Towards a Taxonomy of Privacy Concerns of Online Social Network Sites Users

A Case Study of Facebook Beacon

J a m a l  A r s h a d

Master of Science Thesis
Stockholm, Sweden 2010
TRITA-ICT-EX-2011:1
Towards a Taxonomy of Privacy Concerns of Online Social Network Sites Users: A Case Study of Facebook Beacon

Jamal Arshad (681231-7235)

August 2010

Department of Computer and Systems Sciences
Stockholm University / Royal Institute of Technology

This thesis corresponds to 20 credits for the author.
Abstract

More than half a billion people use Online Social Network Sites (OSNS) today. They disclose personally identifiable information such as names, date of birth, email address, phone numbers, and home address on their profiles. Some OSNS users also disclose their political and religious views and personal interests. The huge commercial potential of OSNS users’ information and the integration of OSNS with third party advertisers and/or aggregators pose threats to users’ privacy. This thesis reports a study which contributes to our understanding of the form and nature of online privacy by critically analysing the privacy concerns related with the failed launch of Facebook’s advertising tool Beacon. Beacon is an interesting case study because it highlighted the privacy concerns of OSNS users. Qualitative data was gathered from 29 weblogs (blogs) representing user opinions (492 comments) published between 6th November 2007 (when Beacon was launched) and 28th February 2008 (when commentary had dwindled). A thematic analysis of the blogs contributed in the development of a taxonomy of privacy concerns of OSNS users specifically related with the third party information use. Noticeably, the concerns such as commercialism, terms of service (TOS), lack of user control, lack of user awareness and data protection influence user perceptions of online privacy. The limitations and key implications for designers and service providers of OSNS are also discussed.

Keywords: Weblogs (blogs), social network sites (OSNS), privacy, thematic analysis, Facebook
CHAPTER ONE: Introduction

“Like “usability” and “security,” privacy is a holistic property of interactive systems, which include the people using them. An entire system may be ruined by a single poorly implemented component that leaks personal information…” (Iachello and Hong, 2007:p.5)

1.1 Chapter Overview

This research aims to investigate privacy concerns of Online Social Network Sites (OSNS) users which occur as a consequence of the interaction between online social network sites (OSNS) and third party companies. This chapter begins with the research background and motivation. The previous research on privacy concerns of OSNS users is introduced in the second section. In the third section, the aims and objectives of current research are highlighted. In the next section the research approach used to achieve aims and objectives of the current research is introduced. Finally, the structure of the thesis is presented.

1.2 Research Background

More than half a billion users use online social network sites (OSNS) today. They disclose personally identifiable information such as names, date of birth, email address, phone numbers, and home address on their profiles. Some OSNS users also disclose their political and religious views and personal interests. Although OSNS users gain many benefits of using OSNS such as: enjoyment (Rosen and Sherman, 2006; Sledgianowski and Kulviwat, 2008); self-presentation (Boyd, 2007); the ability to maintain social ties (Ellison et al., 2007); and the convenience of maintaining and developing relationships (Krashnova et al., 2010) but the use of SNS is not without a cost e.g. privacy concerns. Noticeably, almost all OSNS offer free membership
which means all of these services are free to OSNS users. Then the question arises how OSNS service providers are managing the availability of all these services for OSNS users for free. Perhaps, the answer lies in the commercial use of personal information of OSNS users by OSNS service providers and third party companies. For instance, Facebook runs more than 55000 external applications which have access of OSNS users’ data (Krishnamurthy and Wills, 2010). Therefore the huge commercial potential of OSNS users’ information and the integration of OSNS with the third party companies pose threats to users’ privacy because user information may be used in acts such as stalking, identity theft etc.

Whilst the general perception about users of OSNS is that they do not value privacy as they disclose too much private information about them on their profile and do not use privacy controls provided by OSNS (Gross and Acquisti, 2005). Likewise, Acquisti and Gross (2006) report entirely strange user behaviour in OSNS where in users on the one hand show concern about their privacy, yet they continue to disclose too much personal information on their profiles. However, the counter argument suggests that users do not utilize privacy controls as they feel privacy controls are not sufficient (Boyd, 2008) and also very complex (Strater and Richter, 2007) and perhaps, they do not understand privacy controls. Additionally, the current media coverage and privacy debate on the internet is changing this perception (Nov and Wattal, 2009; Young and Anabel, 2009). Gurses et al. (2008) found 124 articles on privacy breaches in online social networks between Oct 2007 and March 2008. Similarly, another study highlights that privacy concerns of users in OSNS affect user behaviour such that users try to a) control the disclosure of information and b) adopt more strict privacy controls (Nov and Wattal, 2009).

This thesis reports a case study of Facebook’s personalised marketing tool Beacon which was withdrawn just one month after its launch in November 2007 due to severe user criticism on privacy grounds (Facebook, 2007). Facebook is the largest OSNS today with over 500 million users (Facebook, 2010a). It is also the second leading website by data traffic and number of page views (see Alexa.com).
The primary premise of Beacon was to leverage social networks by allowing third party companies such as eBay, Fandango and Travelocity to encourage OSNS 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. When such an action is performed on a participating company website, a Beacon alert occurred informing the user that it is going to automatically share their ‘story’ with their Facebook friends unless the user choose to specifically opt-out of that particular action (Jamal and Cole, 2009). However, Beacon was withdrawn after a month by Facebook because of severe user criticism on privacy grounds.

Although privacy has been defined variously by various researchers (see chapter two for details), the most widely used privacy definition is that of Westin’s (1967) who defined privacy as the ability of individuals, groups, or institutions to control when, how, and to what extent information about them is communicated to others (Westin, 1967). The control can be exercised if individuals have the ability to manage the collection, processing and dissemination of information by others. However, this thesis utilizes the narrow view of privacy as data protection offered by the Council of Europe’s 1981 Convention for the Protection of Individuals with regard to the Automatic Processing of Personal Data. It is because of the nature of the advertising tool Beacon which used OSNS users’ personal data automatically without even the knowledge of or notice to the users.

As such the Council of Europe’s 1981 Convention for the Protection of Individuals with regard to the Automatic Processing of Personal Data required that personal information must be:

- Obtained fairly and lawfully;
- Used only for the original specified purpose;
- Accurate and up to date;
- Accessible to the subject;
- Kept secure; and
- Destroyed after its purpose is completed.

Robert Ellis Smith, editor of Privacy Journal, regards this interpretation being too
narrow and offers a broader definition: “the desire by each of us for physical space where we can be free of interruption, intrusion, embarrassment, or accountability and the attempt to control the time and manner of disclosures of personal information about ourselves” (2000).

For the purposes of this study, the narrow legalistic view of data protection offers a basis for considering the automatic processing aspects embodied within Beacon. It is combined with the broader perspectives of privacy which offer a useful starting point for considering user concerns within online social networks. Together, these two approaches to privacy provide the theoretical lens for analysing OSNS users’ privacy concern related with the launch of Beacon.

Accordingly significant number of studies have highlighted privacy concerns of OSNS users (e.g. Acquisti and Gross, 2006; Bonneau et al.,2009; Boyd and Ellison, 2007; Debatin et al, 2009; Gross and Acquisti, 2005; Rosenblum , 2007; Strater and Lipford, 2008). The media also have highlighted privacy concerns of OSNS users (Gurses et al., 2008). The privacy concerns OSNS users such as the damaged reputation due to rumour and gossip, stalking, phishing, identity theft, and the use of personal data by third parties are only few to name (Boyd and Ellison, 2007; Debatin et al., 2009; Gross and Acquisti, 2005).

However, majority of the research studies have focussed on privacy concerns which arise as a consequence of either excessive disclosures or the lack of the use of protection strategies by OSNS users (e.g. Acquisti and Gross,2006; Barnes, 2006; Boyd and Ellison,2007; Govani and Pashley, 2005; Strater and Lipford, 2008; Tufekci, 2008). Therefore the focus of majority of the privacy studies in OSNS was on tiny visible part of the iceberg (see figure 1.1) which comprises of only the users’ profiles and their interactions (Debatin et al., 2009). Whilst, there are fewer research studies which have focussed on the much larger invisible part of that iceberg which is constantly being fed with huge amount of OSNS users’ data , shared/leaked to third parties by OSNS operators (Bonneau et al., 2009; Debatin et al., 2009; Krishnamurthy and Wills, 2010).
In order to fill this gap in the current knowledge, this thesis, therefore aims to investigate privacy concerns of OSNS users related with the launch of Facebook’s personalised advertising tool Beacon and in doing so contributes in the development of a taxonomy of privacy concerns specifically related with the interaction between OSNS and third party companies.

1.3 Research Motivation

The researcher derives motivations for the current research because of the number of reasons which consequently become research motivations and argues that researching privacy in OSNS is more challenging and critical than for other websites (both e-commerce and other).

Firstly, due to the nature of online information which is: persistent (it persists almost permanently as compared to speech or other information shared in offline life); searchable (a part or all information can be searched with search engines); replicable (the information can be easily copied); and invisible audiences (difficult to tell who views the information) (Boyd, 2008). These characteristics also apply to digital data. These characteristics pose threats to online information and many studies have
highlighted privacy concerns of OSNS users (e.g. Gross and Acquisti, 2005; Boyd, 2008).

Secondly because the traditional context of privacy existed within a one-to-one environment of information disclosure, online social networks are based on information broadcast principles (Jamal and Cole, 2009). Digital information is characterised by an extremely porous nature whilst the network is designed to support widespread information dissemination (Rosenblum, 2007). Consequently, Dwyer et al. (2007) argues that “privacy within OSNS 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. The researcher argues that the failed launch of Beacon is the typical example of this viewpoint.

Thirdly, because of the depth and vast amount of personal information shared by over half a billion people on OSNS (Krishnamurthy and Wills, 2010; Strater and Lipford, 2008). OSNS users often willingly share personally identifiable information such as name, date of birth, and phone number (Acquisti and Gross, 2006; Gross and Acquisti, 2005; Krishnamurthy and Wills, 2010) which consequently poses serious threats to their privacy.

Fourth is the ability of online companies to track users’ online activities and link gathered behavioural information (e.g. which sites they visit) to other online and offline information (FTC, 2009). Likewise, Krishnamurthy and Wills (2010) demonstrate third party advertisers and/or aggregators track online users’ activities pervasively without their knowledge to send them relevant advertisements (Krishnamurthy and Wills, 2010). The interest in tracking online users’ activities by advertisers and/or aggregators has increased significantly because of the intimacy and availability of up-to-date personal information of online social network sites users (OSNS) (Bonneau et al., 2009) which consequently poses threats to users’ privacy in OSNS.

Also as the OSNS users’ demographics in the US suggests that 65% of the users are teens (12-17 years) and 35% of the users are adults (18 years and above) (Lenhart,
2009), therefore the majority of OSNS users (teens) are vulnerable to threats associated with the leakage of data on OSNS.

Finally, because OSNS operators allow access of user information (often without the knowledge of OSNS users) to external applications and third party companies who can then link it with already kept behavioural data of users (e.g. which sites they visit) (Krishnamurthy and Wills, 2010; Strater and Lipford, 2008) and hence pose threats to users’ privacy in OSNS. Also, the leakage of privacy by third parties is more serious given the huge commercial potential for the use of OSNS users’ information (Felt and Evans, 2008) and the use of sophisticated tracking software and data mining by third party companies (Acquisti and Gross, 2006; Debatin et al., 2009; Jones and Soltren, 2005; Krishnamurthy and Wills, 2010).

1.4 Research Aims and Objectives

The aim of this research therefore, is to investigate the privacy concerns of OSNS users associated with the failed launch of the Facebook’s personalised advertising tool Beacon. As such the research has the following objectives:

1. To investigate what view of privacy users of OSNS have in relation to sharing of their personal information with third party companies;
2. To investigate the privacy concerns of OSNS users specific with the use of Beacon by Facebook;
3. To investigate whether the privacy concerns can be arranged in the form of a taxonomy of privacy specifically related with third party information sharing in OSNS.
As such the study contributes in the current knowledge base in the following three ways:

1. It focuses on users’ online behaviour specific to third party information sharing in OSNS rather than the generic information disclosure behaviour in OSNS;
2. It focuses on highlighting privacy concerns of OSNS users specific to third party information sharing as well develops an understanding of the complexity of the management of OSNS users’ personal information;
3. It develops a privacy taxonomy specific to OSNS users in the context of the interaction between OSNS and the third party companies.

1.5 The Research Methods Approach

This research adopts an interpretive research approach which is based on the ontological belief that knowledge about a reality is gained through language, consciousness and shared meaning (Klein and Myers, 1999). In the context of this research, the knowledge pertains to the reasons why OSNS users criticised Beacon which can best be obtained through language (user commentary of blogs published as a reaction to the launch of Beacon) and shared meaning of OSNS users.

After selection of the philosophical basis (epistemology) such as interpretivism, the next task is to select appropriate research method. Although several research methods exist within an interpretive research approach such as action research, case study, ethnography and hermeneutics. However, case study research approach as propounded by Yin (2002) is adopted in this research as being most appropriate which allowed the researcher to examine users’ behaviour relating to the failed launch of Facebook’s personalised advertising tool Beacon. The Beacon case study allowed the researcher to investigate privacy concerns of OSNS users which occur as a consequence of the interaction between third party companies and OSNS. As such neither of the other research methods such as action research, ethnography and hermeneutics could have been used in this particular research. For instance, ethnography could not be chosen because it was impossible for the researcher to
study the privacy concerns in Beacon by immersing into the field.

Finally the next choice which the researcher has to make was to use appropriate data collection and analysis method. The researcher argues that blogs was the only sources of data available given the speed of users’ reactions and the geographical distribution of users across the world. Also because of the access to information which could not have been possible without users’ blogs which were published as a reaction to the launch of Beacon. Furthermore, the use of blogs data provided a most suitable means of collecting reliable user opinions because blogs offer real-time, unedited user commentaries (Gruhl et al, 2005; Thelwall and Hasler, 2006) and also represent a rich source of qualitative data which is unbiased by a research process (Jones and Alony, 2008).

