoln (Eds.), Handbook of qualitative research. Thousand Oaks, CA: Sage.

Gurses ,S., Rizk, R., and Gunther,O.,(2008) "Privacy Design in Online Social Networks: Learning from Privacy Breaches and Community Feedback", in proceedings of International conference on Information systems (ICIS 2008), Paris.

Harris Interactive. (2002). Privacy on and off the Internet: What consumers want. Hackensack, NJ: Author.

HIPPA (1996).The Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy Rule. Accessed on Oct 20, 2009 from: http://www.hhs.gov/ocr/privacy/

Hughes, J and Jones, S. (2003). Reflections on the Use of Grounded Theory in Interpretive Information Systems Research. ECIS 2003 Proceedings. Paper 62.

Hui, K.L., Teo, H.H., Lee, S.Y.T.(2007). The Value of Privacy Assurance: An Exploratory Field Experiment. MIS Quarterly 31, 19–33 (2007)

Iachello, G. and Hong, J. (2007). End user Privacy in Human Computer Interaction. Foundations and trends in Human-Computer Interaction Vol 1 , No 1 , pp1-137

Jamal, A. and Cole, M. (2009) Rethinking Privacy in Social Networks: A Case Study of Beacon. Proceedings of 4th Mediterranean Conference on Information Systems, Athens, Greece.

Johnson, C. Y. (2009). Project 'Gaydar'. Accessed on 10 jan,2010 from : http://www.boston.com/bostonglobe/ideas/articles/2009/09/20/project_gaydar_an_mit_experiment_raises_new_questions_about_online_privacy/

Johnston, S.J. (2007). Microsoft Buys Facebook Stake. http://www.internetnews.com/busnews/article.php/3707121. Retrieved on June 10, 2008.

Joinson, A.N. (2008). Looking at, Looking up or Keeping up with People? Motives and Use of Facebook. The proceedings of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 2008, pp. 1027-1036.

Jones, M. and Alony, I.(2008). Blogs - the new source of data analysis. Journal of Issues in Informing Science and Information Technology, vol. 5, 2008, 433-446.

Jones, H., and Soltren, J. H. (2005). Facebook: Threats to privacy. December 14, 2005. Retrieved Nov 20, 2008, from http://www-swiss.ai.mit.edu/6805/student-papers/fall05-papers/facebook.pdf

Kaplan, B. and Maxwell, J.A.(1994). Qualitative Research Methods for Evaluating Computer Information Systems, in Evaluating Health Care Information Systems: Methods and Applications, J.G. Anderson, C.E. Aydin and S.J. Jay (eds.), Sage, Thousand Oaks, CA, pp. 45-68.

Kirkpatrick, D. 2007. As Facebook takes off, MySpace strikes back. www.bauer.uh.edu/cox/wordDocs/AsFacebookTakesOff.doc . (Accessed Aug 2010).

Klingsheim, A. N. and Hole, K. J. (2008). Identity Theft: Much too Easy? A Study of Online Systems in Norway. In Proc. 12th International Conference on Financial Cryptography and Data Security (FC08), number 5143 in LNCS, pp. 192–196, Cozumel, Mexico, January 28–31 2008.

Klein, H., and Myers, M. (1999). A Set of Principles for Conducting and Evaluating Interpretive Field Studies. MIS Quarterly (23:1), 1999, pp. 67-93.

Korzaan, M., Brooks, N., and Greer, T. (2009). Demystifying Personality and Privacy: An Empirical Investigation into Antecedents of Concerns for Information Privacy. Journal of Behavioral Studies in Business (1), pp. 1-17.

Krishnamurthy, B. and Craig E. Wills.(2006). Generating a privacy footprint on the Internet. In

Proceedings of IMC, October 2006.

Krishnamurthy, B., Malandrino, D. and Wills,. E.C. (2007). Measuring privacy loss and the impact of privacy protection in web browsing. In Proceedings of the Symposium on Usable Privacy and Security, pages 52{63, Pittsburgh, PA USA, July 2007.

Krishnamurthy, B. and Wills, E., C. (2008). Characterizing privacy in online social networks. Proceedings of the first workshop on online social networks, Seattle, WA, USA.

Krishnamurthy, B. and Wills, E.C. (2010). On the Leakage of Personally Identifiable Information via Online Social Networks. SIGCOMM Computer Communication Review , Volume 40 Issue 1.

Kuzel. (1992). Sampling in quantative inquiry. In: Crabtree BF, Miller WL, eds. Doing qualitative research. California: Sage, 1992; 31–44.

Land, F. (1992). The Information Systems Domain. In R.D. Galliers (Ed.). Information Systems Research: Issues, Methods and Practical Guidelines. Blackwell Scientific Publications. ISBN: 0632028645

Lam, I.F., Chen, K.T., and Chen ,L.J. (2008). Involuntary Information Leakage in Social Network Services. IWSEC 2008, LNCS 5312, pp. 167–183.

Lee, A.S. (1994). Electronic Mail as a Medium for Rich Communication: An Empirical Investigation Using Hermeneutic Interpretation. MIS Quarterly (18:2), pp. 143-157.

Lee, M.,(2009). Friendster to be sold by month's end: Reuters. Accessed on Dec 04,2009 from : http://www.reuters.com/article/idUSTRE5B25X020091204

Leedy, P.D. and Ormrod, J.E. (2005). Practical Research: Planning and Design. 8th ed., Pearson Merrill Prentice-Hall.

Lenhart, A., and Madden, M.(2007). Teens, Privacy & Online Social Networks. Retrieved on Oct 12,2009 from: http://www.pewinternet.org/

Lewis, K., Kaufman, J., & Christakis, N. (2008). "The taste for privacy: An analysis of college student privacy settings in an online social network". Journal of Computer-Mediated Communication 14 (1): 79–100. doi:10.1111/j.1083-6101.2008.01432.x.

Lenhart, A. (2009). The Democratization of Online Social Networks. PEW INTERNET & AMERICAN LIFE PROJECT. Accessed on Oct 20,2009. http://www.pewinternet.org/Presentations/2009/41--The-Democratization-of-Online-Social-Networks.aspx

Lewis, C. C., and George, J. F.(2008). Cross-cultural deception in social networking sites and face-to-face communication. *Computers in Human Behaviour*. Li, C., and Bernoff, J. (2008). Groundswell: Winning In A World Transformed By Social Technologies", Harvard Business Press, Boston, MA, 2008.

LinkedIn Press (2013). LinkedIn reaches 200 million members worldwide. http://press.linkedin.com/News-Releases/165/LinkedIn-reaches-200-million-members-worldwide. Accessed 10 Jan 2013.

Livari, J., R. Hirschheim, H. Klein.(1998). A paradigmatic analysis contrasting information systems development approaches and methodologies. Information Systems Research. 9(2) 164–193.

Livari, J., R. Hirschheim, H. Klein.(1999). Making sense of the methodology jungle: A three-tiered framework for information systems development. Working Paper.

Lwin, M., Wirtz, J., and Williams, J. D.(2007). Consumer online privacy concerns and responses: a power-responsibility equilibrium perspective. Journal of the Academy of Marketing Science (35:4), pp. 572-585.

Malhotra, N. K., Kim, S. S., and Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. Information Systems Research (15:4), pp. 336-355.

Martin, K.E. (2010). Business Roundtable · Institute for Corporate Ethics. Case BRI-1 006 (A). Facebook(A):Beacon and Privacy. Kirsten E. Martin

Martin, P. Y., & Turner, B. A. (1986). Grounded theory and organizational research. The Journal of Applied Behavioral Science, 22(2), 141-157.

Martin, David, Hailin Wu, and Adil A. (2003). Hidden Surveillance by Web Sites: Web Bugs in Contemporary Use. Communications of the ACM, 46 (December), 258–64.

Margulis, S. T. (2003). On the Status and Contribution of Westin's and Altman's Theories of Privacy. Journal of Social Issues (59:2), pp. 411-429.

Marketaire (2011). LinkedIn Ads Get Social Network Enhancement. Accessed 04 Jan 2013 from: http://marketaire.com/2011/06/27/linkedin-ads-get-social-network-enhancement/

Matavire, R. and Brown, I. (2008). Investigating the use of "Grounded Theory" in information systems research. In Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology (SAICSIT '08). ACM, New York, NY, USA, 139-147. DOI=10.1145/1456659.1456676 http://doi.acm.org/10.1145/1456659.1456676

McCallister,E, Grance, T. and Scanfone, K.(2009). Guide to protecting the confidentiality of personally identifiable information (PII) (draft), January 2009. NIST Special Publication 800-122. http://csrc.nist.gov/publications/drafts/800-122/Draft-SP800-122.pdf.

Michael, J. (1994). Privacy and Human Rights .UNESCO Publishing. ISBN 1-85521-381-8.

Miles, M. and Huberman, A. (1994). Qualitative Data Analysis: An Expanded Sourcebook. Sage, Newbury Park, California.

Mingers, J. (2001). Combining IS Research Methods: Towards a Pluralist Methodology. Information Systems Research, Vol. 12, No. 3, pp. 240–259

Moore, S. and Read, I.(2006). Collective organisation in small- and medium-sized enterprises – an application of mobilisation theory. Human Resource Management Journal, **16**(4): 357–375.

Moran, D.(2008). Announcing Facebook Connect. Retrieved 08 Dec 2008. http://developers.facebook.com/blog/post/108/

Morrissey, B. (2009). Connect the thoughts. The analysis of social networking data promises to take behavioural targeting to a new level. But will it work? Brandweek; Jun 29, 2009; 50, 26; ABI/INFORM Research pg. AM10.

Ng, K. and Hase, S. (2008). Grounded Suggestions for Doing a Grounded Theory Business Research. The Electronic Journal of Business Research Methods Volume 6 Issue 2 2008, pp. 155 - 170, available online at www.ejbrm.com

Nov, O., and Wattal, S., (2009). Social Computing Privacy Concerns: Antecedents & Effects. CHI 2009 ~ Privacy and Trust. April 6th, 2009 ~ Boston, MA, USA

Narayanan, A. and Shmatikov, V. (2009) De-Anonymizing Social Networks. In Proc. of 30th IEEE Symposium on Security and Privacy, Oakland, CA, May 2009, pp. 173-187. IEEE Computer Society, 2009.

Nissenbaum H. (2004). Privacy as contextual integrity. Washington Law Review 79:119-158.

Ohm, P. (2009). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. University of Colorado Law Legal Studies Research Paper No. 09-12. Available at: SSRN: http://ssrn.com/abstract=1450006.

Orlikowski, W.J. and Baroudi, J.J.(1991). Studying Information Technology in Organizations: Research Approaches and Assumptions. Information Systems Research (2) , pp. 1-28.

Okazaki, S., Li, H., and Hirose, M. (2009). Consumer Privacy Concerns and Preference for Degree of Regulatory Control. Journal of Advertising (38:4), pp. 63-77.

Palen, L. and Dourish,P. (2003). Unpacking "privacy" for a networked world. In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '03). ACM, New York, NY, USA, 129-136. DOI=10.1145/642611.642635 http://doi.acm.org/10.1145/642611.642635

Patil, S., Kobsa, A. (2004). Instant Messaging and Privacy. In: Proceedings of HCI 2004. pp 85{88, http://www.ics.uci.edu/~kobsa/papers/2004-HCIkobsa.pdf .

Patton, M.Q.(1990). Qualitative evaluation and research methods, second edition. Sage.

Perez, J.C. (2009). Facebook Will Shut Down Beacon to Settle Lawsuit. Accessed 02 Oct 2009 http://www.pcworld.com/article/172272/facebook_will_shut_down_beacon_to_settle_lawsuit.html

Phelps, J., Nowak, G., and Ferrell, E. (2000). Privacy Concerns and Consumer Willingness to Provide Personal Information. Journal of Public Policy and Marketing (19:1), pp. 27-41.

Preibusch, S., Hoser, B., Gurses, S., & Berendt, B. (2007). Ubiquitous social Networks — opportunities and challenges for privacy-aware user modeling. Proceedings of Workshop on Data Mining for User Modeling. Corfu, Greece. Retrieved November 12, 2008 Http://vasarely.wiwi.hu-berlin.de/DM.UM07/Proceedings/05-Preibusch.pdf

Raynes–Goldie, K.(2010) Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook. *First Monday*, volume 15, number 1, athttp://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2775/2432, accessed December 2010.