Additionally, the use of blogs data in text form from the actual subjects (participants) relieved the researcher to transcribe (which was required in case of audio interview data) that may cause errors during transformation or conversion (Razera et al., 2010) and hence helped the researcher to maintain data quality.

Thematic analysis method is used to analyse blogs data as being most appropriate. Among the reason of using thematic analysis is its flexibility as compared to other methods of analysis such as conversation analysis and interpretative phenomenological analysis which either are tied to or originate from epistemology or theoretical considerations (Braun and Clarke, 2006). Moreover, conversation analysis cannot be adopted in this thesis as a mode of analysis as it deals with the audio or video data unlike the textual data of users’ blogs.

Another dominant motivation to use thematic analysis is because it is a foundational qualitative method which should be used as a first qualitative method by the early researchers (Braun and Clarke, 2006). Therefore recognising the diverse and complex nature of qualitative methods (Holloway and Todres, 2003) and the fact that the researcher did not possess past experience of conducting qualitative data analysis, thematic analysis was adopted in the current thesis.

See chapter three for details of the research approaches used in the current research.
1.6 Structure of the Thesis

Chapter 1: Introduction
Research background, motivations, aim and objectives and Research Methods Approach

Chapter 2: Extended Background
Literature Review (OSNS and Privacy)

Chapter 3: Research Methods Approach
Research Approach and Methods

Chapter 4: Case Study: Thematic Analysis of Blogs
Study set-up, Research Design and Data Collection

Chapter 5: Empirical findings
A Taxonomy of Privacy Concerns of OSNS

Chapter 6: Discussions and Synthesis of Results
Analysis of Data and Discussion

Chapter 7: Conclusions, limitations and Future Research
Conclusions, Contribution, Limitations and Future Research
CHAPTER TWO: Extended Background

2.1 Chapter Overview

The previous chapter highlighted the research problem addressed in this thesis which is reiterated here as the investigation of the privacy concerns of OSNS users related with the interaction between third party companies and OSNS. This chapter provides the foundation by laying out the theoretical dimensions of the research, and looks at the challenging issue of privacy in OSNS with a view to identify gaps in the current research and to justify further investigation of the issue. First section introduces the characteristics of OSNS, characteristics of OSNS users, information disclosure characteristics, and the third party interaction with OSNS. The next section details theories of information privacy. Finally, the discussion of privacy concerns in OSNS is presented.

2.2 Online Social Network Sites (OSNS)

Online social network sites (OSNS) have experienced an unprecedented growth in past few years. Today, over half a billion people use online OSNS (Krishnamurthy and Wills, 2009b). The examples of popular OSNS include Bebo, Facebook, Friendster, Hi5, MySpace, Orkut and Twitter. Facebook is the largest OSNS with over 500 million users (Facebook, 2010a) and is the 2nd leading website for data traffic (see Alexa.com).

According to Boyd and Ellison (2007) OSNS are “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).
The definition highlights three characteristics of OSNS. First, they allow users to create an online identity called profile where a user can “type oneself into being” (Sunden, 2003, p. 3).

Profiles include information such as name (first name, middle name and surname), address, email address, gender, date of birth, relationship status (single, married and other), education, work information (such as university name, course name, employer name, job title, work address), political and religious views, photos, and other information. Additionally, profiles are linked to other members’ pictures, videos and blogs.

OSNS users’ profiles can be divided into public and private parts. The public part of profiles can be seen by all OSNS users as well as others who do not have an account or even web crawlers (Bonneau et al., 2009a). A person starts by creating a profile by filling in personal information such as name, email address, gender, location, school, work place, interests, photo, etc. Of this information, the user can make some information private (meaning that only his/her friends or he/she himself/herself can view that information and no one else can view it) and leave other information public (meaning that everyone including strangers who either has a SNS profile or not can view that information). On the other hand SNS such as Facebook also allows access to users’ public profiles listings (containing information such as user name, photo and photos of eight random friends) to those who have not logged in their accounts including web crawlers (Bonneau et al., 2009b). The default settings of profiles on most of the OSNS are public. However, OSNS users can change public default settings to private. Many studies show that a significant majority of OSNS users do not change default settings of their profiles. Only 1.2% of college Facebook users at Carnegie Mellon University changed the searchability of their thumbnail profiles (consisting of names and photos only); 0.06% of users changed their profiles’ visibility (Gross and Acquisti, 2005) and 99% of users in a study of 67000 Twitter users did not change default privacy settings of their profiles (Krishnamurthy et al., 2008). However, it is not easy to draw a line between what is public or private in online OSNS. As Boyd (2007a) states “What it means to be public or private is quickly changing before our eyes” (p.1).
Second, OSNS allow users to create relationships (contacts) with other OSNS users, which enable them to form a social network of online friends— that they can communicate and share information. The number of contacts determines the strength of a social network of a user which consequently determines the strength of a OSNS. For example, an update on a user’s page with a network of 20 relationships is broadcasted not only to these 20 contacts, but also to a very large number of sub-contacts which a user can barely realize (Dube and Adomaitis, 2009). Therefore, OSNS nourish on relationships (Dube and Adomaitis, 2009).

Finally, OSNS allow users to traverse their own as well as friends’ social networks and thus become part of a larger network called community. As members of a community hold common beliefs or interests, groups are founded within OSNS that share commonalities such as alumni of a high school (Dube and Adomaitis, 2009).

Boyd and Ellison’s (2007) definition and terminology is criticized by Beer (2008) and Lang (2007) that it is too broad to include sites like YouTube (www.Youtube.com) which is predominately a video sharing site and the interaction between users is not the main focus of this site (Thelwall, 2009). Although YouTube does allow creating a profile page called channels (in YouTube) and also allows people to add friends which are two main characteristics of a social network site. However, the primary focus of users in YouTube is to watch and share videos. Whereas OSNS aim to facilitate users’ interaction and not just sharing of videos. Unconvincingly, however, YouTube can be viewed as a navigating social network as it allows users to find videos by browsing selected video posters and their friends (Thelwall, 2009).

2.2.1 Brief History and Typology

Boyd and Ellison (2007) presented a detailed history of social network sites in their editorial introduction to a journal special issue on social network sites (figure 1 highlights complete timeline of OSNS). According to them, the features of social network sites matured from relatively unsuccessful experiments. The first OSNS SixDegrees.com was launched in 1997, which allowed users to create profiles, list their friends and browse their friends’ list in 1998. Although these features existed
before the launch of sixdegrees.com, they were combined in its launch for the first time. For example, profile feature existed in the dating and community sites and list of friends existed in AIM and ICQ buddy lists (Boyd and Ellison, 2007). Sixdegrees attracted millions of users, but could not sustain the business and closed in 2000. The reasons for its failure as suggested by Boyd and Ellison (2007) include: most users did not have many online friends; the site offered limited functionality apart from adding friends and most users were not interested in meeting strangers.

Many other OSNS launched between 1997 and 2000 allowed creating profiles and articulated lists of friends (Boyd and Ellison, 2007). The OSNS namely AsianAvenue, BlackPlanet, and MiGente allowed users to create personal, professional and dating profiles. Interestingly, they allowed identifying friends on their personal profiles without seeking approval from them (O. Wasow, personal communication, August 16, 2007: referred in Boyd and Ellison, 2007) unlike today’s OSNS which require approval from connections before their names could be published on their friends’ profiles. Similarly, LiveJournal and Korean virtual world Cyworld were launched in 1999, but Cyworld offered OSNS features in 2001 (Kim and Yun, 2007). In 2000, LunarStorm—a Swedish web community also entered in the OSNS stream by including friends lists, guest books and diary pages (D. Skog, personal communication, September 24, 2007: in Boyd and Ellison, 2007). However, the first wave of OSNS ended in 2001. To sum up, the first wave of OSNS offered features such as: creating profiles (personal, professional and dating); and create friends lists, guest books and diary pages.

The second wave of OSNS began in 2001 to help people leverage their business networks (Boyd and Ellison, 2007). The examples of OSNS included are Ryze, Tribe.net, LinkedIn, and Friendster. Ryze was introduced in 2001 primarily for San Francisco business and technology community, including entrepreneurs of OSNS like Tribe.net, LinkedIn and Friendster (A. Scott, personal communication, June 14, 2007: in Boyd and Ellison, 2007). Interestingly, originators of OSNS namely Ryze, Tribe.net, LinkedIn, and Friendster were closely linked as they thought they can help each other instead of competing (Festa, 2003). However, only Friendster and LinkedIn gained mass popularity; Ryze and Tribe.net could not gain much popularity (Boyd and Ellison, 2007).
Launched in 2002, Friendster was designed to compete with popular dating site match.com (Cohen, 2003: in Boyd and Ellison, 2007). However, Friendster was focused to help friends-of-friends unlike the match.com which focused on introducing strangers with common interests (Boyd and Ellison, 2007). Bloggers, attendees of the Burning Man arts festival, and gay men were the early adopters of Friendster (Boyd, 2004). Friendster gained wide spread popularity in the mid of 2003 (O’Shea, 2003). However, its popularity began to fade due to number of factors as highlighted by Boyd (2006) including : technical such as servers were not equipped well to cope with the growth; social such as problems in maintaining cultural balance between users with the rapid increase in the number of users ; and mistrust between users and the site due to deletion of accounts of Fakesters (that their accounts were still liked by people for fun) and genuine users that chose not to display their real picture). While the popularity of Friendster was fading in the U.S, at the same time, it was booming in the Philippines, Singapore, Malaysia, and Indonesia (Goldberg, 2007). Probably that is why Friendster had an offer of $100 million from an Asian buyer as Lee (2009) reports. However, Arrington (2009) recently claimed that this deal is finalised for only $24.6 million.

Finally, the third wave of OSNS began in 2003 when OSNS hit the mainstream which has prompted social software analyst Shirky (2003: p.1) to introduce the term YAOSNS: "Yet Another Social Networking Service.". Trying to replicate the success of Friendster, most OSNS became profile-centric (Boyd and Ellison, 2007). MySpace launched in 2003, began competing sites namely Friendster, Xanga, and AsianAvenue. Particularly, MySpace management wanted to target disgruntled users of Friendster (T. Anderson, personal communication, February 2, 2006: in Boyd and Ellison, 2007). Because of improved personalised services to users and joining of disgruntled users of Friendster in large numbers, MySpace grew quickly (Boyd and Ellison, 2007). Boyd and Ellison (2007) report group of music bands which were expelled from Friendster membership due to non compliance of profile regulations. Likewise, by that time social network facilities have matured to include: group formation, blogging services, improved personalisation services, and ability to develop third applications e.g. compare movies (Boyd and Ellison, 2007).
A significant majority of popular OSNS today (2009) are introduced after 2003. Examples of popular OSNS include Facebook, MySpace, Hi5, Flicker, Orkut and Twitter.

Facebook, one of the largest OSNS today has over 500 million users (Facebook, 2010). It was launched in early 2004 as a Harvard-only OSNS (Cassidy, 2006), but in September 2005 it allowed high school students, professionals inside corporate networks and later in 2006 allowed everyone (Boyd and Ellison, 2007). This decision has allowed Facebook to grow broadly as well exponentially (Boyd and Ellison, 2007). However, gaining access to corporate networks still require a valid.com address, where as the approval of administrator is required for high school network (Boyd and Ellison, 2007). The evolution of OSNS demonstrates that these sites are now organised around people and not interests unlike the other websites (Boyd and Ellison, 2007; Dube and Adomaitis, 2009).
Thelwall (2009) suggests a typology (see figure 2.2) to categorize OSNS based on three purposes: socializing, networking, and social navigation. In socializing OSNS, all activities centre on social communication between members. Examples are MySpace, Hi5, Bebo, Facebook, and Cyworld (Thelwall, 2009). In networking OSNS, the activities centre on non-social communication between members. For example, in LinkedIn communication mostly centre on establishing business contacts. Finally, in social navigation OSNS browsing mostly takes place to find a certain type of information. Examples include YouTube, digg, CiteUlike. For example, in YouTube browsing takes place to find certain video data. Social navigation resembles the ordinary navigation as the purpose of both is to find certain information, but they
are different such that social navigation relates to finding information posted by other people unlike the ordinary navigation on websites managed by companies e.g. a bank’s website.

Figure 2.2 Examples of sites with varied purposes for OSNS friendship
(Source: Thelwall, 2009)

### 2.2.2 OSNS Usage

We have seen a tremendous growth in the user base of OSNS and half a billion people use OSNS today (Krishnamurthy and Wills, 2009). In order to understand this phenomenon, Thelwall (2009) poses legitimate questions: why so many people use OSNS? , are OSNS a passing phenomenon or do they have significant power to sustain? Because OSNS satisfy a deep human need that is the desire to investigate and gossip about human relationship (Donath, 2007; Tufekci, 2008), they are likely to sustain unless replaced by something more powerful satisfying the same need (Thelwall, 2009). Donath (2007) argues that as OSNS use an interface that facilitates gossiping, therefore, they are proficient in transmitting information (such as attributes of friends and acquaintances as well as their relationships). This enables OSNS to increase the scale of our social world (Donath, 2007). Thelwall (2009) endorses this argument for the longevity of OSNS as we do not want to lose track of our social
interactions and contacts. Tufekci (2008) demonstrates in a study of 713 U.S college students that, social grooming (such as gossips, small talks) is an important phenomenon of social network use and people who do not value social grooming are significantly less likely to use OSNS. As social grooming helps to create relationship and relationship creation is an important characteristic of community formation. Therefore, social grooming helps in community building.