Robey, D. (1996). Diversity in information systems research: Threat, promise and responsibility. Information Systems Research, 7(4), pp. 400–408.

Rosenblum, D. (2007). What Anyone Can Know: The Privacy Risks of Social Networking Sites. IEEE Security and Privacy, vol. 5, no. 3, pp. 40-49.

Rotenberg, M. (2000). Protecting Human Dignity in the Digital Age. In proceedings of the third United Nations Educational, Scientific and Cultural Organization Congress on ethical, Legal and Societal Challenges of Cyberspace. Retrieved on July 14, 2009 from: http://webworld.unesco.org/infoethics2000/report_151100.html.

Schmidt, J. (2007). Blogging practices: An analytical framework. Journal of Computer-Mediated Communication, (12), 1409–1427.

Schoen, S. (2009). What Information is "Personally Identifiable ? Electronic Privacy Foundation. Accessed on Oct 10, 2009 from: http://www.eff.org/deeplinks/2009/09/what-information-personally-identifiable.

Schonfeld, E. (2007). Is Beacon Inflating Facebook's Visitor Numbers? Accessed on dec 10, 2008. http://techcrunch.com/2007/12/06/is-beacon-inflating-facebooks-visitor-numbers/

Shaughnessy, H. (2011). LinkedIn Social Advertising - Privacy Blunder or Missed Revenue Share Opportunity?. Accessed Jan 03, 2013 from: http://www.forbes.com/sites/haydnshaughnessy/2011/08/09/linkedin-social-advertising-privacy-blunder-or-missed-revenue-share-opportunity/

Sheehan, K.B. (2005). In Poor Health: An Assessment of Privacy Policies at Direct-to-Consumer Web Sites. Journal of Public Policy & Marketing, 24 (fall), 273–83.

Shirky, C. (2003). People on page: YASNS… Corante's Many-to-Many. Accessed Dec 05, 2009 from : http://many.corante.com/archives/2003/05/12/people_on_page_yasns.php

Smith, H. J., Milberg, S. J. and Burke, S .J. (1996) Information privacy: Measuring individuals' concerns about organizational practices. MIS Quart. 20(2) 167–196.

Smith, H. J., Dinev, T., and Xu, H. (2011). Information Privacy Research: An Interdisciplinary Review, MIS Quarterly, Vol. 35, No. 4, pp. 989-1015.

Son, J., Y., and Kim,S.,S.,(2008) Internet Users' Information Privacy-Protective Responses: A Taxonomy and a Nomological Model. MIS Quarterly Vol. 32 No. 3, pp. 503-529/September 2008

Solove, J.,D. (2006). A Taxonomy of Privacy. University of Pennsylvania Law Review, vol 154 No.3.

Solove, J., D. (2008). Understanding Privacy. Harvard University Press , 2008.

Stake, R.E.(1995). The art of case study. Thousand Oaks, CA: Sage Publications.

State of the Blogsphere (2011). State of the Blogosphere 2011: Part 2. Accessed in Jan 2012 from: http://technorati.com/social-media/article/state-of-the-blogosphere-2011-part2/

Stewart, K. A., A. H. Segars. (2002). An empirical examination of the concern for information privacy instrument. Information Systems Research. 13(1) 36–49.

Stone, B.(2008). Facebook Aims to Extend Its Reach Across the Web. Retrieved 10 Dec 2008http://www.nytimes.com/2008/12/01/technology/internet/01facebook.html?pagewanted=1&_r=2&partner=rss&emc=rss

Stone, E. F., Gardner, D. G., Gueutal, H. G., and McClure, S. (1983). A Field Experiment Comparing Information-Privacy Values, Beliefs, and Attitudes across Several Types of Organizations. Journal of Applied Psychology (68:3), pp. 459-468.

Strater, K. and Lipford, H.R.(2008). Strategies and struggles with privacy in an online social networking community. Proceedings of the 22nd British HCI Group Annual Conference on HCI 2008: People and Computers XXII: Culture, Creativity, Interaction - Volume 1.

Strauss, A.L., and Corbin, J. (1990) Basics of Qualitative Research. Grounded Theory Procedures and Techniques, Sage, Newbury Park, CA.

Strauss, A.L., and Corbin, J. (1998) Basics of qualitative research : techniques and procedures for developing grounded theory, Thousand Oaks, CA: Sage.

Sunden, J. (2003). Material Virtualities. New York: Peter Lang.

Swearingen, J. (2008). Four Ways Social Networking Can Build Business. Accessed on Dec 20,2009 from : http://www.bnet.com/2403-13070_23-219914.html

Sweeney, L. (2000) Uniqueness of simple demographics in the US population. Technical Report LIDAPWP4, Data Privacy Laboratory, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Tashakkori, A., C. Teddlie. (1998). Mixed Methodology: Combining Qualitative and Quantitative Approaches. Sage Publications, London, U.K.

Thelwall, M. & Hasler, L. (2007). Blog search engines. *Online Information Review*, 31(4), 467-479

Thelwall, M. (2009). Social network sites: Users and uses. In: M. Zelkowitz (Ed.), Advances in Computers 76. Amsterdam: Elsevier (pp. 19-73).

Thurm, S., & Kane, Y.I. (2010). Your apps are watching you: A WSJ investigation finds that iPhone and android apps are breaching the privacy of Smartphone users. The Wall Street Journal. Retrieved from http://online.wsj.com/article/SB10001424052748704694004576020083703574602.html#articleTabs%3Darticle.

Tsotsis, A. (2010). 250 Million People Now Using Facebook Connect Every Month. Retrieved on 10 Dec 2010 from: http://techcrunch.com/2010/12/08/250-million-people-now-connecting-via-facebook-connect/

Tufekci, Z. (2008 a) Can you see me now? Audience and disclosure regulation in online social network sites. Bulletin of Science, Technology and Society, 28(1), 20-36.

Tufekci, Z. (2008 b) Grooming, Gossip, Facebook and MySpace: What Can We Learn About These Sites From Those Who Won't Assimilate?. Information, Communication, and Society, 11 (4), 544 - 564.

Urquhart, C. (2001) An encounter with grounded theory: tackling the practical and philosophical issues. In: Qualitative Research in IS: Issues and Trends, Trauth, E. (ed.), pp. 104–140. Idea Group Publishing, Hershey, PA, USA.

Urquhart, C., Lehmann, H. and Myers, M.(2010). Putting the 'theory' back into grounded theory: guidelines for grounded theory studies in information systems. Information Systems Journal. 20, 357–381.

U.S code (2008). Privacy of customer information. Accessed on Oct 20,2009 from : http://www.law.cornell.edu/uscode/47/222.html

Van Maanen, J. (1983). Qualitative Methodology. Sage, London.

Van Maanen, J. (1998). Different strokes: Qualitative research in Administrative Science Quarterly from 1956 to 1996. In J. Van Maanen (ed.), Qualitative studies of organisations: ix-xxxii. Thousand Oaks, CA: Sage.

Waldo, J., Lin, H., & Millett, L.I. (2007). Engaging privacy and information technology in a digital age. Washington, D.C.: National Academies Press.

Walsham, G. (1995a). The Emergence of Interpretivism in IS Research, Information Systems Research, 6(4): 376-394.

Walsham, G. (1995b). Interpretive Case Studies in IS Research: Nature and Method, European Journal of Information Systems, 4: 74-81.

Walsham, G. and Waema, T. (1994). Information Systems Strategy and Implementation: A Case Study of a Building Society. ACM Transactions on Information Systems (12:2), pp. 150-173.

Warren, S,D. and Brandeis, L.D. (1890) The Right to Privacy, 4 HARV. L. REV. 193, 195-196

Webbmedia (2009). Facebook Connect. Retrieved on 20 March 2009 from: http://www.webbmediagroup.com/upload/pdf/WMG_FBConnect.pdf

Weick, K. E. (1993). The collapse of sense making in organizations: The Mann Gulch disaster. Administrative Science Quarterly, 38: 628–652.

Westin ,A.,F., (1967). Privacy and freedom. Atheneum, New York.

Blog #	Publication	Number of comments posted by users on or after publication of a blog	Total comments	Number of unique users who posted comments

Westin, A. F. (2003). Social and Political Dimensions of Privacy. Journal of Social Issues (59:2), pp. 431-453.

Wood-Harper, T. and Wood, B. (2005).Multiview as social informatics in action: past, present and future. Information Technology & People, 18(1): 26-32.

Wu Kuang-Wen , Shaio, Y. H, David C. Y. , Irina, P. (2012). The effect of online privacy policy on consumer privacy concern and trust, Computers in Human Behaviour, Volume 28, Issue 3, May 2012, Pages 889-897

Xu, H., Dinev, T., Smith, H. J., and Hart, P. (2008). Examining the Formation of Individual's Information Privacy Concerns: Toward an Integrative View, Proceedings of 29th Annual International Conference on Information Systems (ICIS), Paris, France, Paper 6.

Xu, H., Dinev, T., Smith, H. J., and Hart, P. (2011). Information Privacy Concerns: Linking Individual Perceptions with Institutional Privacy Assurances, Journal of the Association for Information Systems, Vol.12, No.12, pp. 798-824.

Yin, R.K. (2008) Case Study Research: Design and Methods, 4th edition, Sage Publications, Inc.

Yue, C., Xie, M. and Wang, H.(2007). Automatic Cookie Usage Setting with Cookie Picker. Proceedings of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2007. DSN apos;07 Volume , Issue , 25-28 June 2007 Page(s):460 – 470.

Zhao, Shanyang, Grasmuck, S., and Martin, J.. (2008). Identity Construction on Facebook: Digital Empowerment in Anchored Relationships. Computers in Human Behaviour, 24. 1816-1836

Zimmer, M.(2008).The Externalities of Search 2.0: The Emerging Privacy Threats when the Drive for the Perfect Search Engine meets Web 2.0. First Monday, Volume 13, Number 3 - 3 March 2008

APPENDIX A -
Table 1- Time series Data -Beacon blogs

		Ist day	2nd day	3rd day	More than 3 days		
1	09 Nov 07	8	3	1	1	13	10
2	21 Nov 07	8	1	0	2	11	8
3	20 Nov 07	2	0	0	0	2	2
4	12 Feb 08	22	16	7	2	47	41
5	07 Nov 07	3	0	0	0	3	2
6	06 Nov 07	0	3	0	0	3	2
7	11 Feb 08	40	0	0	0	40	32
8	30 Nov 07	6	8	2	4	20	17
9	31 Dec 07	3	2	0	1	6	4
10	29 Nov 07	46	10	2	2	58	50
11	05 Dec 07	7	3	3	0	13	10
12	06 Nov 07	90	40	10	6	146	90
13	03 Dec 07	48	8	0	0	56	52
14	05 Dec 07	4	0	0	4	8	5
15	10 Dec 07	22	4	3	3	32	30
16	19 Feb 08	0	4	0	1	5	4
17	22 Feb 08	0	1	0	0	1	1

18	20 Feb 08	2	1	0	0	3	3
19	21 Feb 08	7	1	1	0	9	5
20	06 Feb 08	1	1	0	0	2	2
21	14 Feb 08	5	1	0	0	6	4
22	21 Jan 08	1	0	0	0	1	1
23	20 Jan 08	1	0	0	0	1	1
24	02 Jan 08	1	0	0	0	1	1
25	04 Jan 07	2	2	0	0	4	2
26	29 Nov 07	12	3	2	1	18	15
27	21 Nov 07	10	3	1	2	16	10
28	22 Nov 07	7	2	1	1	11	9
29	05 Dec 07	2	0	0	0	2	2
30	03 Dec 07	50	10	5	4	69	58
31	30 Nov 07	60	10	6	2	78	50
32	01 Dec 07	140	15	5	5	165	100
33	27 Nov 07	3	1	0	1	5	4
34	05 Dec 07	64	5	2	1	72	62
35	18 Apr 08	26	6	3	4	39	32

36	11 dec 07	10	2	2	1	15	10
37	21 sept 09	14	4	2	3	23	18
38	10 dec 07	10	3	3	3	19	15
39	03 dec 07	3	2	0	0	5	2
40	23 Nov 07	4	1	0	0	5	2
41	30 Sept 08	4	1	0	0	5	2
42	18 Sep 09	2	0	0	0	2	2
43	21 Sep 09	1	0	0	0	1	1
44	12 Sep 09	1	0	0	0	1	1
45	19 Sep 09	100	20	20	7	147	70
	Total	**852**	**197**	**81**	**62**	**1190**	**844**
	% of total	**72%**	**17%**	**6%**	**5%**		