Similarly, Joinson (2008) reports the primary motivations to use Facebook in a study of 137 users. The primary motivations include: social connection (e.g. desire to connect and communicate); shared identities (e.g. joining groups or events); photographs (e.g. sharing, posting, tagging); content (e.g. applications, games etc.); social network browsing (e.g. viewing profiles of acquaintances or strangers) and social investigation (finding and contacting new people). These findings corroborate with the research of (Donath and Boyd, 2004; Ellison et al., 2006) who also show people use OSNS to publish personal profiles, share photos and videos in order to create and maintain virtual networks of friends.

Apart from the social motivations to use OSNS, however, Rosenblum (2007) finds that people use OSNS because they offer a simple and usable communication platform. In a more recent study of students’ use of OSNS in India, Agarwal and Mital (2009) identify three factors: widening of perspective (e.g. better job prospects, understanding business environment, career building etc.); sharing of opinions (e.g. seek help from people, make plans with your friends/contacts etc.) and personal socialization (e.g. make new friends, share opinions, stay in touch with friends/family/contacts/strangers etc.). However, two of the three factors (sharing of opinions and personal socialization) have already been reported in the research of (Donath and Boyd, 2004), Donath (2007), Ellison et al. (2006) and Joinson (2008). Nevertheless, the finding widening of perspective is interesting especially in cultural context. As the study was conducted in an Indian university, one of the motivations of students could be to go abroad (e.g. U.S or Europe) for better job prospects or higher education and OSNS may help them to seek useful information as a very high percentage of users belong to US or Europe.
Interestingly, OSNS have also been used by businesses. Businesses launch various OSNS applications and integrate them into corporate environment to enhance marketing and organisational performance (Li and Bernoff, 2008).

### 2.2.3 User Characteristics

Different generic OSNS (socializing OSNS) with the same features may have different user audiences. For example, Facebook users are more educated than MySpace users (Boyd, 2007), most likely because of Facebook’s origin in education (Thelwall, 2009). Also, Hargittai (2007) in a study of 1000 students of an ethnically distributed the U.S. university in 2007 found that 79% used Facebook, 55% MySpace and only 6% Xanga. Yet, in many cases it is not easy to understand the difference in user demographics of OSNS. Due to the Pew Internet & American Life project (http://www.pewinternet.org/), teens are the most widely studied OSNS users (Thelwall, 2009). 55% of the U.S. teens in 2006 used a OSNS (Lenhart et al., 2007). Similarly in the U.K., a survey of internet usage in 2007, students were most likely to have an OSNS profile as compared to pensioners who almost had none (Dutton and Elsper, 2007). These findings suggest the adoption pattern of OSNS users which are predominantly teens and young adults.

A recent report by Lenhart (2009) highlights important facts about the U.S. population using OSNS in 2009 such as: 65% teens (12-17 years) use OSNS as compared to 50% in 2006; 35% adults (18 years and older) use OSNS as compared to 8% in 2005; (54% female vs. 46% male) as compared to (55% male vs. 45% female) in 2005 ; a slightly higher proportion of younger adults aged 18-24 (28%) that make up OSNS population as compared with 19% in 2005 and a move from urban (from 41% in 2005 to 29% in 2009) to suburb (56%) and rural (15%) areas. These statistics highlight the growth of OSNS in the U.S. However, as adults make up a larger proportion of the U.S. population, that is why 35% adults exceed in number than 65% of teens.
2.2.4 Information Disclosure Characteristics

Boyd (2008) highlights four distinguishing features of information on OSNS: persistence (information persists more permanently than speech or other information shared in offline life); searchability (all or part of information can be searched with search engines within or outside OSNS); replicability (virtually all information can be easily copied) and invisible audiences (most OSNS do not report who views user information). Moreover, online OSNS are based on information broadcast principles unlike the traditional context of offline information sharing in a one-to-one environment. Also, digital information is characterised by an extremely porous nature whilst the network is designed to support widespread information dissemination (Rosenblum, 2007).

Gross and Acquisti (2005) in a field study of 4000 Carnegie Mellon University students, find out that majority of students have disclosed personal information on their Facebook profiles. For instance, 89% users used their real name on profiles and 90% have put profile image (figure 2.3 highlight other types of information). They argued that such a high percentage of disclosure of real information on profiles may be due to Facebook’s information facilitating design and its perceived connection to a physical and a bounded community (Facebook at that time was only restricted to students). Dwyer et al. (2007) also vouch the association of Facebook with physical entities (e.g. universities). Studies by other researchers corroborate the findings this study (e.g. Acquisti and Gross, 2006; Barnes, 2006; Govani and Pashley, 2005; Strater and Lipford, 2008; Tufekci, 2008).
However, Gross and Acquisti (2005) did not find any substantial difference in the information revelation by males and females except in phone numbers where male disclosure was high (47.1%) as compared to female (28.9%). Contrary to this is the finding in a study of 886 teens; that girls and boys differ in how they think about disclosing information online (Lenhart and Madden; 2007). For instance, girls were found more likely than boys to say that they have posted photos of both themselves and their friends and boys were found more likely than the girls to say they have posted their address, last names and phone numbers. But this disagreement between the findings may be due to the age differences of the participants in both the studies. As in the study of Gross and Acquisti (2005) the age range of participants was 18-24 years which was 12-17 years in the later study. Lenhart and Madden (2007) found that older teens (15-17) were more likely to share personal information and that is why the difference between disclosures of overall information between two groups (teens and adults) was not substantial (see table 2.1).

<table>
<thead>
<tr>
<th>Information item</th>
<th>Adults (Source: Gross and Acquisti;2005)</th>
<th>Teens (Source: Lenhart and Madden;2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Name</td>
<td>89%</td>
<td>82%</td>
</tr>
<tr>
<td>Profile image</td>
<td>90%</td>
<td>79%</td>
</tr>
<tr>
<td>Date of birth</td>
<td>87%</td>
<td>Not available</td>
</tr>
<tr>
<td>Home town</td>
<td>72%</td>
<td>61%</td>
</tr>
<tr>
<td>Home Address</td>
<td>51%</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Phone number</td>
<td>40%</td>
<td>Not available</td>
</tr>
<tr>
<td>School name</td>
<td>87%</td>
<td>49%</td>
</tr>
<tr>
<td>AIM screen name</td>
<td>77%</td>
<td>40%</td>
</tr>
</tbody>
</table>

Table 2.1: Comparison between information disclosure by adults (18-24) and teens (12-17)

Similarly, the studies of Campbell et al (2001) and Fogel and Nehmad (2008) highlight the effect of gender on information revelation behaviour. Campbell et al. (2001), in a study of 468 internet users, found that males have more awareness of privacy issues than females and that is why they are willing to engage in a more risky behaviour (e.g. divulging personal information). Presumably the males who have more awareness of privacy issues may also be familiar with privacy assurance mechanisms and techniques (e.g. encrypted transactions, anonymous browsing etc.) (Campbell et al., 2001). Perhaps, the same argument explains the finding of Acquisti and Gross (2006) that the privacy concerned individuals reveal great amount of information on their profile. Campbell et al (2001) also find that experienced internet users are more likely to engage in risky behaviour than their non experienced counterparts and the experienced male internet users are more likely to engage in risky behaviour than their experienced female users. Likewise, Fogel and Nehmad (2008) find that women (in OSNS) show more concerns about information disclosure and privacy than men. Their study also, corroborates the findings of Gross and Acquisti (2005) and Lenhart and Madden (2007) that male are more likely to disclose phone number and home addresses on OSNS.

Given the growth of OSNS in the U.S., the finding suggests that the geography influences the information disclosure and privacy needs of OSNS users (Lenhart, 2009).

2.2.5 Disclosure of Personally Identifiable Information in OSNS

The studies of Acquisti and Gross (2006), Gross and Acquisti (2005) and Krishnamurthy and Wills (2009b) highlight the thorny issue of disclosure of personally identifiable information on OSNS either voluntarily or inadvertently. Personally identifiable information (PII) is defined as below.
Johnson III (2007), the Deputy Director of Management, defines personally identifiable information (PII) as “the information which can be used to distinguish or trace an individual’s identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother’s maiden name, etc.” (p.1). A similar definition by the US Government Accountability Office (GAO) (2008) refers PII as containing all information that can be used to locate or identify an individual, such as names, aliases, Social Security numbers, biometric records, and other personal information that is linked or linkable to an individual.

The user information such as name (first and last), location(city), zip code, street, address, email address, telephone numbers, and photos (both personal and group) represent parts of PII in an OSNS (Krishnamurthy and Wills (2009b). They also included parts of information which can be linked to one of the above such as gender, birthday, age or birth year, schools, employer, friends and activities/interests.

The executives of a privacy coalition (Schwartz et al., 2007), in a report forwarded to the Federal Trade Commission (FTC), define PII as information that can directly or indirectly:

(1) “ identify an individual, including but not limited to name, address, IP address, SSN and/or other assigned identifier, or a combination of unique or non-unique identifying elements associated with a particular individual or that can be reasonably associated with a particular individual, or

(2) permit a set of behaviours or actions to be consistently associated with a particular individual or computer user, even if the individual or computer user is never identified by name or other individual identifier. Any set of actions and behaviours of an individual, if those actions create a uniquely identified being, is considered PII because the associated behavioural record can have tracking and/or targeting consequences “ (p.6).

Recently, Narayanan and Shmatikov (2008) demonstrated that users can also be identified from their online behaviours (such as which sites they visit and what
information they seek). They identified two individuals based on their preferences, recommendations, and transaction records.

McCallister et al. (2009) present an exhaustive list of personally identifiable information (PII) which includes name, personal identification number, address, asset information, telephone, personal characteristics and information identifying personally owned property. The examples of each type of personally identifiable information are shown in table 2.2.

<table>
<thead>
<tr>
<th>Personally Identifiable Information and examples</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>full name, maiden name, mother’s maiden name, or alias</td>
</tr>
<tr>
<td>Personal identification number</td>
<td>SSN, passport number, driver’s license number, taxpayer identification number, patient identification number, and financial account or credit card number</td>
</tr>
<tr>
<td>Address information</td>
<td>street address or email address</td>
</tr>
<tr>
<td>Asset information</td>
<td>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</td>
</tr>
<tr>
<td>Telephone numbers</td>
<td>mobile, business, and personal numbers</td>
</tr>
<tr>
<td>Personal characteristics</td>
<td>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</td>
</tr>
<tr>
<td>Information identifying personally owned property</td>
<td>vehicle registration or identification number and title numbers and related information</td>
</tr>
<tr>
<td>Information about an individual that is linked or linkable to one of the above</td>
<td>date of birth, place of birth, race, religion, weight, activities, or employment, medical, education, or financial information</td>
</tr>
</tbody>
</table>

Table 2.2: Personally Identifiable Information and examples (McCallister et al., 2009:p.2)

These definitions highlight two characteristics of PII: an ability to identify an individual and the combining ability of non-identifiable information to a specific person. Together, these two characteristics present a broad scope of PII to include virtually all types of information. Specifically, the combining ability of PII with non-PII and vice versa, offers limitless opportunities to third party trackers to identify individuals. An interesting finding suggests that the birth date, five digit zip code and gender can be used to identify 87% Americans (Sweeney, 2000). This is because the apparently anonymous information narrows down the population to such an extent that the combination of (date of birth, zip code and gender) uniquely identifies
majority (87%) of Americans (Schoen, 2009). Also, Federal Trade Commission (FTC) almost a decade ago, in a report to Congress, warned the possibility of linking online profile information with the offline data collected from consumers (FTC, 2000).

Unlike the linking of online profile information (mostly non-PII) with offline PII, Krishnamurthy and Wills (2009b) demonstrate that OSNS leak users’ PII to third party companies either directly or indirectly—who can then link this with the behavioural information (e.g. which web sites they visit, which advertisements they click and other information) they kept about users. This may de-anonymize or de-identify OSNS users. They explained how this leakage takes place in a straightforward manner. When users log in to OSNS, these sites includes advertising and tracking code which can read users’ id (a unique number pointing to every SNS user profile). Therefore these third party advertising companies can access users’ profile such that their OSNS account becomes available to third party tracking companies who can access OSNS profiles and record information on the files they keep on people. Their empirical research of 12 OSNS (Bebo, Digg, Facebook, Friendster, Hi5, Imeem, LiveJournal, MySpace, Orkut, Twitter, Xanga and LinkedIn), shows that all OSNS except Google owned Orkut did not leak PII to third party companies.

The examples demonstrate that the distinction between identifiable and non identifiable information has now become meaningless: a position that is highlighted in the FTC recent report which clarifies that the self regulatory principles for online behavioural advertising (details follows later in this chapter) shall apply to all practices involving information without any distinction being made to identifiable or non identifiable information (FTC, 2009).

However, the ease of de-anonymization or re-identification of individuals may have long term consequences for law makers, website service providers and online users alike. It is challenging for law makers because “these advances should trigger a sea change in the law, because nearly every information privacy law or regulation grants a get-out-of jail-free card to those who anonymize their data” (Ohm, 2009:p.4). For instance, The Federal Telecommunications privacy laws (U.S code, 2008), the Health Insurance Portability and Accountability Act (HIPPA, 1996) and the EU Data
Protection Directive (1995) all treat PII as sensitive information and protect it legally. Similarly, website service providers emphasize the protection of PII of users and devise privacy policies to protect it. Internet users often have intuitions about certain types of information; whether it is anonymous or not and use self management strategies for the protection of PII through disclosure. For instance, most commonly used self management strategies are: providing false information, avoiding information disclosures, and blocking of cookies.