Privacy Themes- First Round of Open Coding

Nodes

Name	References
Business integrity	69
Transparency	40
User control	38
Automatic dissemination	37
Not opt-in	30
Third party data leak	28
Universal opt-out	28
Lack of incentive	24
Permanent information removal	24
Broadcast data	24
Uninformed user tracking outside SNS	22
No consent	21
Sell information to advertisers	20
Intrusion	19
Monetise personal information	19
No Notice	18
Public disclosure of private info	18
Deceptive practices	17
Granular control	14
No awareness of Business use of information	12
Embarrassing disclosures	12
Lack of trust on user data handling	10
Behavioral Targeting	9
Privacy diffusion	8
Identity theft	7
Information Ownership	7

Figure 1: Privacy themes first round of open coding

Nodes

Name	References
Push marketing	6
Interception	6
Inability to delete account	6
Sharing identity info with third parties	6
Stalking	5
Use of information by employers	5
Online tracking	5
Data retention	4
Tagging Photo	4
Security of data	4
Secondary use	3
False anonymity	3
Data collection even after opt-out	3
Online archeology	3
Third party info sharing	3
Invasion of personal space	3
Information aggregation	3
Repeated privacy breaches	3
Self Accountability	3
Cross-pollination of information	2
Terms of Service	2
Unsolicited ads	2
Spam	2
Disregard to users' privacy	2
Complex privacy controls	2
Push private-public boundaries	2

Figure 2: Privacy themes first round of open coding

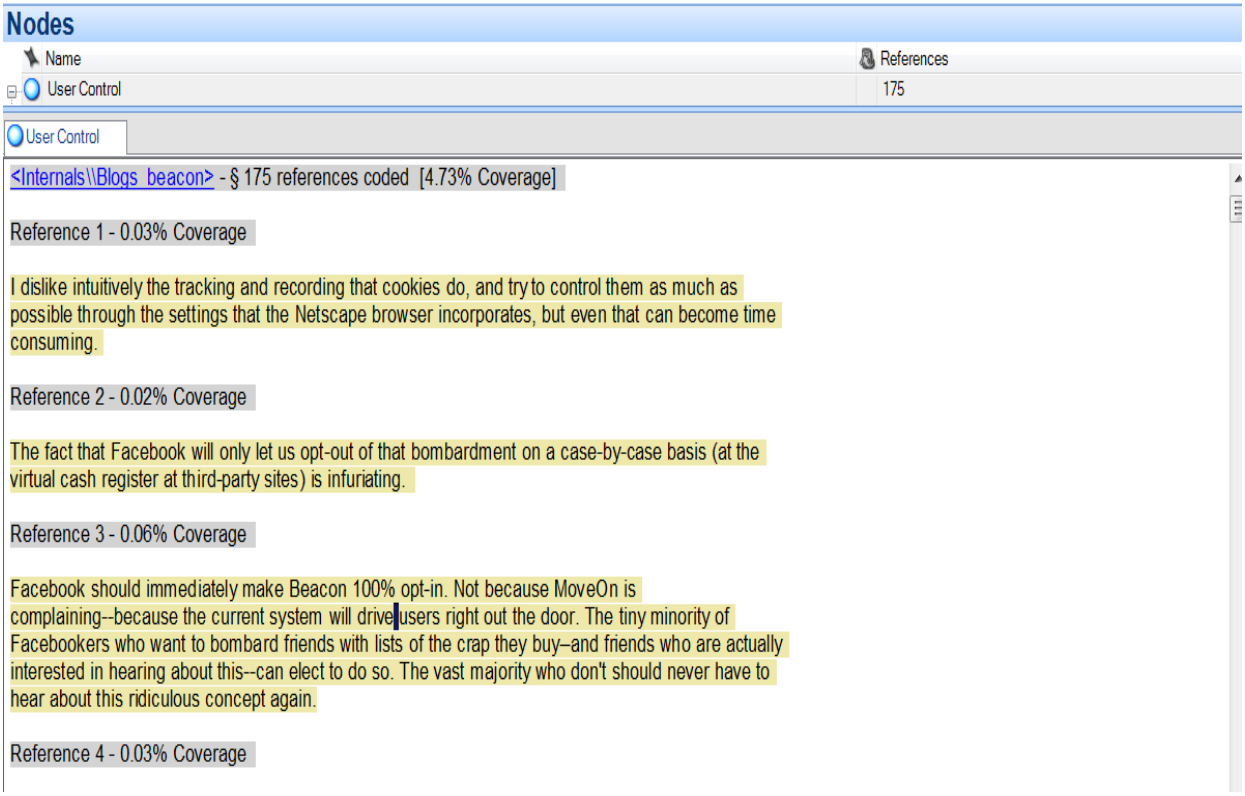
Nodes

Name	References
No awareness of changes in privacy controls	2
Fourth party info sharing	1
Push messaging	1
Social control	1
Fair data use policy	1
Limited coverage of laws	1

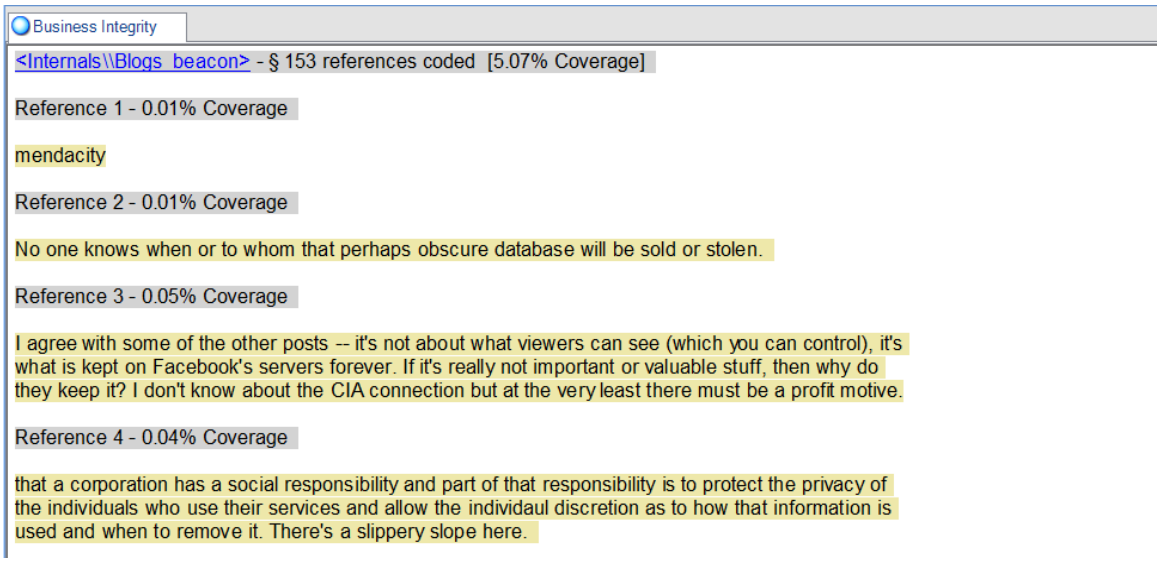
Figure 3: Privacy themes first round of open coding

CODE EXAMPLES (Beacon Case)

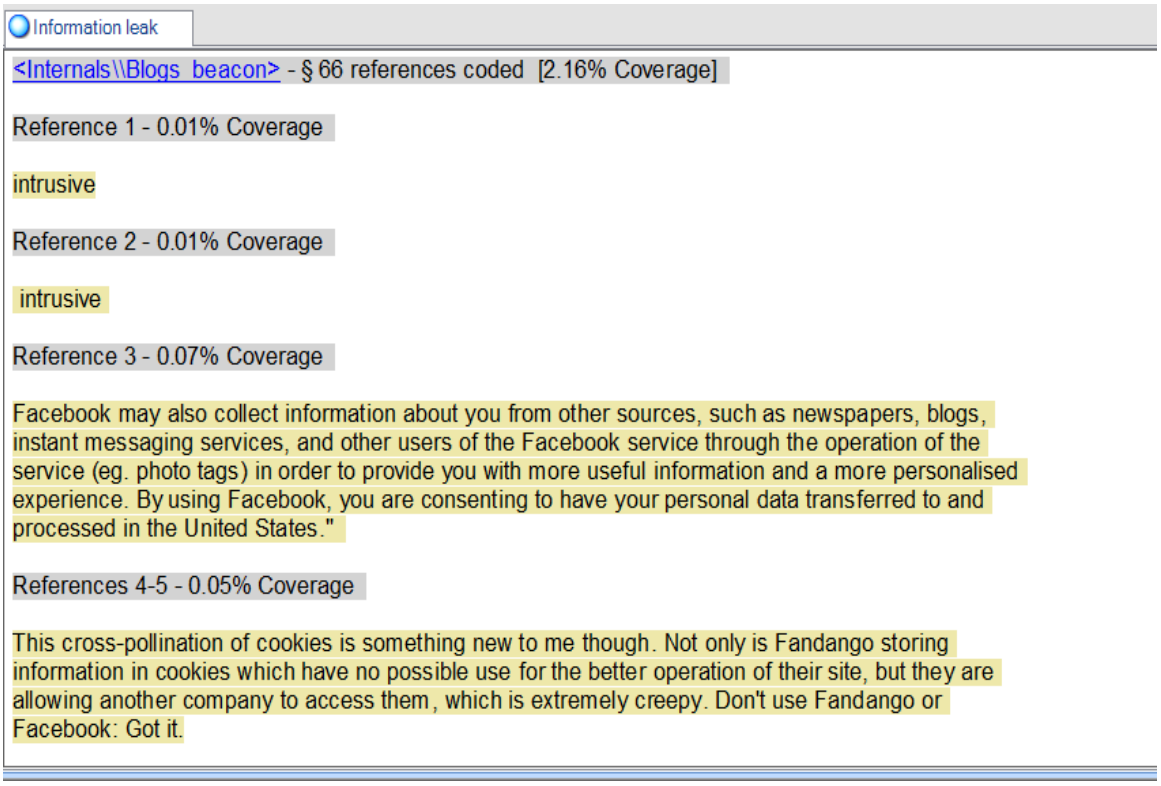
User Control



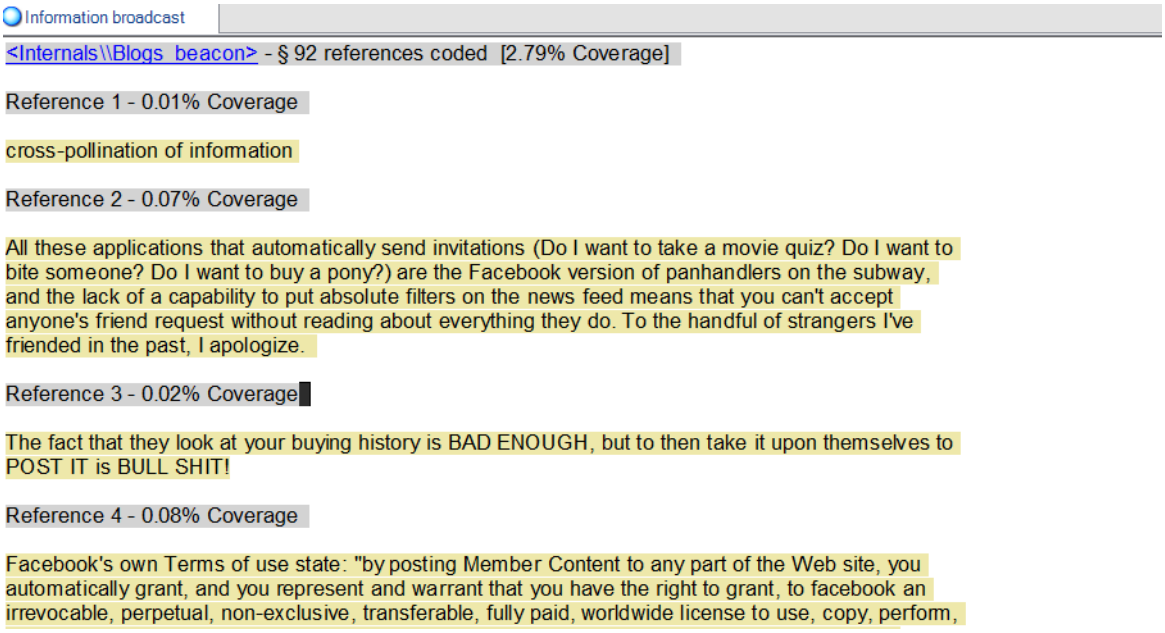
Business Integrity



Information Leak



Information Broadcast



APPENDIX B – CASE OF CONNECT LAUNCH

Time series data –blogs comments postings

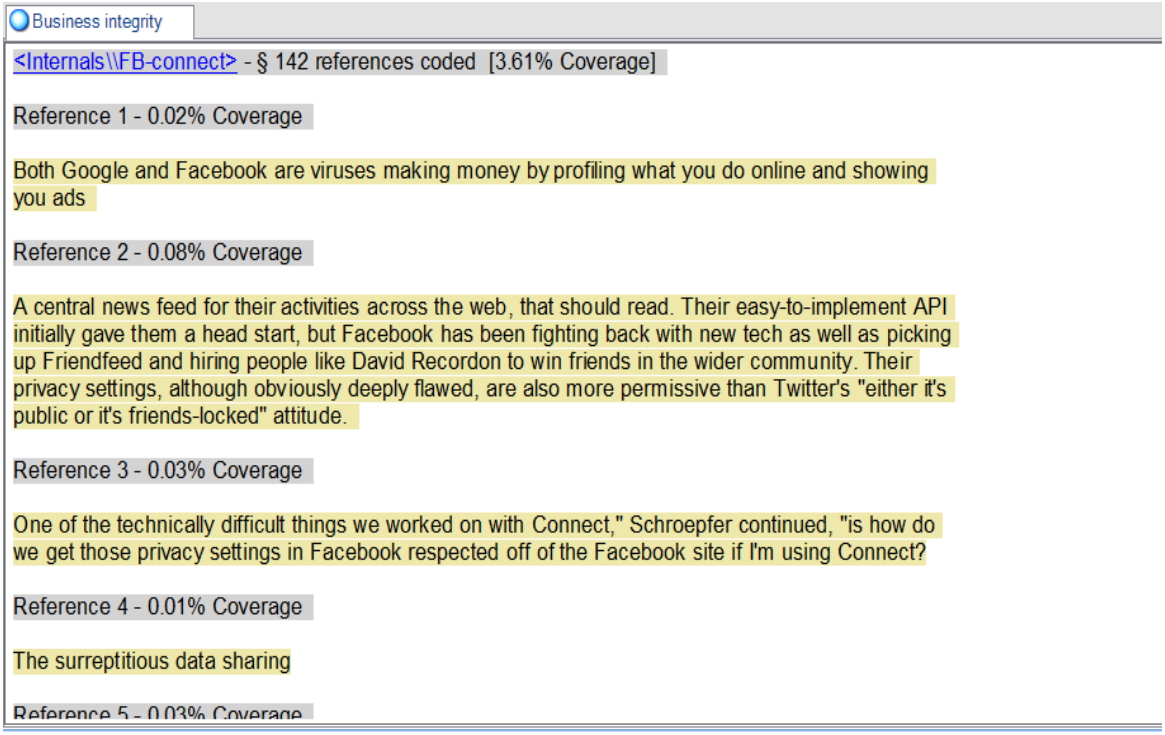
Blog #	Publication Date	Number of comments posted by users on or after publication of a blog				Total comments	Number of unique users who posted comments
		Ist day	2nd day	3rd day	More than 3 days		
1	24 Oct 10	044	3	2	0	49	49
2	15 Jan 10	48	2	1	0	51	51
3	14 July 10	4	0	0	0	4	4
4	21 May 10	72	5	1	0	78	72
5	30 Nov 10	41	4	3	1	49	42
6	21 Apr 10	6	0	0	1	7	6
7	22 Dec 09	13	3	0	0	16	9
8	10 June 10	3	0	0	0	3	3
9	05 July 10	2	0	0	0	2	2
10	25 Dec 08	21	0	0	0	21	21
11	26 May 09	3	0	0	0	3	3