These studies have demonstrated that OSNS leak information (either PII or non-PII) to third party trackers or advertisers who can link it with already kept behavioural information of the same users. Thus, posing a threat to re-identify or de-anonymize OSNS users—which can have long term consequences on the privacy of OSNS users. This also have long term consequences for law makers, OSNS service providers, website service providers and third party trackers or advertisers. However, the focus of this research is to investigate the effect of leakage of information to third party advertisers on the privacy of OSNS users. To date, there is paucity of research which has demonstrated the effect of leakage of information to third party advertisers and it’s linking with the behavioural information on the privacy of OSNS users; and also to investigate the effect of its subsequent linking with the behavioural information already collected by third party companies on the privacy of OSNS users.

2.2.6 Business Model of Social Network Sites

Primarily, three business models exist for SNS: advertising (e.g., Facebook), micropayments through supporting gift exchange function (e.g., Cyworld) and premium membership fee for enhanced features (e.g., flicker, mixi) (Thelwall, 2009). Because of the huge potential of using users’ personal data for targeted advertising, Thelwall (2009) suggests advertising will remain the predominant source of revenue for SNS. Targeted advertising (also called behavioural advertising) is the practice of sending relevant advertisements to users based on their behaviour (CIPPIC, 2008). Likewise, Morrissey (2009) acknowledges the importance of the use of users’ data by advertisers as “universe of social network sites presents a tempting pool of data for advertisers to use in order to improve targeting techniques”. Moreover, the CEO of Media6° (a marketing company), highlights the importance of SNS’ data as “It’s
getting back to the old adage that bird’s of a feather flock together”. Similarly, Preibusch et al. (2007) demonstrate that most SNS use users’ profile data for behavioural advertisement to generate revenues. For instance, BBC News (2008) reported that credit companies are advertising on SNS for cheap loans for people with poor credit ratings. Also, SNS allow advertisers and advertising networks to track users’ behavioural as well as profile data to send them targeted advertisements (Krishnamurthy and Wills, 2008). Advertising networks (e.g. DoubleClick) are the companies that provide content as well as advertisements to websites owned by first party servers (e.g. Facebook) (Krishnamurthy and Wills; 2009b). Some internet companies e.g. search engine providers also need to access the personal data of SNS users to improve their searching techniques (Zimmer, 2008).

The SNS are building advertisement alliances with giants like Google and Microsoft (Johnston, 2007). For instance, Microsoft spends $240 million to buy just 1.6% stake in Facebook, given Microsoft an exclusive right (only in the U.S.) to provide banner and sponsored links on Facebook. Similarly, Johnston (2007) reports that Google signed its exclusive advertising deal with MySpace in 2006. Bernoff and Li (2008) suggest that advertisers will use innovative ways to leverage SNS in order to make better connections with their customers. Facebook’s personalised advertising tool Beacon was an effort to enable SNS to leverage businesses such as eBay, Fandango and Travelocity to allow users to share various actions amongst their friends via automatic news feed. One month after the launch, Facebook withdrew Beacon (make it opt-in instead of opt-out) because of severe negative user commentary (highlighting privacy issues) offered via blogs and mirrored in the printed press.

2.2.7 OSNS and Third Party Interaction

Riphagen (2008), a researcher at Electronic Privacy Information Centre (EPIC), talked about generations of OSNS interaction with third parties, which has now been evolved into third generation. The first generation is concerned with pushing in information from third parties websites to OSNS. For example, the case of Facebook’s Beacon (which is the main focus in this thesis) was designed to sends information of user actions (e.g. purchase of an item) performed at third party
websites such as eBay, Fandango and Travelocity to Facebook users’ profiles (see chapter 4 for details). Whilst, the second generation allows OSNS to push out information to third parties and pull in processed information back to OSNS. OSNS give access of user information to third party application developers in order to build applications for OSNS who push out that information from OSNS and pull in processed information. For instance, Facebook runs more than 55000 external applications which have access of OSNS users’ data (Krishnamurthy and Wills, 2010). Finally, the third generation of interaction of OSNS and third parties is characterized by both pushing out information from OSNS to third parties and pulling in information from third parties websites to OSNS.

2.3 Information Privacy Concerns and Social Network Sites

Information privacy (or simply privacy in the context of this thesis) is “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others” (Westin, 1967:p.7). Likewise, Stone et al. (1983) defines information privacy as “the ability of the individual to personally control information about one’s self”. Perhaps, this perspective of information privacy is the most widely used in IS literature. However, several subsequent studies reported privacy as a mutli-dimensional concept rather than single dimensional phenomenon based on control over personal information. Malhotra et al. (2004) view information privacy as a multifaceted phenomenon and develop internet users’ information privacy concerns (IUIPC) consisting of three factors: collection of personal information by an online marketer; user control over collected information and the awareness of user on how the collected information about them will be used. Here distinction between ‘data controller’ and ‘data processor needs to be done. European Union (EU) in DIRECTIVE 95/46/EC1 defines data controller as “the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data” whereas data processor “mean a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller”. Malhotra et al. (2004) operationalized the control aspect of

information privacy and suggested that it can be exercised via approval, modification, and the choice to opt-in or opt-out. IUIPC allow two way communications –from users to web site. However, as they apply this theory in the e-commerce environment where information sharing is mandatory (Nov and Wattal, 2009) whereas the social networks, the information flow is more complicated: from a user to an SNS; from one user to another and from a user to third party trackers even without user knowledge.

Gross and Acquisti (2005) depicts this phenomenon very well in online SNS: “the relation between privacy and a person’s social network is multi-faceted. In certain occasions we want information about ourselves to be known only by a small circle of close friends, and not by strangers. In other instances, we are willing to reveal personal information to anonymous strangers, but not to those who know us better” (p.2). This perspective on privacy relate to Altman’s (1975) work, who states that management of privacy is a boundary negotiation process in which people switch between the private and public space.

However, as discussed in chapter 1, this thesis utilizes the narrow view of privacy as data protection offered by the Council of Europe’s 1981 Convention for the Protection of Individuals with regard to the Automatic Processing of Personal Data. It is because of the nature of the advertising tool Beacon which used OSNS users’ personally data automatically without even the awareness or notice of users.

As such the Council of Europe’s 1981 Convention for the Protection of Individuals with regard to the Automatic Processing of Personal Data required that personal information must be:

- Obtained fairly and lawfully;
- Used only for the original specified purpose;
- Accurate and up to date;
- Accessible to the subject;
- Kept secure; and
- Destroyed after its purpose is completed.

Robert Ellis Smith, editor of Privacy Journal, regards this interpretation being too narrow and offers a broader definition: “the desire by each of us for physical space
where we can be free of interruption, intrusion, embarrassment, or accountability and the attempt to control the time and manner of disclosures of personal information about ourselves” (2000).

For the purposes of this study, the narrow legalistic view of data protection offers a basis for considering the automatic processing aspects embodied within Beacon. It is combined with the broader perspectives of privacy which offer a useful starting point for considering user concerns within online social networks. Together, these two approaches to privacy provide the theoretical lens for analysing OSNS users’ privacy concern related with the launch of Beacon.

Online users’ privacy concerns have risen in this decade, many surveys report (Bettina, et al., 2005; Equifax, 1996; Harris & Westin, 1998; Smith, et al., 1996; Westin, 1997). Son and Kim (2008) suggest that “successfully addressing information privacy issues in an online environment is particularly relevant to the growth of the information age. This is especially true for online companies because their success and quality of customer service hinge largely on their ability to collect and analyze a large amount of personal information about Internet users (p.504). This applies very well to all online tracking companies (e.g. DoubleClick) who target SNS users for advertisement. However, the responses of users to information management practices (collection, processing, storage and dissemination) of third party advertisers should be given valued. For example, users do not like that information about them is retained for long time (Canadian Press, 2008; Centre for Democracy and Technology, 2008; Smith, 2008). The information privacy is viewed as a critical ethical issue which needs attention from scholars and practitioners (Smith et al., 1996).

The focus of information privacy in the past research has been to determine the motivating factors to divulge personal information and the factors inhabiting them in divulging personal information (Son and Kim, 2008). Many constructs have been used in the past, but internet users’ information privacy concerns (IUIPC) have received much attention as a belief that determine their willingness to disclose or not to disclose information (Malhotra et al.,2004; Smith et al.,1996; Stewart and Segars,2002). For instance, Smith et al. (1996) developed a multidimensional concern for information privacy scale (CFIP). It was designed to record individuals’ concerns about organisational practices. CFIP consisted of four dimensions and 15 items. The
four dimensions are: collection, unauthorized secondary use, improper access, and
errors. However, CFIP was used mostly offline or direct marketing context (Malhotra et al., 2004). Stewart and Segars (2002) empirically tested CFIP with 355 participants. Communication was mostly in one direction unlike Malhotra et al.’s
(2004) IUIPC which mostly allowed two way communications.

Recently, Son and Kim (2008) developed taxonomy of information privacy-protective responses (IPPR). IPPR consists of six types of behavioural responses: refusal, misrepresentation, removal, negative word-of-mouth, complaining directly to online companies, and complaining indirectly to third-party organizations. They grouped these six behavioural responses to three items: information provision (refusal and misrepresentation), private action (removal and negative word-of-mouth), and public action (complaining directly to online companies and third party organizations). Subsequently, they developed a nomological model to test IPPR which included three salient antecedents of IPPR: concerns for information privacy, perceived justice, and societal benefits from complaining.

However, as IPPR are only applicable to online situations where users have knowledge about collection of information by companies (e.g., e-commerce websites), therefore, IPPR cannot appropriately be applied to situations where companies collect users’ data without their knowledge. Specifically, in the case of Beacon OSNS were neither aware nor served any prior notice that Beacon is running on these third party sites and will share users action to their profile in Facebook.

Moloney and Bannister (2008) summarized the work of six academic surveys relating to information privacy (see table 2.3). Hui et al. (2007) investigated the effect of monetary compensation on reluctance to disclosed information. They found out that monetary compensation decrease the reluctance to disclose information. This finding is in line with the finding of Hui et al. (2006) which suggests the use of seven types of benefits including monetary benefit to motivate online users to disclose personal information to internet businesses. Dhillon and Moores (2001) findings are the unsolicited emails and identity theft.

<table>
<thead>
<tr>
<th>Summary of Information Privacy Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information privacy Concerns</td>
</tr>
<tr>
<td>Collection of personal information</td>
</tr>
</tbody>
</table>
Unauthorized secondary use of personal information | Milberg et al. (1995)  
| Ackerman et al. (1999)  
| Harris Interactive (2002)  

Improper access to personal information | Smith et al. (1996)  
| Dhillon and Moores (2001)  
| Milberg et al. (1995)  
| Ackerman et al. (1999)  
| Harris Interactive (2002)  

Errors with stored personal information | Smith et al. (1996)  
| Milberg et al. (1995)  

Identification theft through personal information once it has been collected | Dhillon and Moores (2001)  
| Ackerman et al. (1999)  

Receiving unsolicited email. | Dhillon and Moores (2001)  

Monetary incentives decrease reluctance to impart personal information | Hui et al. (2007)  

The more information requested, the less likely the subjects were to disclose it | Hui et al. (2007)  

The purpose for collecting information. | Ackerman et al. (1999)  

| Table 2.3: Summary of findings from six Information Privacy surveys  
(Source: Moloney and Bannister, 2008: p.25)  

Another stream of literature to study privacy is the use of taxonomy. For instance, Solove (2006) takes the same position to solve privacy related issues and argues “to develop a taxonomy that focuses more specifically on the different kinds of activities that impinge upon privacy. [He]endeavours to shift focus away from the vague term "privacy" and towards the specific activities that pose privacy problems. Although various attempts at explicating the meaning of "privacy" have been made, few have attempted to identify privacy problems in a comprehensive and concrete manner” (p.482). One of the other contributions that current research makes is develop taxonomy of privacy concerns of OSNS specifically relating to third party information sharing. As Solove (2006) suggests those individuals, institutions and governments all engage in activities that might be conflicting, thus cause social friction. The privacy is the relief from social friction. For example, in case of SNS, third party organizations who track users (without consent) may be involved in harmful activities (Solove, 2008; Solove, 2006). As there is coincidence of thought in Solove’s (2008, 2006) view of privacy problems and the privacy problems of SNS (the state of privacy can be predicted using privacy taxonomy of Solove (2008, 2006). For instance, privacy concern stalking relate to activity exposure, identity theft with identification and so on. As Solove (2008) states that “the taxonomy is an attempt to
identify and understand the different kinds of socially recognized privacy violations” (p.10).

<table>
<thead>
<tr>
<th>Privacy Taxonomy</th>
<th>Activities creating privacy problems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information collection</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Surveillance</strong></td>
<td>Involuntary gathering of information, mostly without the awareness of subject</td>
</tr>
<tr>
<td><strong>Interrogation</strong></td>
<td>Involuntary gathering of information with conscious awareness</td>
</tr>
<tr>
<td><strong>Aggregation</strong></td>
<td>The gathering together of information about a person.</td>
</tr>
<tr>
<td><strong>Identification</strong></td>
<td>Identification enables us to attempt to verify identity</td>
</tr>
<tr>
<td><strong>Insecurity</strong></td>
<td>Problem caused by the way our information is handled and protected.</td>
</tr>
<tr>
<td><strong>Secondary use</strong></td>
<td>The use of data for purposes unrelated to the purposes for which the data was initially collected without the data subject's consent.</td>
</tr>
<tr>
<td><strong>Exclusion</strong></td>
<td>The failure to provide individuals with notice and input about their records.</td>
</tr>
<tr>
<td><strong>Information processing</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Breach of confidentiality</strong></td>
<td>Is not simply that information has been disclosed, but that the victim has been betrayed</td>
</tr>
<tr>
<td><strong>Disclosure</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Exposure</strong></td>
<td>Involves the exposing to others of certain physical and emotional attributes about a person.</td>
</tr>
<tr>
<td><strong>Increased accessibility</strong></td>
<td>The existence of information in a government database can increase the potential accessibility of that information.</td>
</tr>
<tr>
<td><strong>Blackmail</strong></td>
<td>Consensual sexual relations conducted in private</td>
</tr>
<tr>
<td><strong>Appropriation</strong></td>
<td>use of one's identity or personality for the purposes and goals of another.</td>
</tr>
<tr>
<td><strong>Distortion</strong></td>
<td>Consists of the dissemination of false or misleading information about individuals.</td>
</tr>
<tr>
<td><strong>Information dissemination</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Invasion</strong></td>
<td>Concerns invasive acts that disturb one's tranquillity or solitude.</td>
</tr>
<tr>
<td><strong>Intrusion</strong></td>
<td>Concerns invasive acts that disturb one's tranquillity or solitude.</td>
</tr>
<tr>
<td><strong>Decisional interference</strong></td>
<td>Involves the government's incursion into the data subject's decisions regarding her private affairs</td>
</tr>
</tbody>
</table>

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

2.4 Information Privacy Concerns in OSNS
Whereas the traditional context of privacy existed within a one-to-one environment of information disclosure, online social networks are based on information broadcast principles (Jamal and Cole, 2009). Digital information is characterised by an extremely porous nature whilst the network is designed to support widespread information dissemination (Rosenblum, 2007). Consequently, Dwyer et al. (2007) argue that “privacy within SNS 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.

The researchers, privacy advocates, surveys and the media (electronic and print) have raised privacy concerns associated with the disclosure of too much personal information by SNS users (Acquisti and Gross, 2006; Barnes 2006; Gross and Acquisti, 2005; Krishnamurthy and Wills, 2008; Strater and Lipford, 2008). The issue of privacy concerns (e.g. unauthorised collection and secondary use of user information, and identity theft) becomes more complicated when the social network service providers, in order to raise funds, use or sell user data for targeted advertisements (Jamal and Cole, 2009; Krishnamurthy and Wills, 2008; Preibusch et al., 2007). On the other hand, the general perception about users of SNS is that they do not value privacy as they disclose too much private information about them on their profile (Gross and Acquisti, 2005). However, current media coverage and privacy debate on the internet is changing this perception (Nov and Wattal, 2009; Young and Anabel, 2009). Gurses et al. (2008) found 124 articles on privacy breaches in online social networks between Oct 2007 and March 2008. A study highlight that privacy concerns of users in SNS affect user behaviour such that users try to a) control the disclosure of information and b) adopt more strict privacy controls (Nov and Wattal, 2009). However, Acquisti and Gross (2006) report entirely strange users’ behaviour in SNS when they show concern on the privacy of information, yet they continue to disclose too much personal information on their profiles.

As Etzioni (1999) suggests that the first step to address any privacy issue is to see whether there is a problem or not. However, in the context of OSNS, the situation is more complex and it is not easy to tell (specially in advance) whether a certain action (of a user, OSNS service providers or third party companies) will cause a privacy
breach because a) users exhibit a strange behaviour b) technological platform is built
to support dissemination of information and not to control it, and c) the integration
between third party companies and OSNS and the use of OSNS users’ data. As such
this thesis contributes to our understanding of what factors cause users’ privacy
concerns in OSNS.

However, few privacy concerns have been highlighted from literature as follows which
are only representative of the many of the privacy concerns of OSNS.

i) Identity Theft

Gross and Acquisti (2005) find that as Facebook is bounded with the physical world
(e.g. university), therefore, it facilitates users to disclose real name, image, address,
date of birth, school name, AIM screen name and many other type of information
which is identifiable. Hence, the data identifiability is a significant risk on identity
theft of a person (Krishnamurthy and Wills, 2009a). “Identity theft is the wrongful
obtaining and using of another person’s identifying information in some way that
involves fraud or deception, typically for economic gain” (Hotaling, 2008). It is also
called re-identification (Gross and Acquisti, 2005; Ohm, 2009). Under the new
development of fading of boundary between PII and non-PII — the challenge to
protect user from this threat is enormous and requires multi-fold efforts from users,
policy makers, service providers and designers.

ii) Stalking

According to Gross and Acquisti (2005), Krishnamurthy and Wills (2009a) highlight
the concern of physical stalking from an adversary who can determine the physical
location of user for larger portion of the day. Apart from physical stalking, a user is in
danger of cyber-stalking especially through using AOL instant messenger (AIM).
However, male teenage users seem to manage this privacy concern through providing
wrong information (Lenhart and Maddel, 2007).

iii) Building a Digital Dossier
Users disclose too much personal information on SNS (Gross and Acquisti, 2005) which help SNS as well as third party advertisers (Krishnamurthy and Wills; 2009b) to build digital profile of a users (European Network and Information Security Agency (ENISA) (2007). Also, the unique characteristics of digital information (e.g. persistence, replicability, searchability, invisible audience) (Boyd, 2007) pose a serious threat on user privacy as the information may remain for indefinite time and also can get into wrong hands. As most of that information is identifiable the damage may get worse.

iv) Secondary Privacy Diffusion

Krishnamurthy and Wills (2009a) identify a new privacy issue “secondary privacy diffusion” which has not been raised in privacy literature until recently. This results when privacy person is damaged either intentionally or inadvertently by another user. They explain this privacy damage with the example. “Some Internet services allow customers to provide email addresses of other Internet users so that these other users can be invited to an event or to send copies of restricted online articles to non-subscribers. Event organizing sites host content of interest to the event which can be updated by the invited parties. However, the supplied addresses become known to the service without any prior approval necessarily obtained from these other Internet users resulting in secondary leakage. The relationship between the supplier of the email address and non-subscribers can be stored by the article site. For example, the forwarding of a news article of restricted sites to someone else may give an indication of the recipient's interest or political leanings”(p.549). However, they admitted that with the current available tools, the damage from secondary privacy diffusion cannot be avoided before it occurs. Yet, they offer a notification solution method after that secondary leakage has occurred. According to them, the user is notified when this leakage occurs (if user does not like this leakage of information), the source website is requested not to disclose that information to avoid any further damage. However, their suggested method involves huge cost of monitoring public websites, SNS and blogs to determine that the leakage has occurred. Also, their proposed solution itself is privacy invasive as it requires monitoring of public websites, SNS and blogs.
2.5 Third Party Interaction and Information Privacy Concerns

The interaction between third party companies and OSNS and the huge commercial potential of the use of personal information of OSNS users pose severe threats to OSNS users’ privacy. Privacy concern refers the extent to which an individual is concerned about organisational practices relating to the collection and use of his/her information (Smith et al. 1996). The privacy concern perspective of privacy is the most widely used which is adopted in the context of current thesis. As such, little research has been conducted on privacy concerns of OSNS users which occur as a consequence of the interaction between third party companies and OSNS. Therefore, in order to fill this gap in the current knowledge base, this research is conducted to investigate privacy concerns of OSNS related with the launch of Facebook’ advertising tool Beacon. Beacon study represents a typical case of the third party interaction of OSNS.

The next chapter discusses the research approach adopted in this thesis and chapter 4 details the case study of Beacon.
CHAPTER THREE: Research Methods Approach

3.1 Chapter Overview

The purpose of this chapter is to reflect upon various research approaches used along with their implications on the current research. First section briefly details the research paradigms and methodology together with the choice of appropriate research approach adopted in this thesis. The next section discusses the research methods adopted in this thesis together with the justification. Third section discusses why blogs are used as a data source in this thesis. Finally, thematic analysis as a data analysis method is discussed.

3.2 Research Paradigms and Methodology

The terms paradigm, methodology, method and technique have been used differently by different researchers. A paradigm consists of general set of philosophical assumptions such as ontology (what is assumed to exist), epistemology (what is a valid knowledge - the nature of valid knowledge), ethics or axiology (what is valued or considered right), and methodology (set of activities aim to produce valid and reliable research results) (Mingers, 2001). Also, the distinction between a method and methodology is not very clear. As Mingers (2001) reports, some authors have simply used these terms interchangeably such as Tashakkori and Teddlie (1998) or Livari et al. (1998), whilst, these are also used differently in North America and Europe (Livari et al., 1999).

However, this thesis adopts Mingers’s (2001) view of methodology – a general and less prescriptive than a method and that describes methodology as – a set of guidelines or activities which assist in producing reliable and valid results followed in a certain research paradigm. Whereas a method consists of a set of procedures and techniques used to gather and analyse data (Strauss and Corbin, 1998). Therefore a research methodology often contains various research methods.
Although various categorisations of research methods are used, but the most common distinction is between quantitative and qualitative (Myers, 1997). Myers (1997) suggests quantitative research methods are most suitable in the natural sciences to study natural phenomena. Survey methods, laboratory experiments, formal methods are commonly used quantitative methods. Whereas the qualitative research methods are most suitable in the social sciences to study social and cultural phenomena (Myers, 1997). Action research, case study, and ethnography are the widely used qualitative research methods.

As discussed in chapter 1, this thesis adopted a qualitative research method which allowed the researcher to understand people and the social context within which they live (Myers, 1997). Specifically, it allowed the researcher to study information privacy concerns of OSNS users associated with the launch of Facebook Beacon. Because the quantification of textual data mostly result in the loss of the social and institutional context in which people live (Kaplan and Maxwell, 1994).

However, whatever research method researchers use, they will have to assume what constitute a valid research (Myers, 1997). Three research categories are proposed by Orlikowski and Baroundi (1991) which are: positivist, interpretive and critical.

The underlying assumption behind positivist research is that reality is objectively given and can be measured independent of the researcher and his/her instruments and the objective is to test a theory (Myers, 1997). Example of a positivist qualitative research is that of Yin’s (2002) work on case study.

Whereas the underlying assumption of interpretive research is that knowledge about a reality is gained through language, consciousness and shared meaning (Klein and Myers, 1999). Therefore interpretive research studies allow researchers to understand phenomena by interpreting what people say or believe. This research which is aimed at investigating privacy concerns of OSNS users related with the launch of Beacon follows interpretive research approach which allows the researcher to understand social phenomena of privacy in OSNS. Another motivation of following interpretive research approach is because the objective of interpretive research is to understand the full complexity of human sense making rather than predefining dependent and
independent variables (Kaplan and Maxwell, 1994) which of course is not the objective of the current research. Here the intention is to get deep insights into the complex issue of privacy in OSNS, therefore interpretive research approach seems most appropriate. Klein and Myers (1999) suggested principles for conducting interpretive field studies (including in depth case study and ethnographies) which were adhered to by the researcher in the current research which is also a form of in depth case study of Beacon blogs published as a reaction to the Beacon launch. The researcher argues that the application of these principles helped a lot to conduct the current research. However, the use of inductive approach allowed the researcher to avoid having preconceptions which may be conflicting with the actual findings (from the data). This is also echoed by Glaser and Straus (1967) who warned the use of thorough literature search before conducting research as:

“. . . carefully to cover 'all' the literature before commencing research increases the probability of brutally destroying one's potentialities as a theorist, (p. 253)”.

See table 3.1 for summary of the principles proposed by Klein and Myers (1999) for interpretive field studies. I argue that the principles are quite useful especially for inexperienced qualitative researchers to guide them how to conduct a valid qualitative research by avoiding research biases etc.

<table>
<thead>
<tr>
<th>Summary of Principles for Interpretive Field Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1) The Fundamental Principle of the iterations</strong></td>
</tr>
<tr>
<td>between interdependent parts and the whole that they</td>
</tr>
<tr>
<td>form allow us to understand human meaning. This</td>
</tr>
<tr>
<td>principle of human understanding applies to all other</td>
</tr>
</tbody>
</table>

Page 41 of 88
| Hermeneutic Circle | principles.  
Example: Lee’s (1994: as referred in Klein and Myers, 1999, p.72) study of information richness in e-mail communications. It iterates between the separate message fragments of individual e-mail participants as parts and the global context that determines the full meanings of the separate messages to interpret the message exchange as a whole. |
| 2) The Principle of Contextualization | The social and historical context of the research must be critically reflected so as to allow intended audience to make sense of the situation under study.  
Example: Ciborra et al. (1996: as referred in Klein and Myers, 1999, p.72) after discussing the historical forces that led Fiat to establish a new assembly plant, show how old Fordist production concepts still had a significant impact even after the radical changes in work organisation and operations. |
| 3) The Principle of Interaction Between the Researchers and the Subjects | The interaction between the researchers and participants allows to critically reflect on how the research materials are socially constructed.  
Example: Trauth (1997: as referred in Klein and Myers, 1999, p.72) explains how her understanding improved as she became selfconscious and started to question her own assumptions. |
| 4) The Principle of Abstraction and Generalization | “Requires relating the idiographic details revealed by the data interpretation through the application of principles one and two to theoretical, general concepts that describe the nature of human understanding and social action.  
Example: Monteiro and Hanseth’s (1996: as referred in Klein and Myers, 1999, p.72) findings are discussed in relation to Latour’s actornetwork theory.” |
| 5) The Principle of Dialogical Reasoning | The possible contradiction between preconceived theories which guide the research and the actual findings must be sensitive with subsequent revision iterations.  
Example: Lee (1991: as referred in Klein and Myers, 1999, p.72) describes how Nardulli (1978) came to revise his preconceptions of the role of case load pressure as a central concept in the study of criminal courts several times. |
| 6) The Principle of Multiple Interpretations | Different interpretations of the same story or events expressed by the different participants should be taken care of.  
| 7) The Principle of Suspicion | Possible “biases “and systematic “distortions” in the stories/narratives collected from the participants should be taken care of.  
Example: Forester (1992: as referred in Klein and Myers, 1999, p.72) looks at the facetious figures of speech used by city planning staff to negotiate the problem of data acquisition |

(Table 4.1: Summary of Principles for Interpretive Field Studies)  
(Adapted from: Klein and Myers (1999:p.72))

Finally, the critical research assumes that social reality can be produced and reproduced because it is historically constituted (Myers, 1997). Myers (1997) further
elaborates that critical research focuses on the oppositions, conflicts and contradictions in society.