12	04 Dec 08	15	1	0	0	16	16
13	16 Dec 08	6	0	0	0	6	6
14	27 Jan 10	30	1	0	0	31	25
15	29 Apr 09	42	16	2	2	62	57
16	04 Dec 08	23	0	0	1	24	22

17	04 Dec 08	17	0	0	0	17	15
18	04 Dec 08	125	6	1	1	133	122
Blog #	**Publication Date**	**Number of comments posted by users on:**				**Total comments**	**Number of unique users who posted comments**
		Ist day	**2nd day**	**3rd day**	**More than 3 days**		
19	04 Dec 08	36	2	0	0	38	35
20	26 Mar 10	99	5	0	8	112	103
21	01 July 10	2	0	0	0	2	2
22	14 Mar 09	8	0	0	0	8	7
23	21 Dec 09	8	0	0	0	8	7
24	30 Aug 10	2	0	0	1	3	3
25	21 Oct 09	33	2	1	0	36	33
26	20 July 09	43	1	2	0	46	42
27	01 Sep 10	5	0	0	1	6	6
28	01 Feb 09	77	3	0	1	81	74
29	19 May 10	7	0	0	1	8	7
30	17 May 10	8	0	0	0	8	7
31	01 Jan 09	4	0	0	6	10	9
32	11 Dec 08	6	0	0	0	6	6
33	02 Dec 08	7	1	0	0	8	7
34	10 Dec 08	5	0	0	0	5	5