### 3.3 Case Study Research Method Adopted in this Thesis

After selecting the appropriate research paradigm (in this case interpretive research), the next task of the researcher is to select appropriate research method/s. As Myers (1997) defines a research method “is a strategy of inquiry which moves from the underlying philosophical assumptions to research design and data collection”. Consequently, research method guides the researcher how to collect data. The examples of qualitative research methods include action research, case study research, ethnography and grounded theory.

However, this thesis adopts case study research method because of the complex social nature of the problem i.e. investigating privacy concerns of OSNS relating to launch of Beacon tool is such that the case study research approach was found most appropriate. The case study research is particularly useful when developing a new theory or testing an existing theory (Yin 1994; Benbasat et al., 1987). Therefore, as the aim of this thesis is to investigate privacy concerns of OSNS related with the launch of Beacon which is entirely a new phenomenon and hence case study research is applied to develop a new theory in this thesis (taxonomy of privacy concerns of OSNS users).

Yin (2002) defines case study as an empirical inquiry that:

> “Investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and the context are not clearly evident” (Yin, 2002:p.13).

Because of its suitability to information systems (IS) research (Myers, 1997), it is the most commonly used qualitative research method in IS (Orlikowski and Baroudi, 1991; Alavi and Carlson, 1992). However, case study research may be exploratory, descriptive and explanatory which would depend on the type of research questions such as what, how and why (Yin, 2002). As such the current research falls into an
exploratory type as it aims to answer what type of question (what are the privacy concerns of social network site users related with the launch of Beacon?). Case study research method is highly flexible and hence allows much freedom to the researchers.

3.4 Thematic Analysis as Modes of Analysis

After the researcher has selected appropriate research approach (interpretive in this thesis) and research method (case study in this thesis), the next task is to select an appropriate data analysis method. However, the term “modes of analysis “is preferred over “data analysis “as suggested by Myers (1997) being more appropriate for qualitative research whereas “data analysis” is more appropriate for quantitative research. The modes of analysis are “different approaches to gathering, analysing and interpreting data” (Myers, 1997).

Given the speed of both the users’ reactions and Facebook management almost in the real time, blogs data was collected (Jamal and Cole, 2009). The researcher argues that blogs was the only sources of data available given the speed of users’ reactions and the geographical distribution of users across the world. Also because of the access to information would not have been possible without users’ blogs , which were published as a reaction to the launch of Beacon. Furthermore, the use of blogs data provided a most suitable means of collecting reliable user opinions because blogs offer real-time, unedited user commentaries (Gruhl et al, 2005; Thelwall and Hasler, 2006) and also represent a rich source of qualitative data which is unbiased by a research process (Jones and Alony, 2008).

The researcher argues that in this particular instance of Beacon, blogs was the only source to access users’ opinions particularly those who were affected from Beacon because the stories of their actions performed on other websites were shared with their friends on Facebook without their knowledge. Indeed there was no other way to even learn which users were affected and let’s not think about how to access them for collecting their opinions. Blogs are frequently modified web pages with dated entries in reverse chronological order (Bortree, 2005; Buckingham and Willett, 2006;
Schmidt, 2007). Gruhl et al, (2005) argued that blogs are more reliable source of public opinion insights because they cannot be edited. Methodologically, Thelwall and Hasler (2006) suggest that blogs are a useful means of acquiring a relevant set of opinions or attitudes towards an event or topic. Pikas (2005) highlight the potential of using blogs to gather consumer opinions or comments in the advertising industry. Likewise, Jones and Alony (2008, pp 439) argued that blogs present a “tempting source of qualitative data”. They advocated the benefits of using this novel data source in qualitative research such as “convenience and ease of access to codified data” , “richness and depth of information available”, and “unbiased by research process”(Jones and Alony, 2008: pp439-440).

Additionally, the use of blogs data in text form from the actual subjects (OSNS users) relieved the researcher to transcribe it (which was required in case of audio interview data) that may cause errors during transformation or conversion (Razera et al., 2010) and hence helped the researcher to maintain data quality.

Indeed blogs offer novel and cost effective source to gather data, yet there are few studies which have exploited the use of blogs. To the best knowledge of the researcher as of today, none of the research has used blogs as a means to gather opinions OSNS users regarding their privacy. Therefore this research also contributes in the current knowledge base to exploit the use of blogs as a means to gather opinions of OSNS users regarding their privacy. However, to ensure the privacy of bloggers, no mention of blogger name was made rather only the sequence numbers are used to refer to bloggers. It is also mentioned here that all comments which were also gathered did not have any identifying information since the commentators only use pseudonyms and not their real names. Also, it is confirmed that blogs data will be used only for research purposes and will be destroyed on conclusion of research.

Thematic analysis method is used to analyse blogs data as being most appropriate. Thematic analysis method is used to identify, analyze and report patterns within data (Braun and Clark, 2006). The patterns which emerge from the data are often called themes. However as Braun and Clark (2006) suggest, a theme should capture something important about the research question. Thus not every pattern which
emerges from data becomes a theme; instead it should tightly link with the research question. Thematic analysis can be performed by following one of the two approaches; inductive or deductive. Inductive analysis allows researchers to generate themes directly from the data (Frith and Gleeson, 2004), just like grounded theory approach (Braun and Clark, 2006). Thus, the themes developed are tightly linked to data themselves (Patton, 1990). Whilst the deductive analysis also called theoretical analysis is driven by the theoretical or analytic interest of the researcher.

This research followed inductive analysis approach which is predominately data driven as this allowed the researcher to get avoid any research bias which otherwise could have affected the results because of being bound too much of the theoretical consideration as in the case of deductive analysis.

Among other reasons of using thematic analysis is its flexibility as compared to other methods of analysis such as conversation analysis and interpretative phenomenological analysis which either are tied to or originate from epistemology or theoretical considerations (Braun and Clarke, 2006). Moreover, conversation analysis cannot be adopted in this thesis as a mode of analysis as it deals with the audio or video data unlike the textual data of users’ blogs.

Another dominant motivation to use thematic analysis is because it is a foundational qualitative method which should be used as a first qualitative method by the early researchers (Braun and Clarke, 2006). Therefore recognising the diverse and complex nature of qualitative methods (Holloway and Todres, 2003) and the fact that the researcher did not posses past experience of conducting qualitative data analysis, thematic analysis was adopted in the current thesis.

A six-phased approach to as propounded by Braun and Clarke (2006) was used (see next chapter for detail).
CHAPTER Four: Thematic Analysis of Blogs – A Case Study

4.1 Chapter Overview

This chapter reports the case study of thematic analysis of users blogs published as a reaction to Beacon launch. The first section introduces the Beacon tool. Second section details the research design of the case study which includes discussion of the preparation of data set, selection of data source, selection of search keywords and sources of data collection, and data set generation. Final section elaborates thematic analysis process followed in this research.

4.2 Beacon Introduction

On the 6th November 2007 the hugely popular OSNS, Facebook, launched a new marketing tool called Beacon. Beacon was intended to provide an alternative approach to personalized marketing as a new way of “socially distributing information” (Facebook, 2007). The primary premise was to connect social networks by allowing online businesses such as eBay, Fandango and Travelocity to encourage users to share various actions amongst their friends via automatic news feed. The actions might involve posting an item for sale, purchasing an item such as a cinema ticket or holiday and renting a movie. When such an action is performed on a participating third party website (such as eBay), a Beacon alert occurred informing the user that it is going to automatically share their ‘story’ with their Facebook friends unless the user choose to specifically opt-out of that particular action (Jamal and Cole, 2009).
One month after the launch, Facebook withdrew Beacon because of severe negative user commentary offered via blogs and mirrored in the printed press (Jamal and Cole, 2009).

Below are the three versions of Beacon alert which were served when users performed actions on third party websites.

![Figure 4.1 Early Beacon Warning](image)

Figure 4.1 is an early Beacon warning which indeed was elusive since there was no prior notice given to user about what is Beacon and what it is going to do. Also, there was no choice to accept or reject the offer of sending story to user profile unless user click the button ‘See More’ which prompted figure 4.2 to the user.

![Figure 4.2 Beacon Interface When Users Clicked ‘See More’](image)

However, the Beacon interface yet fails to provide an explicit choice to accept the offer for sending a story to Facebook user profile, which was introduced in the final Beacon interface in figure 4.3. Clearly, Beacon interface violated many of the Nielson’s (1990) popular heuristics including ‘user control and freedom’ because of the limited choices it offered to users (Jamal and Cole, 2009b).
4.3 Research Design

Thematic analysis study is divided into two complementary steps i) preparation of data set and ii) thematic analysis process. Figure 3.1 depicts the study process well. Braun and Clark (2006) suggested that the thematic analysis process assumes a “good quality” data corpus and data set which should define criteria regarding what, why and how data was collected. Consequently, the data preparation steps involved various checks explained below.

I. Preparation of Data Set

Preparation of data set involved elementary steps such as selection of data source, defining search keywords, choosing sources to gather data, and applying several checks to optimise the search. Due care was exercised in the execution of these steps which ensured “data quality” as proposed by Braun and Clark (2006). Thelwall and Hasler (2006) also highlighted the need for “an appropriate blog search to yield a set of relevant posting[s]”.

II. Selection of Data Source

Given the speed of reaction – both from the users and Facebook management, blog data was used as the means to collect an accurate representation of real-time user opinion. The researcher argues that in this particular instance of Beacon, blogs was the only source to access users’ opinions particularly those who were affected from Beacon because the stories of their actions performed on other websites were shared with their friends on Facebook without their knowledge. Indeed there was no other
way to even learn which users were affected and let’s not think about how to access them for collecting their opinions. Blogs are frequently modified web pages with dated entries in reverse chronological order (Bortree, 2005; Buckingham and Willett, 2006; Schmidt, 2007). Gruhl et al, (2005) argued that blogs are more reliable source of public opinion insights because they cannot be edited. Methodologically, Thelwall and Hasler (2006) suggest that blogs are a useful means of acquiring a relevant set of opinions or attitudes towards an event or topic. Pikas (2005) highlight the potential of using blogs to gather consumer opinions or comments in the advertising industry. Likewise, Jones and Alony (2008, pp 439) argued that blogs present a “tempting source of qualitative data”. They advocated the benefits of using this novel data source in qualitative research such as “convenience and ease of access to codified data”, “richness and depth of information available”, and “unbiased by research process” (Jones and Alony, 2008: pp439-440).

III. Selection of Search Keywords and Sources of Data Collection

Due care was exercised in the selection of search keywords and choosing sources to collect blogs so that relevant blogs can be gathered. Keywords were defined to search and collect most relevant blogs. Specifically, three keywords were defined: “Facebook”, “Beacon” and “Privacy” and their combinations resulted in three search strings: “Facebook Privacy”; “Facebook Beacon”; “Beacon Privacy”, which were finally used to search relevant blogs. Noticeably, all three search strings have either a word “Facebook” or “Beacon”. This was done to restrict the result to only blogs related to Facebook Beacon. In order to optimize our search to blogs only, we used blog search engines and media sites so that the search results could be restricted to blogs only. However, due care was exercised in the selection of these blogs related search engines and media sites so that wide coverage of blogs on the topic should be gathered.

The blogs were gathered using blog search engines namely: Google Blog Search (http://blogsearch.google.com/), Technorati (http://technorati.com/), 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 (http://www.sociableblog.com/). The blogs published between 6th November, 2007 (date of the launch of Beacon) and 28th February, 2008 were included as the user commentary has dwindled on most of the blogs beyond that date. Only the blogs returned on first three pages of blog search engines were collected to restrict to the most relevant blogs.

IV. Data Set Generation

Initially 95 blogs were collected containing 568 data opinions which were saved into a text file called data corpus. The data corpus contained 234 A4 size pages of text. Subsequently blogs were checked against redundancy through same title or same author information. Consequently data corpus was updated. However, relevancy of blogs could not be checked unless the researcher had read them. Therefore to establish relevancy of blogs, the researcher skimmed each blogs to ensure it discussed privacy leakage of Facebook users by Beacon. Additionally to allow for more coverage of user opinions, the blogs with at least three comments/ opinions were considered. Therefore the final set of blogs meeting both the relevancy and more coverage of opinions criteria were saved into a text file called data set. The final data set contained 29 blogs which was (31%) of the original 95 blogs. However, the final data set still contained 492 data comments and represented 87% of the original total 568 comments. The data set finally contained 159 A4 size pages of text (word file available on request).
Figure 4.4 Data preparation and thematic analysis process
4.4 Thematic Analysis Process

Thematic analysis method is used to identify, analyze and report patterns within data (Braun and Clark, 2006). The patterns which emerge from the data are often called themes. However as Braun and Clark (2006) suggest, a theme should capture something important about the research question. Thus not every pattern which emerges from data becomes a theme; instead it should tightly link with the research question.

This research followed inductive analysis approach which is predominately data driven as this allowed the researcher to get familiarised with the phenomenon of third party privacy leakage in OSNS. The researcher argues this would have not been possible in the other theoretical analysis approach and that too in very early stage of the research.

The blogs were analysed using the 6-phased approach to thematic analysis proposed by Braun and Clarke (2006). Figure 4.4 depicts all six phases of thematic analysis process. A brief description of six phases of thematic analysis process is presented in table 4.1.

QSR Nvivo version 8- a qualitative data analysis tool was used to better manage thematic analysis process. The choice of Nvivo was made because of its capability to support all six steps in the thematic analysis process. Consequently the data set text file was loaded in Nvivo to begin the analysis.

Thematic analysis began with the researcher reading and re-reading the blogs data to get familiarised with the data set. Although the reading and re-reading process was laborious and time taking given the volume of text data (159 A4 size pages in this case), yet it was enormously helpful for the researcher to a) get familiarised with the phenomenon of third party privacy leakage in OSNS and b) to understand the nature of privacy in OSNS. Thus the benefits of this process were far more over reaching given the early stage of the research. Additionally, as Braun and Clarke (2006) suggested this repeated and active reading (reading with a purpose to search for
patterns or themes) provides the bedrock for the whole analysis process. Therefore due care was exercised to this first phase of thematic analysis. The 6-phased framework which adopted an iterative approach also helped the cause.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description of the process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <strong>Familiarizing yourself with your blogs data:</strong></td>
<td></td>
</tr>
<tr>
<td>Reading and re-reading the blogs data, noting down initial ideas.</td>
<td></td>
</tr>
<tr>
<td>2. <strong>Generating initial codes:</strong></td>
<td></td>
</tr>
<tr>
<td>Coding interesting features of blogs data in a systematic fashion across the entire data set, collating blogs data into each code.</td>
<td></td>
</tr>
<tr>
<td>3. <strong>Searching for themes:</strong></td>
<td></td>
</tr>
<tr>
<td>Collating codes into potential themes, gathering all blogs data relevant to each potential theme.</td>
<td></td>
</tr>
<tr>
<td>4. <strong>Reviewing themes:</strong></td>
<td></td>
</tr>
<tr>
<td>Checking if the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic ‘map’ of the analysis.</td>
<td></td>
</tr>
<tr>
<td>5. <strong>Defining and naming themes:</strong></td>
<td></td>
</tr>
<tr>
<td>Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.</td>
<td></td>
</tr>
<tr>
<td>6. <strong>Producing the report:</strong></td>
<td></td>
</tr>
<tr>
<td>The final opportunity for analysis. Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.</td>
<td></td>
</tr>
</tbody>
</table>

*Table 4.1 Phases of Thematic Analysis (Adapted from Braun and Clark, 2006)*
After getting familiarised with the data set, the researcher started coding the data. 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 OSNS users associated with the third party privacy leakage caused by the Beacon. Consequently, in the first iteration, 176 codes were generated which were grouped into 64 themes. However, 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 focus. In the second iteration the codes were checked to see if they corresponded to the coded extracts and the entire data set and whether the grouping sounds logical. Additionally, all codes were checked that they are not redundant to remove the duplicate ones. The second iteration also gave the researcher an opportunity to read entire data set again to see there was no omission of the important patterns in the data set. In the third and final iteration all the themes were reviewed and themes with diverse data were split as new themes whilst themes which were coded on related data were merged. At this stage, all the themes represented separate data and could not be further refined or eliminated. The final list consisted of 48 themes which are represented as a thematic map in the next chapter.

Given the nature and scope of this exploratory case study, however, all privacy concerns were coded and grouped into main themes rather than focussing on some particular themes which was not the objective of this study. Consequently, the thematic map represented taxonomy of privacy concerns of OSNS users associated with third party leakage of information. So far to the best knowledge of the researcher, this taxonomy is the first which group privacy concerns of OSNS users related with third party.
CHAPTER Five: Empirical Findings and Analysis

5.1 Chapter Overview

This chapter reports the empirical findings of this research and their critical analysis.

5.2 Empirical Findings and Analysis

As discussed in previous chapter, inductive method of coding was followed to generate themes from the data (159 A4 size pages loaded into Nvivo 8 qualitative data analysis tool). Also, the codes were generated and grouped as themes in an iterative process (three iterations were made because until then all redundant codes had been merged and no further redundancy was found). Thematic analysis process resulted in the generation of final list of 48 privacy themes (privacy concerns which was the focus of this study. Thematic map of privacy concerns of OSNS users is generated (see figure 5.2) which represented a sort of taxonomy of privacy concerns.

Privacy themes were arranged in three levels i.e. from general to specific. By referring Figure 5.2, the analysis showed 5 main themes: commercialism, Terms of service (ToS), lack of information control, lack of user awareness and data protection.
Figure 5.1 also represent count of major privacy themes. The 5 core themes were also quantified using a simple frequency count. All the occurrences of the major themes were added as well as their corresponding sub-themes.
Lack of user control, data protection and commercialism received the majority of counts (79, 53 and 38 respectively). Lack of user awareness and ToS though received fewer counts as compared to lack of user control, data protection and commercialism, yet they were significant enough to warrant attention because they captured privacy concerns of OSNS users and contributed to the understanding of the overall research.
question. As Braun and Clarke (2006) suggested that quantifying themes does not establish the importance of a theme, which rather depends more on how the theme contributes to answer the overall question. Therefore the purpose of simple frequency count was not to establish the importance of a theme but rather to give additional information to the reader.

The sub-themes represented in the thematic map provided greater clarity regarding the nature and form of the core themes. Therefore, third party paid advertisement and the selling of personal information were identified and related to commercialism. Table 5.1 represents sample text extracts taken from blogs together with the code. Also following text extracts taken from blog data helped the researcher to develop themes of around privacy concerns of OSNS. These extracts have been published in (Jamal and Cole, 2009). For example, blogger [8] commented:

“I think that this is a wonderful example of advertisers going too far and actually discouraging people from using their products”.

Whilst blogger [13] stated:

“I just don’t see how Beacon benefits me as a user. Facebook has cash on hand; they should be focusing 100% on how to benefit users rather than how to monetize them”.

Perhaps the users were found performing a sort of risk analysis at this occasion to evaluate the benefits of information transfer with its costs. This perspective has been advocated well in privacy calculus approach which proposes that users perceive benefits of information disclosures to offset the risks associated with it (Culnan and Armstrong, 1999; Dinev and Hart, 2006). The privacy calculus perspective is “the most useful framework for analyzing contemporary consumer privacy concerns” (Culnan and Bies, 2003: p.326). Viewing privacy as risk analysis in online social networks allows us to interpret individual user’s privacy interests in terms of an exchange where users disclose personal information in return for some benefits (Xu et al., 2010).
Moreover, lengthy, obscure, autocratic and irrevocable rules which users agreed to abide by in order to use Facebook were grouped under the generic heading of ToS (Jamal and Cole, 2009). This was briefly echoed by commentator [7] on blog [1] saying:

“Facebook's Terms of Service are long, legalistic, onerous, and absurdly overreaching and self-serving”.

The general inability of the user to determine how their personal information was used by third parties (people and organisations) was labelled ‘lack of control’. Commentator [26] on blog [7] stated:

“The recent moves... into advertising, [mean] that other users can post information about you, be it true or false, or in or out of context. That means that if I have a profile but am not an active user, and an old friend posts an incriminating picture from years ago, it will show up as a photo link when other users visit my page, without my knowing it”.

Additionally, the commentator [16] on blog [7] noted that:

“You can't stop your Facebook friends from tagging you in inappropriate photos, and you can't stop them from posting inappropriate things on your "wall"”.

These concerns were precisely summarised by the following comment:

“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” (blogger [9]).
Besides the obvious lack of respect for privacy, the Beacon feature seems to be promoting a level of consumerism that I would hope very few people support, I can’t think of anyone who judges their friends by the shoes they buy”
“What give them the right to distribute and monetize my activities”
“I think that this is a wonderful example of advertisers going too far and actually discouraging people from using their products”.
“I just don’t see how Beacon benefits me as a user. Facebook has cash on hand; they should be focusing 100% on how to benefit users rather than how to monetize them”.

“Has Facebook been signing agreements with online commerce companies so that whenever I make any sort of online purchase -- or sign up for anything, or just do anything -- it’ll show up on my Facebook page as advertising?”
“They will probably take another page from the Google book and use Facebookers information to help advertisers without “proactively notifying users.”

“Facebook’s Terms of Service are long, legalistic, onerous, and absurdly overreaching and self-serving”.

“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.”

“Also alarming is the apparent cross-pollination of information between Facebook and Fandango”.

“Many, many social networking sites might allow you to take information off of your profile, but you cannot delete your account–esp. if it’s a free account.”

“The recent moves to expand its reach beyond college campuses and into advertising, is that other users can post information about you, be it true or false, or in or out of context. That means that if I have a profile but am not an active user, and an old friend posts an incriminating picture from years ago, it will show up as a photo link when other users visit my page, without my knowing it”.

“You can't stop your Facebook friends from tagging you in inappropriate photos, and you can't stop them from posting inappropriate things on your “wall”. You can delete the tags, and you can delete the posts, but obviously there is a time delay problem to delete them.

(\textit{Table 5.1: Examples of Text Extracts and related codes})

The theme “lack of control” stands in contrast to the theme ‘lack of awareness’ which highlighted that users were not informed about the nature and consequences of the Beacon as it related to their SNS information behaviour. For example blogger [1] commented:

\textit{“Has Facebook been signing agreements with online commerce companies so that whenever I make any sort of online purchase -- or sign up for anything, or just do anything -- it'll show up on my Facebook page as advertising?”}
The final theme labelled as “data protection” is related with the legal aspects of data collection, storage and dissemination of personal information. The blogger [4] commented:

“Many, many social networking sites might allow you to take information off of your profile, but you cannot delete your account—esp. if it’s a free account.”

Also, blogger [1] remarked:

“Also alarming is the apparent cross-pollination of information between Facebook and Fandango”

As such the theme “data protection” resembles to widely referred concerns for information privacy (CFIP) scale developed by Smith et al. (1996) to represent individuals privacy concerns against organizational practices (Malhotra et al.,2004). CFIP covers four aspects of data: collection, unauthorized secondary use, improper access, and errors. Although these dimensions of privacy were developed keeping in view the offline environment, but these also apply in online setting. The subtheme “cross pollination of information” is surprising and it specifically relates to OSNS. This is evident from the automatic transfer of information between OSNS and third party companies via Beacon.
CHAPTER SIX: Discussion, Conclusions, Limitations and Future Recommendations

6.1 Chapter Overview

This chapter begins with the discussion of results of thematic analysis of blogs. Then, the conclusion is reported which briefly summarises the research overview, accomplishment of the aims and objectives of research, major findings of research, and research contributions and originality. Finally, discussion of the limitations of the current research and recommendations for future research are made.

6.2 Discussion

The results of thematic analysis corroborate the findings of a great deal of the previous work in privacy (Jamal and Cole, 2009). The theme lack of control fits well with the work of Westin (1967) and Malhotra et al. (2004) who found that self-regulation is an important determinant of privacy. Westin (1967) defined privacy as “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others”. This perspective was applied by Malhotra et al. (2004) in an online setting. They argued that user control is an important component of online privacy which can be exercised via approval, modification, and choice to opt-in or opt-out. Beacon did not give users control on how, when and to what extent their information is communicated with others. Moreover, users were neither consented prior to the automatic transfer of their personal information from third party company website to Facebook newsfeeds nor they were given choices to alter their decision to opt-in or opt-out of Beacon.

Therefore this research study suggests that the designers and administrators of OSNS should consider control as an important factor that might affect the future growth of OSNS user base and thus may result in loss of business value accumulated by a SNS (Jamal and Cole, 2009).
The next two core themes (data protection and commercialism) reflect the thorny issue of who owns the information posted on a social network. Because users on the one hand want to protect their information against corporate use so in order to maintain their confidentiality. However, the findings also highlight the issue of rewards offered to users in return for corporate access and use of personal information (Jamal and Cole, 2009). To what extent it is reasonable to expect people using free OSNS while sharing personal information on a voluntary basis services not to be sources of personalised advertising? This suggests that the value of privacy needs to be radically redefined to permit free commercialisation of personal data – but within certain legally defined and society-agreed contexts (Jamal and Cole, 2009). Perhaps, users seem to perform risk analysis at this occasion to evaluate the benefits and the costs of information use to third parties. This perspective has been well advocated in privacy calculus approach -that user view benefits of information disclosures off set the risks associated with it (Culnan and Armstrong, 1999; Dinev and Hart, 2006). The privacy calculus perspective is the “the most useful framework for analyzing contemporary consumer privacy concerns” (Culnan and Bies, 2003: p.326). Viewing privacy as risk analysis in online social networks allows us to interpret individual user’s privacy interests in terms of an exchange where users disclose personal information in return for some benefits (Xu et al., 2010).

However, users require a greater degree of education and awareness about their information behaviour to see if this approach is acceptable. Because the visibility, purpose and presence of Beacon was not made clear to the user prior to browsing actions, the integrity of the third party organizations to safeguard the interests of end users was questioned (Jamal and Cole,2009). Perhaps, the users were not sure if their individual interests would be protected when their personal browsing habits and purchases were automatically distributed across their ‘friends’ network. Users perceived this transfer unfair which calls for an approach to involve justice framework to evaluate the fairness of the information transfer practices of companies. Justice framework has recently gained importance to study privacy phenomenon in an online setting where users perform risks and benefits analysis before they share personal information (Xu et al., 2010).
Therefore organisational misuse of users’ information both by the OSNS and the third party companies and malicious user activity is unclear. It may be that this is an on-going balancing act between self-regulation and legal protection that depends on the context of information use and requires a maturation of online information behaviour (Jamal and Cole, 2009). This means that privacy calculus approach seems appropriate where by users evaluate the risks and benefits of information transfer to third parties. In so doing, users may evaluate privacy self regulation by social network site providers.