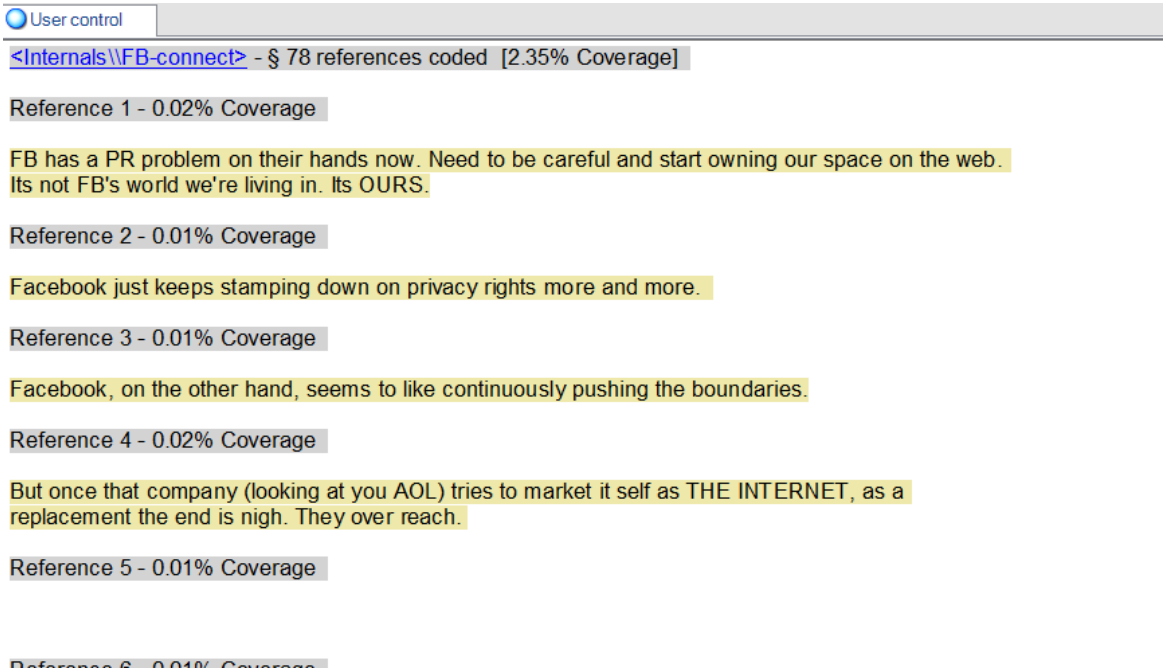
35	20 July 09	50	4	3	0	57	46
	Total	**915**	**59**	**16**	**24**	**1014**	**924**
	% of total	**90.24 %**	**5.82%**	**1.58%**	**2.36%**		

Coding examples: The case of Connect

Business integrity



User control


User control

<Internals\\FB-connect> - § 78 references coded [2.35% Coverage]

Reference 1 - 0.02% Coverage

FB has a PR problem on their hands now. Need to be careful and start owning our space on the web. Its not FB's world we're living in. Its OURS.

Reference 2 - 0.01% Coverage

Facebook just keeps stamping down on privacy rights more and more.

Reference 3 - 0.01% Coverage

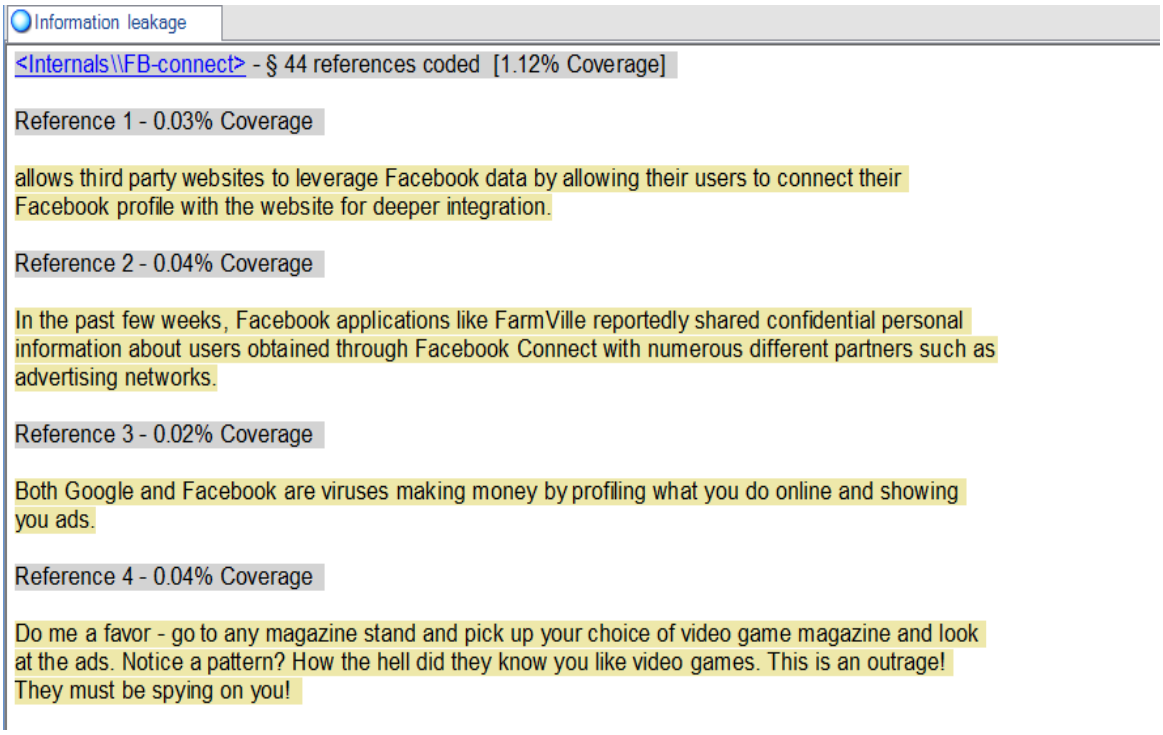
Facebook, on the other hand, seems to like continuously pushing the boundaries.

Reference 4 - 0.02% Coverage

But once that company (looking at you AOL) tries to market it self as THE INTERNET, as a replacement the end is nigh. They over reach.

Reference 5 - 0.01% Coverage

Information leak


Information leakage

<Internals\\FB-connect> - § 44 references coded [1.12% Coverage]

Reference 1 - 0.03% Coverage

allows third party websites to leverage Facebook data by allowing their users to connect their Facebook profile with the website for deeper integration.

Reference 2 - 0.04% Coverage

In the past few weeks, Facebook applications like FarmVille reportedly shared confidential personal information about users obtained through Facebook Connect with numerous different partners such as advertising networks.

Reference 3 - 0.02% Coverage

Both Google and Facebook are viruses making money by profiling what you do online and showing you ads.

Reference 4 - 0.04% Coverage

Do me a favor - go to any magazine stand and pick up your choice of video game magazine and look at the ads. Notice a pattern? How the hell did they know you like video games. This is an outrage! They must be spying on you!