6.3 Conclusions

6.3.1 Research Overview

This research aimed at investigating the privacy concerns of online social networks users associated with the interaction between third party companies and the OSNS. Research followed interpretive based qualitative case study approach to accomplish the research aims objectives by conducting thematic analysis of users’ blogs which were published as a reaction to the launch of Facebook’s personalised advertising tool Beacon. Beacon was launched on 06 November 2007 by Facebook to leverage social networks to use it by third party companies (45 partner companies including eBay, Fandango and Travelocity) for personalised marketing. However, Beacon was severely criticised by users which was mirrored in the press soon after its launch. Consequently, Beacon was withdrawn by Facebook one month after its launch whilst Facebook CEO, Mark Zuckerberg publically admitted that they have made a mistake and apologised users (Facebook, 2007). This study, therefore, investigated what are the privacy concerns of OSNS related with the launch of Beacon. Blogs data was collected due to the real time reaction both from the users as well as Facebook management and a thematic analysis of blogs was conducted which offered understanding to the nature and form of privacy concerns of OSNS related specifically with the launch of Beacon. In so doing, this case study contributed in the development of taxonomy of privacy concerns of OSNS users specifically related with the third party information use. Noticeably, the concerns such as commercialism,
terms of service (TOS), lack of user control, lack of user awareness and data protection influence user perceptions of online privacy. Beacon was an interesting study as it helped the designers and service providers of online social networks to understand the nature and form of privacy concerns of OSNS related with the failed launch of Beacon and learn lessons from this case study.

6.3.2 Accomplishment of Research Aims and Objectives

The aim of current research was to investigate the privacy concerns of online social network sites related with the sharing and use of personal information of OSNS users by third party companies and OSNS. Research objective was to develop taxonomy of privacy concerns of OSNS which will enable the designer, service providers and policy makers to understand the nature and form of online privacy in social networks.

Interpretive approach followed in this research helped the researcher to make sense of the meaning people assign to words referred in blogs data collected. Thematic analysis of blogs data collected from 95 blogs (which were finally refined for redundancy etc to 29 blogs containing 468 users’ comments) enabled the researcher to represent privacy concerns in the form of taxonomy arranged in three levels of privacy concerns- from more general to specific. The researcher argues that the developed taxonomy of privacy concerns is a good starting point for the designers and service providers of OSNS and the policy makers to consider valuing privacy of personal information of OSNS users (currently more than 500 million users use OSNS). On the hand the analysis highlighted the thorny issue of information ownership whereby on the one hand OSNS users willingly disclose their personal information (even identifiable information such as name, phone number, date of birth, email address and others) and use OSNS services free and expect that neither third party companies nor OSNS should use their information for personalised advertising. The finding of rewards attached with the use of personal information by third party companies and OSNS offers interesting debate.
As such the results of this research accomplished its aims and objectives to understand the form and nature of privacy concerns of OSNS users and in doing so result in the development of privacy taxonomy specifically related with OSNS.

6.3.3 Research Findings
The case study of Facebook’s personalised advertising tool Beacon offers interesting findings as it helped to improve our understanding of the form and nature of online privacy in social network sites specifically related with the use and sharing of personal information of OSNS users by OSNS and the third party companies. Specifically, the findings suggest issues such as commercialism, lack of control, lack of awareness (of information use and sharing by third party companies, terms of service and data protection (Jamal and Cole, 2009). These privacy concerns are arranged in three levels of granularity (see figure 3.2) – starting from the general (left side) to more specific (right side). Together, 48 privacy concerns are reported which represents a sort of taxonomy of privacy concerns of OSNS users. Whilst most of these privacy concerns corroborate well from the literature, some privacy concerns, however, are surprising (new in the context of OSNS) such as behavioural advertising and cross pollination of information between different sites. Additionally, the research also highlighted the complexity of information ownership and management in OSNS. Finally, the research findings are represented as taxonomy of privacy concerns of OSN users.

6.3.4 Research Contributions and Originality
The case study of the thematic analysis of user blogs published as a consequence to the launch of Beacon contributes to our understanding of the form and nature of online privacy. The critical analysis resulted in the development of a thematic map which group privacy concerns in three levels of granularity- from more general (e.g. commercialism ) to more specific (e.g. behavioural advertising). Therefore the result represents a sort of taxonomy of privacy concerns of OSNS users which occur as a consequence of the third party use of personal information of OSNS users either.
Beacon is an interesting case study because it highlighted the privacy concerns of OSNS users specifically related with third party use of information. The developed taxonomy of privacy concerns, being the first in OSNS, may contributes to the understanding of the designers that the applications (such as Beacon) which use personal information of OSNS without their knowledge, consent and choice of opt-in should be build such that these concerns are properly taken care of by the designers. For example, as most of users show concern of the lack of universal opt-out of Beacon, the designers should learn lesson to develop such application by considering privacy concerns of OSNS users. Likewise, the study contributes to the understanding of OSNS service providers to appreciate the concerns of OSNS by providing better control and better transparency of the use of personal information of OSNS users. Not of least importance of this study is for policy makers to think how to better protect consumers’ privacy by framing either new laws or introducing self regulation specifically in OSNS. The study also contributes to enhance awareness of third party use of personal information among the users of social network sites and help them educate on how to better protect their privacy by either utilising more stricter privacy controls or by adopting appropriate responses such as limited disclosure of personal information on profile, falsifying information and complaining directly to the online social networks service providers.

Finally the research also contributes in the use of research methodology followed in this research. For example, the study is first of its kind which has critically analysed users’ blogs by using thematic analysis.

Given the real time speed of reactions both from the users as well as Facebook management, the study critically analysed blogs when no such study existed which has empirically evaluated privacy problems of Beacon at that time, thus the present study addressed this gap in the current knowledge base and hence is original.

**6.4 Research Limitations**

Although this research study offered many benefits to the designers, service
providers and users of social networks as well as law makers but this study is not without any limitations. One limitation of this research is use of qualitative case study research approach and hence the findings cannot be generalised. Also, the study could have been benefitted from other methods of data collection such as in depth interviews (which could not have been done perhaps due to access problem).

6.5 Recommendations for Future Research

The exploratory nature of the current research helped identify issues such as commercialism of personal information of social network sites, behavioural advertising and information ownership, however, in-depth analysis and study of these issues is recommended for future research. Also, to generalise the findings of this research, future research is recommended which could benefit from development of hypotheses and survey items from the current findings and then empirical testing of the hypotheses. As the current research only investigated privacy concerns of social network users related with the launch of Beacon which represents a case of third party information leakage, therefore future research is required to understand phenomenon of third party privacy leakage as suggested by Krishnamurthy and Wills (2010).
References


URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5454790&isnumber=5454428


Smith, S., (2008). The Rising Price of Privacy. EContent; Apr 2008; 31, 3; ABI/INFORM Research pg. 18


Young, A.L. and Anabel Quan-Haase (2009). Information revelation and internet privacy concerns on social network sites: a case study of Facebook. Proceedings of the fourth international conference on Communities and technologies, University Park, PA, USA.

Appendix A (Initial List of Codes)

<table>
<thead>
<tr>
<th>Sno</th>
<th>Name</th>
<th>References (count)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A false sense of Anonymity</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Ability to aggregate data</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>ability to control the data</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Ads are going to feel like content</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Ads Misleadingly integrated with main pages</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>advertising</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>alert a user’s friends to their latest purchases.</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>animated ads cover up content</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>automatic grant of irrevocable, transferrable right to use, manipulate and distribute information and prepare derivative works</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>automatic sharing of purchases information to friends</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>automatically send invitations</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>availability of choice to reveal information</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>Awareness of the risky nature of digital information is vital</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>Bad javascript; animated ads cover up content,</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>Bail out of Advertisers</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>Behavior is hard to hide</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>behavioral targeting,</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>Better user experience as a compensation to giving up information</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>Big brother problem</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>Broadcasting data collected from outside sites to friends and acquaintances</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>Careful disclosure of information on profile</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>change of operational agreement without notice</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>CIA and facebook connections</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>collect information about you from other sources, such as newspapers, blogs, instant messaging services, and other users of the Facebook</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>concern over knowledge of online activities by friends</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Concern over sharing of personal and shopping information</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Consent to share information</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>consumerism or commercialism</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>control over their identities</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>cookie-driven market research</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>costs outweigh benefits</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>cross-pollination of information</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Data Collection</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Data Protection</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Deletion of information on advertisement sites in case of user opt out</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>deletion of profile information is not easy</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Design team didn't know about effective adv</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>developers thinking in single-user mode</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>difficult to opt out</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Disclosure of information on news feed of friends' of friends</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Disclosure of relationship information in news feed</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Display of Tagged photos in friend's profile</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Dissemination of information of users to large number of people</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Do not publish information sharing of which is not safe</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Don't want to Public my personal data such as purchases etc to all friends</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>embarrassing disclosures.</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Embarrassing level of transparency</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Erase data if user opts-out</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Exploiting kids information</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Exposing personal choices</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Fine tuning control over information</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>flow of data from an external site to Facebook</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>focused ads</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>giving away your privacy for the profit of others</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>grant of sublicenses to use, manipulate and distribute information</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>Hard to understand that privacy is not a binary property</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>Hidden opt out</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>huge Privacy Policy</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>identity theft</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>Implications of becoming a fan of a product</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>Implications of publishing user pictures with Advertisements</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>Inadvertant sharing of private information with work colleagues</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>Inappropriate defaults</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>incentive for a user to lend their identity</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>Incompetence to handle user data</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>Inform users as to the implications of becoming a &quot;fan~</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>Information collection from other sources</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>Information ownership</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>Irrational desire for control by decision makers</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Irrevocable right to use, copy, perform, display, reformat, translate and distribute information and content</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>Issues offer a new kind of design problems</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>Lack of a capability to put absolute filters on the news feed</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>Lack of awareness how to have control over privacy options</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>Lack of awareness of what information Facebook is extracting from users</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>Lack of information or guidance on deletion of Facebook account</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>Law to delete personal information</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>Law which prohibits keeping of personal information from SNS when we stop using it</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>Life time contract</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Limited use of privacy settings</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>Long, legalistics, onerous, absurdly overreaching and self-serving Terms of Service</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>Look at your buying history</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Make it opt-in</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>Misplaced advertising</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>Mis-use of profile information by Employers</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>More granular opt-out mechanism</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>Move control to publish or not</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>Multi-site registrations</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>Need for proactive online Privacy education</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>Need Privacy laws</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>No Absolute Privacy</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>No awareness of Program running</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>No choice to share information</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>No compensation for Ad referrals</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>No delete option to delete public information</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>No link between desire of privacy and the risk of information being exploited</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>No notice of compensation will not be paid for Ad referrals</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>No notice to share information and of new feature in a site</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>No opting out of advertising</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>No opt-out</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>No permanent deletion of account and data</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>No universal opt out</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>No user control</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>No way to control flow of information</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>Notice to the user should be sent before any information is disseminated</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>Online archaeology of online lives</td>
<td></td>
</tr>
<tr>
<td>106</td>
<td>Opening membership to all</td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>Opportunity to sell information to advertisers</td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>Opt out of Sponsored Ads</td>
<td></td>
</tr>
</tbody>
</table>
Organizations see privacy as a legal compliance problem.

Out of context disclosure of your information by others

Out of context use of information

Overestimate the benefits of their service as compared to the privacy violation

People have privacy boundaries

Perceived safety at the cost of liberty

Permanency of digital information is risky

Perpetual right to use, copy, perform, display, excerpt and distribute such information

Personality is hard to hide

Posting of Buying History

Posting of inappropriate information on your wall by friends

Posting of your information by others

Privacy as a cost of using free services of Facebook, fun and pleasure

Privacy as a legal compliance problem

Privacy is asymmetrical

Privacy is the cost of using internet

Privacy settings should be prominent

Public purchasing habits

Publish information without permission

Push marketing

Retention of copy of information even on deactivation of account

Retention of data when user opts-out

Risk of disclosure of sensitive information at work

Sell information

Shambolic product tarnishing user reputation

Sharing of personal information with strangers by children

Site feedback or bugs report

Social responsibility of SNS to protect privacy

Social stalking

Spam

Tagging of inappropriate photos by friends

The need to educate youth about privacy

Third party access of information

Third party advertisers

Third party applications are hidden

Third party use of information

too much disclosure of personal information by users on their profiles

Too much friends requests, excessive celebrity

total transparency

total user control

Track activity outside the site

Tracking and recording by cookies

Tracking the websites Internet user visits

Tracks activity
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>153</td>
<td>transmit data to third parties</td>
</tr>
<tr>
<td>154</td>
<td>trespassing on personal space.</td>
</tr>
<tr>
<td>155</td>
<td>trouble predicting what will cause privacy outrage</td>
</tr>
<tr>
<td>156</td>
<td>trust and privacy</td>
</tr>
<tr>
<td>157</td>
<td>trusted referral</td>
</tr>
<tr>
<td>158</td>
<td>tweaked the language and the behavior of interface elements.</td>
</tr>
<tr>
<td>159</td>
<td>un-authorized storage of information</td>
</tr>
<tr>
<td>160</td>
<td>undervaluing emotional factors</td>
</tr>
<tr>
<td>161</td>
<td>unsolicited ads</td>
</tr>
<tr>
<td>162</td>
<td>use by predators</td>
</tr>
<tr>
<td>163</td>
<td>use of alternative network</td>
</tr>
<tr>
<td>164</td>
<td>use of information by employers</td>
</tr>
<tr>
<td>165</td>
<td>use of information by outsider applications</td>
</tr>
<tr>
<td>166</td>
<td>use of obsolete information</td>
</tr>
<tr>
<td>167</td>
<td>use of privacy settings to manage privacy of information</td>
</tr>
<tr>
<td>168</td>
<td>user and designers awareness of privacy bugs</td>
</tr>
<tr>
<td>169</td>
<td>user information storage on servers without their knowledge</td>
</tr>
<tr>
<td>170</td>
<td>user-hostile</td>
</tr>
<tr>
<td>171</td>
<td>users love the new content and receptive to display offer</td>
</tr>
<tr>
<td>172</td>
<td>users receive a cut of their Advertising Revenue</td>
</tr>
<tr>
<td>173</td>
<td>users’ response</td>
</tr>
<tr>
<td>174</td>
<td>Valuation due to demographical information of users</td>
</tr>
<tr>
<td>175</td>
<td>without “proactively notifying users.</td>
</tr>
<tr>
<td>176</td>
<td>without consent.</td>
</tr>
</tbody>
</table>